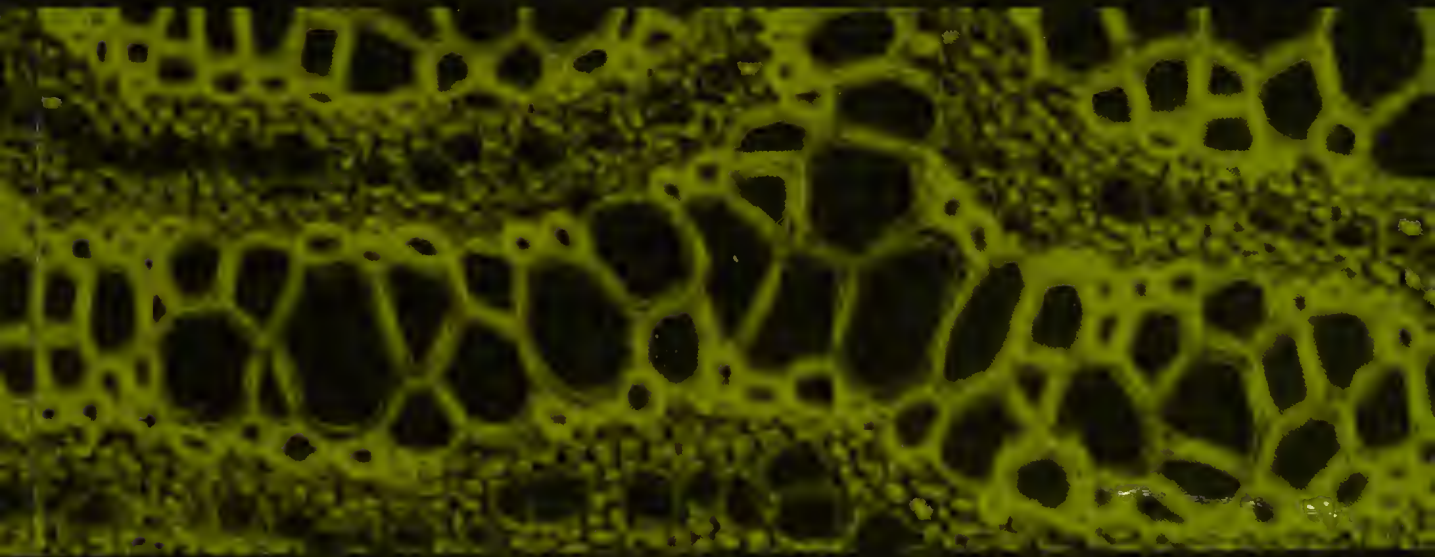


OXFORD

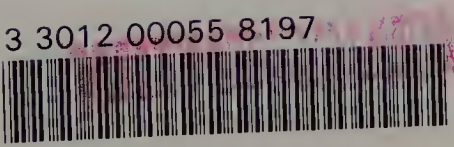
Plant Biotechnology

The genetic manipulation of plants



Adrian Slater, Nigel Scott,
and Mark Fowler

Plant Biotechnology



Plant Biotechnology

The Genetic Manipulation of Plants

Adrian Slater,

Nigel W. Scott

and

Mark R. Fowler

De Montfort University

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Adrian Slater, Nigel W. Scott and Mark R. Fowler 2003

The moral rights of the author have been asserted
Database right Oxford University Press (maker)

First published 2003

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organizations. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer

A catalogue record of this book is available from the British Library
Library of Congress Cataloging in Publication Data

Data available

ISBN 0 19 925468 0

Typeset by SNP Best-set Typesetter Ltd., Hong Kong
Printed in Great Britain
by
Biddles Ltd., Guildford and King's Lynn

Preface

Plant biotechnology has made tremendous progress in recent years and has ‘enjoyed’ a previously unknown level of public awareness. Unfortunately, much of this awareness has arisen from the negative publicity that surrounds GM crops. One only has to think of the media coverage of food safety issues (‘Franken Foods’) or environmental concerns (the Monarch butterfly affair) to gain some appreciation of the public antipathy to this technology. As a result, the GM debate has been fuelled largely by misinformation, and has generated much more heat than light. It is surprising that in the course of this debate, there has not been an accessible textbook available to which serious students of the subject could turn for information and understanding.

The aim of this book is to provide enough information and examples to give the reader a sound knowledge of plant biotechnology in all its various guises, but particularly those related to the genetic manipulation of crop plants. It is not intended to provide an encyclopaedic coverage of the subject; such a task would require a volume far larger than this. As such, this is a textbook, and therefore a learning and teaching aid, rather than an academic treatise. Hopefully, the book also encourages a critical appraisal of plant biotechnology—not just the scientific aspects, but also the economic, social, moral and ethical issues that surround, and, some would say, plague the subject. Although this book is aimed at undergraduate and postgraduate students, we do not assume a huge amount of prior knowledge, and hope that other people will find the book to be accessible, informative, and enjoyable. We hope that this book will make a contribution to the GM crop debate, facilitating a rational exchange of views between informed people.

The first four chapters of the book are designed to provide a more technical introduction to subjects such as gene expression, tissue culture, and plant transformation that enables the remainder of the book to be fully appreciated. They can, of course, be read in their own right and contain information relevant to other areas of biology. The rest of the book looks in detail at various aspects of plant genetic manipulation applied to crop improvement. There are four chapters which deal in depth with the application of plant genetic manipulation to agronomic traits (herbicide, pest, and disease resistance), which can be considered as first-generation plant biotechnology. Three of the remaining chapters deal with more challenging and diverse advances in the

areas of stress resistance, crop yield and quality and molecular ‘pharming’. It is in these areas that GM crops have the potential to produce real, wide-spread social and economic benefits, particularly for developing countries. The final chapter attempts to give an overview of plant biotechnology, past, present, and future, with reference to the legislative framework and economic, social, moral, and ethical issues.

Throughout the book, case-studies are used as extended illustrative examples of particular points that have been made in the main text. These are meant to be read as part of the body of each chapter. However, background information for clarification of advanced concepts, or more advanced information itself is clearly presented in ‘Boxes’, that can be read as and when required. There is, at the end of each chapter, a further reading section, which is not intended to be an exhaustive reference section, but gives enough pointers to allow and encourage further investigation. Various ‘web links’ to informative web-sites are also included in the further reading section. The World Wide Web provides an invaluable resource for investigating plant biotechnology, but care should be exercised when accessing information of dubious provenance. We would, however, encourage the use of the excellent online-journal sites to enable, in many cases freely, an interchange of knowledge and ideas, one of the cornerstones of science.

There is a web site associated with this book, which will contain hyper-linked chapter web links, further references, downloadable figures from the book and update sections. We hope that this web site will help to keep the reader of this book fully informed and up to date with developments in plant biotechnology: www.oup.com/uk/best.textbooks/biochemistry/slater/

We would like to thank all those who have contributed to the making of this book, not least the people whose original investigations are the basis of this book. We apologise unreservedly for any mistakes, all of which are ours, omissions or failure to acknowledge fully. Finally we would like to thank our friends and families, in particular Arlene, Jane, and Hilary, for their continued support and encouragement throughout the writing of this book.

February 2003

Contents

<i>Preface</i>	v
<i>List of Abbreviations</i>	xiii
<i>Foreword</i>	xxi

1 Plant genomes: the organisation and expression of plant genes	1
--	----------

Introduction	1
DNA, chromatin and chromosome structure	1
An introduction to gene structure and gene expression	6
Gene structure and expression in a eukaryotic protein-coding gene	6
Translation	7
Regulation of gene expression	12
Chromatin conformation	14
Gene transcription	14
RNA modification, splicing, turnover and transport	15
Translation	18
Post-translational modification	18
Localisation	19
Protein turnover	19
Conclusions	20
Implications for plant transformation	20
Examples of promoter elements used to drive transgene expression	24
Protein targeting	24
Heterologous promoters	24
Genome size and organisation	25
<i>Arabidopsis</i> and the new technologies	25
Genome-sequencing projects—technology, findings and applications	26
Biotechnological implications of the ‘ <i>Arabidopsis</i> Genome Sequencing Initiative’	30
Crop plant genome sequencing	30

Summary	32
Further reading	33

2 Plant tissue culture 35

Introduction	35
Plant tissue culture	35
Plasticity and totipotency	35
The culture environment	36
Plant cell culture media	36
Plant growth regulators	40
Culture types	43
Callus	43
Cell-suspension cultures	43
Protoplasts	44
Root cultures	45
Shoot tip and meristem culture	45
Embryo culture	45
Microspore culture	45
Plant regeneration	46
Somatic embryogenesis	47
Organogenesis	50
Integration of plant tissue culture into plant transformation protocols	51
Summary	52
Further reading	52

3 Techniques for plant transformation 55

Introduction	55
<i>Agrobacterium</i> -mediated gene transfer	55
The biology of <i>Agrobacterium</i>	55
The Ti plasmid	56
Ti-plasmid features	57
The process of T-DNA transfer and integration	59
Practical applications of <i>Agrobacterium</i> -mediated plant transformation	63
Transformation <i>in planta</i>	66
Direct gene transfer methods	66
Particle bombardment	67
Polyethylene glycol (PEG)-mediated transformation	72
Electroporation	72
Silicon carbide fibres—WHISKERS™	75

Summary	75
Further reading	77

4 Binary vectors for plant transformation 79

Introduction	79
Desirable features of any plasmid vector	79
Development of plant transformation vectors	80
Basic features of vectors for plant transformation	80
Promoters and terminators	82
Selectable markers	87
Reporter genes	88
Origins of replication	93
Co-integrative and binary vectors	93
Families of binary vectors	93
Optimisation	94
Arrangement of genes in the vector	94
Transgene copy number	96
Transgene position	96
Transgene features	97
Clean gene technology	97
Summary	99
Further reading	100

5 The genetic manipulation of herbicide resistance 103

Introduction	103
The use of herbicides in modern agriculture	104
What types of compounds are herbicides?	105
Strategies for engineering herbicide resistance	107
Prospects for plant detoxification systems	121
Commercialisation of herbicide-resistant plants to date	122
The environmental impact of herbicide-resistant crops	125
The development of 'super weeds'	126
Summary	129
Further reading	129

6 The genetic manipulation of pest resistance 131

Introduction	131
The nature and scale of insect pest damage to crops	131

GM strategies for insect resistance: The <i>Bacillus thuringiensis</i> approach to insect resistance	132
The use of 'Bt' as a biopesticide	136
Bt-based genetic modification of plants	137
The problem of insect resistance to Bt	139
The environmental impact of Bt crops	143
The 'Copy Nature' strategy	145
Insect resistant crops and food safety	152
Summary	153
Further reading	153

7 Plant disease resistance 157

Introduction	157
Existing non-GM approaches	157
Plant–pathogen interactions	159
Prokaryotes	160
Fungi and water moulds	160
Viruses	161
Natural disease resistance pathways—overlap between pests and diseases	162
Anatomical defences	163
Pre-existing protein and chemical protection	163
Inducible systems	163
Systemic responses	170
Biotechnological approaches to disease resistance	170
Protection against fungal pathogens	171
Antimicrobial proteins	174
Induction of HR and SAR in transgenic plants	175
Summary	176
Further reading	177

8 Reducing the effects of viral disease 179

Introduction	179
Types of plant viruses	179
RNA viruses	181
Entry and replication—points of inhibition	183
How has industry dealt with viruses?	184

The transgenic approach—PDR	187
Interactions involving viral proteins	188
RNA effects	192
What has been commercialised in the West?	198
Yellow squash and zucchini	198
Papaya	198
Potato	199
Risk	199
Summary	201
Further reading	202

9 Strategies for engineering stress tolerance 205

Introduction	205
The nature of abiotic stress	206
The nature of water-deficit stress	207
Different abiotic stresses create a water deficit	208
Targeted approaches towards the manipulation of tolerance to specific water-deficit stresses	215
Alternative approaches to salt stress	215
Alternative approaches to cold stress	218
Tolerance to heat stress	220
Secondary effects of abiotic stress—the production of reactive oxygen species	222
Strategy 1: Expression of enzymes involved in scavenging ROS	223
Strategy 2: Production of antioxidants	227
Summary	229
Further reading	229

10 The improvement of crop yield and quality 231

Introduction	231
The genetic manipulation of fruit ripening	232
Engineering plant protein composition for improved nutrition	251
The genetic manipulation of crop yield by enhancement of photosynthesis	252
Manipulation of light harvesting and the assimilate distribution—phytochromes	253
Direct manipulation of photosynthesis—enhancement of dark reactions	256
Summary	257
Further reading	258

11	Molecular farming/‘pharming’	261
	Introduction	261
	Carbohydrates and lipids	261
	Carbohydrate production	263
	Metabolic engineering of lipids	270
	Molecular farming of proteins	279
	Production systems	280
	Medically related proteins	289
	Economic considerations for molecular farming	298
	Summary	301
	Further reading	302
12	Future prospects for GM crops	305
	Introduction	305
	The current state of transgenic crops	305
	Who has benefited from these first-generation crops?	307
	What will drive the development of the future generations of GM crops?	309
	Concerns about GM crops	309
	Antibiotic resistance genes	309
	Herbicide resistance and ‘super-weeds’	311
	Gene containment	311
	Big business	312
	The regulation of GM crops and products	316
	The European Union (EU)	316
	The USA	320
	Future developments in the science of plant biotechnology	324
	‘Greener’ genetic engineering	324
	Summary	329
	Further reading	330
	<i>Index</i>	335

Abbreviations

2,4,5-T	2,4,5-trichlorophenoxyacetic acid
2,4-D	2,4-dichlorophenoxyacetic acid
2iP	2-isopentyl adenine
4-MU	methylumbelliferone
4-MUG	4-methylumbelliferryl- β -D-glucuronide
ABA	abscisic acid
ABRE	ABA response element
ACCase	acetyl-CoA carboxylase
ACNFP	(UK) Advisory Committee on Novel Foods and Processes
ACP	acyl carrier protein
ACRE	Advisory Committee on Releases to the Environment
AdoMetDC	adenosyl-L-methionine decarboxylase
AGI	<i>Arabidopsis</i> Genome Initiative
AHAS	acetohydroxy-acid synthase (aka ALS)
AIDS	acquired immunodeficiency syndrome
AFLP	amplification fragment length polymorphism
AIMV	alfalfa mosaic <i>alfamovirus</i>
ALS	acetolactate synthase (or more correctly, acetohydroxy-acid synthase or AHAS)
AMPA	aminomethylphosphonic acid
APHIS	Animal and Plant Health Inspection Service
ArMV	arabis mosaic <i>nepovirus</i>
AuxRE	auxin-response element
AVP	arginine vasopressin
BADH	betaine aldehyde dehydrogenase
BAP	benzylaminopurine
BNA	brazil nut albumin

BNYVV	beet necrotic yellow vein virus
bp	base pair
BSE	bovine spongiform encephalopathy
BSM	bacterial selectable marker gene
<i>Bt</i>	<i>Bacillus thuringiensis</i>
BYDV	barley yellow dwarf <i>luteovirus</i>
bZIP	basic zipper
C	constant region of an immunoglobulin
CDH	choline dehydrogenase
CDK	cyclin-dependent (protein) kinase
cDNA	complementary DNA
CE	coupling element
CMO	choline monooxygenase
CMS	cytoplasmic male sterility lines
CMV	cucumber mosaic <i>cucumovirus</i>
COD	choline oxidase
COR	cold responsive
CP	capsid or coat protein
Cp DNA	chloroplast DNA
CPMR	coat protein-mediated resistance
CpTI	cowpea trypsin inhibitor
CRT	C-repeat
CSSV	cacao swollen shoot <i>badnavirus</i>
CTD	C-terminal domain
CTV	citrus tristeza <i>closterovirus</i>
DAHPh	3-deoxy-D-arabino-heptulosonate 7-phosphate
DEFRA	Department for Environment, Food and Rural Affairs
DHFR	dihydrofolate reductase
DI	defective interfering (RNAs and DNAs)
DMADP	dimethylallyl diphosphate
DNA	deoxyribonucleic acid
DRE	dehydration responsive element (see also LTRE)
DREB1A	dehydration responsive-element binding protein
ds	double-stranded
DST	downstream
ECB	European corn borer (i.e. <i>Ostrinia nubilalis</i>)

EDTA	ethylenediaminetetraacetic acid
EFE	ethylene-forming enzyme
ELISA	enzyme-linked immunosorbent assay
EM	electron microscopy
EPA	Environmental Protection Agency
EPSPS	5-enolpyruvylshikimate-3-phosphate synthase
ER	endoplasmic reticulum
ERA	environmental risk assessment
EU	European Union
FBPase	fructose-1,6-bisphosphatase
FDA	(US) Food and Drug Administration
FDP	farnesyl diphosphate
FFT	fructan–fructan fructosyltransferase
FSE	farm-scale evaluations
G–F–F	glucose–fructose–fructose
GA	gibberellin
GALT	gut-associated lymphoid tissue
GBSS	granule-bound starch synthase
GC–MS	gas chromatography–mass spectrometry
GDP	geranyl diphosphate
GFLV	grapevine fanleaf <i>nepovirus</i>
GFP	green fluorescent protein
GGDP	geranyl geranyl diphosphate
GM	genetically modified
GMO	genetically modified organism
GNA	<i>Galanthus nivalis</i> (snowdrop plant) agglutinin
GOX	glyphosate oxidase
GR	glutathione reductase
GS	glutamine synthetase
GSH	glutathione
GST	glutathione S-transferase
GUS	β-glucuronidase
H	heavy chain of an Ig
HDGS	homology-dependent gene-silencing
HEAR	high-erucic acid rape
HIV	human immunodeficiency virus

HLH	helix–loop–helix
HMG	high-mobility group
HR	hypersensitive response
HSA	human serum albumin
HSE	heat-shock element
HSF	heat-shock factor
HSP	heat-shock protein
HTH	helix–turn–helix
IAA	indole-3-acetic acid
IBA	indole-3-butyric acid
ICP	insecticidal crystal protein
IDP	isopentenyl diphosphate
Ig	immunoglobulin
IRGSP	International Rice Genome Sequencing Project
IPM	integrated pest management
IPR	intellectual property rights
IR	inverted repeats
IRRI	International Rice Research Institute
ISAAA	International Service for the Acquisition of Agri-biotech Applications
ISR	induced systemic resistance
J	joining region of an Ig
JIP	jasmonate-induced protein
JRA	Joint Regulatory Authority
kDa	kilodalton
K_M	Michaelis–Menten constant
L	light chain of an Ig
LD ₅₀	median lethal dose
LEA	late embryogenesis abundant
LEAR	low-erucic acid rape
LRE	light-responsive element
LRR	leucine-rich repeat
LSC	large single copy
LTRE	low temperature-response element (see also DRE)
MAR	matrix-attachment regions
MCPA	2-methyl-4-chlorophenoxyacetic acid

MCS	multiple cloning site
MeJA	methyl jasmonate
MEV	mink enteritis virus
MP	movement protein
mRNA	messenger RNA
MS	Murashige and Skoog (culture medium)
MTA	material transfer agreement
NAA	1-naphthylacetic acid
NADPH	reduced nicotinamide adenine dinucleotide phosphate
NB	nucleotide binding
NBS	nucleotide binding sequence
NLS	nuclear localisation signal
NOA	2-naphthyloxyacetic acid
NOS	nopaline synthase
NPC	nuclear-pore complex
OECD	Organisation for Economic Co-operation and Development
ORF	open reading frame
PAMP	pathogen-associated molecular pattern
PAT	phosphinothricin acetyltransferase
PB	protein bodies
PCR	polymerase chain reaction
PDR	pathogen-derived resistance (originally known as parasite-derived resistance)
PEG	polyethylene glycol
PEP	phospho <i>enol</i> pyruvate
PEPC	phospho <i>enol</i> pyruvate carboxylase
P _{FR}	far-red light-absorbing phytochrome
PG	polygalacturonase
PHA	polyhydroxyalkanoates
PHB	polyhydroxybutyrate
Phy	phytochrome
PME	pectin methylesterase
Pol	polymerase
pp2C	protein phosphatase type 2C
PPT	phosphinothricin (aka glufosinate)
pPTV	plasmid plant transformation vector

var	variety
VLP	virus-like particles
VP	virus particle
VPg	genome-linked protein
WMV	watermelon mosaic <i>potyvirus</i>
X-gluc	5-bromo-4-chloro-3-indolyl β -D-glucuronide
ZYMV	zucchini yellow mosaic <i>potyvirus</i>

Foreword

The “Green Revolution”, led by Norman Borlaug, Monkombu Swaminathan and Gurdev Khush, enabled the world’s food supply to be tripled during the last three decades of the 20th Century. The extraordinary increase in agricultural productivity was made possible by the adoption of genetically improved varieties coupled with advances in crop management. In many countries food supply increased faster than demand and the technological progress contributed to a decrease in the unit cost of production so that farmers were able to share the benefits of the advances with consumers, by offering food to them at lower prices. Intensive (high input/high yield) agriculture has served the populations of the developed countries well but two problems have come to occupy these people. Firstly the full-scale exploitation of intensive agriculture protocols does deliver high yields of high quality produce, but the environmental impact of the processes is often high. Secondly, the farmers have, in fact, the need to dispose of surplus food so that there is a downward pressure on prices in the world market that undermines farmers’ incentives.

The position of the people in the low-income countries contrasts starkly with that in the developed countries. The world’s population has increased from 2.5 billion to 6.1 billion in the last 50 years and it is unlikely to stabilise before 2100 by which time another 3 billion people will inhabit the planet. Most of the increase will occur in the low-income countries where poverty and hunger are already widespread. Each night 800 million people go to bed hungry and suffer from malnutrition, and one-fifth of humankind (about 1.2 billion people) lives on earnings of less than a dollar per day. We must satisfy the need for more food in an environmentally friendly way but we are confronted by major challenges. Prime agricultural land is being diverted to non-agricultural uses to meet the growing demand from housing, urbanisation and industrialisation. Countries inhabited by half the world’s population are already experiencing water crises, while the high agrochemical inputs that maximise yields exert a high environmental impact, which is not acceptable. There is a desperate need to produce more food from less land with less water and reduced agrochemical inputs.

The majority of agricultural scientists, led by Borlaug, Swaminathan and Khush, are convinced that the required high yield/high quality/low cost/low environmental impact crops can be delivered by the exploitation of the tech-

being held together by hydrogen bonding. For this DNA to be packed into a nucleus it must be carefully arranged and higher order structures imposed on it. The necessity for this packing is not surprising when one considers that the DNA in the nucleus of a typical higher eukaryote might be several metres long in total.

BOX 1.1 The chloroplast and mitochondrial genomes of plants

Plastids and mitochondria are thought to have evolved from prokaryotes that were taken up by eukaryotic cells in symbiotic associations very early in their evolution. Although no longer capable of existing independently of the eukaryotic cell, they retain many features of their prokaryotic ancestors. Both organelles have their own genome, although by comparison with the nuclear genome it is extremely simple. Part of the reason for this simplicity seems to be that during evolution many of the genes from these organelles have been 'transferred' to the nucleus. Thus, many proteins that function in chloroplasts or mitochondria are encoded by nuclear genes, the proteins being transported to the organelle.

The genomes of the organelles have maintained many of the features of those of their prokaryote ancestors. Organelle genomes are circular, like bacterial genomes, and do not form complex structures like the chromatin of the nuclear genome. Some of the genes in chloroplasts and mitochondria are arranged in operons, with a common regulatory element or promoter. The promoters of many chloroplast and mitochondrial genes exhibit features of prokaryotic promoters. Protein synthesis in chloroplasts and mitochondria also resembles that of prokaryotes.

The chloroplast

The chloroplast genome is a circular, double-stranded DNA molecule (or chromosome) located in the stroma of the chloroplast. Chloroplast genomes are highly conserved amongst plant species. There is more than one copy of the genome in each chloroplast (the exact number varies during development, but mesophyll cells in young leaves contain about 100 copies of the genome).

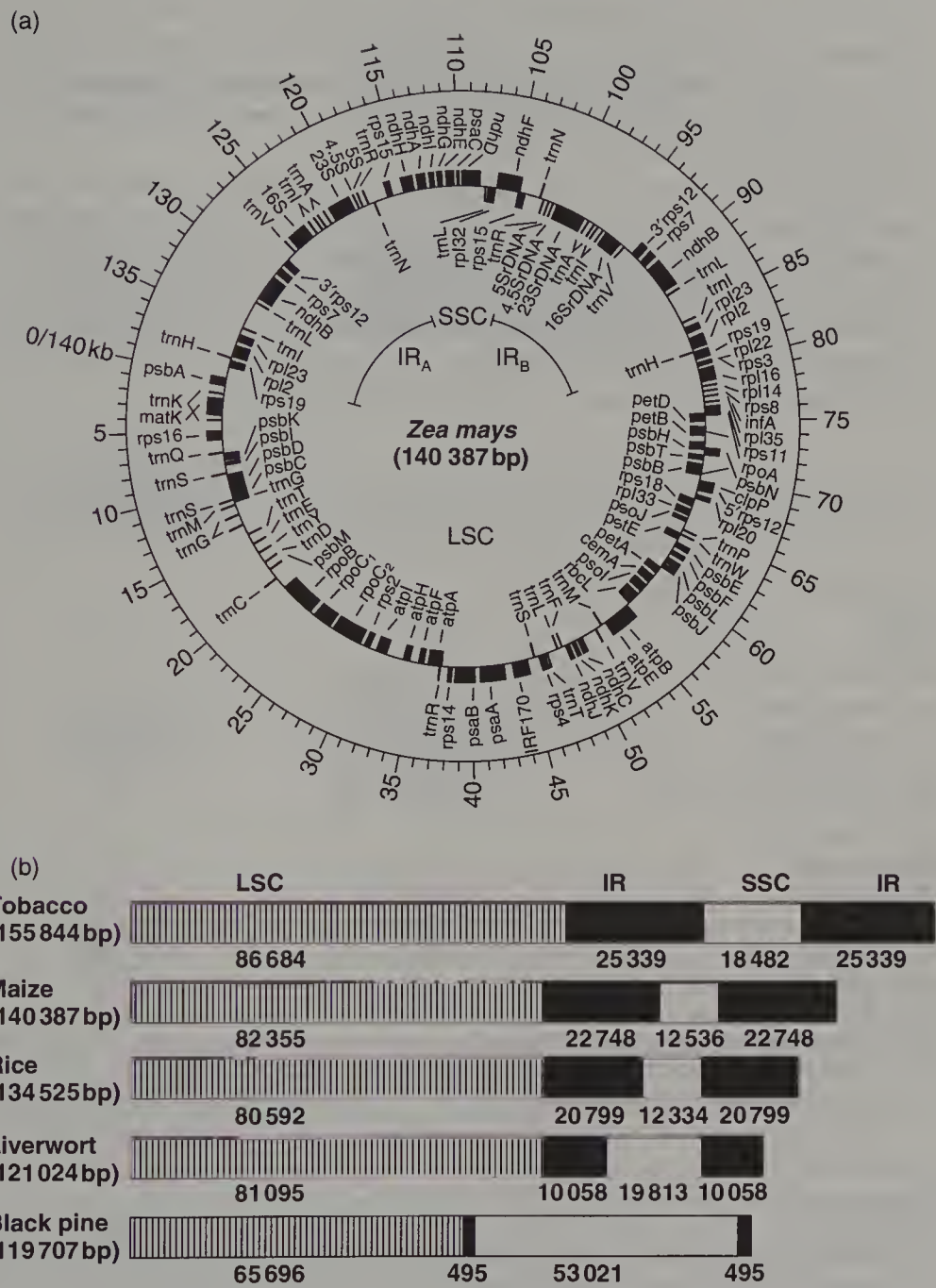
The details of the chloroplast genome can be looked at with reference to the figure, which shows the chloroplast genome of maize in detail and some comparisons with the chloroplast genomes of other species. Most chloroplast genomes are between 120 and 160 kbp in size and contain 120–140 genes. Chloroplast genomes can be categorised into three groups, based on the presence or absence of inverted repeats. The majority of land plants (excluding some legumes and conifers) are characterised by the presence of inverted repeats and have group II genomes.

As can be seen from the table, most of the genes present in the chloroplast genome are involved in either protein synthesis or photosynthesis (although the majority of the genes encoding proteins for photosynthesis are located in the nucleus). Many of these genes are in clusters, this allows expression in the form of large polycistronic primary transcripts which are processed to oligo- or monocistronic mRNAs.

The mitochondrial genome

Mitochondria, which are common to all eukaryotes, also contain their own genome. Plant mitochondrial genomes, which are much larger than those of mammals and yeast, range in size from about 200 kbp in *Brassica* spp. to over 2500 kbp in *Cucumis melo* (muskmelon). This variability stems, in part, from the

BOX 1.1 Continued



(a) Maize chloroplast genome map. The chloroplast genome of maize is 140 387 base pairs. It contains a pair of inverted repeats (IR_A and IR_B) which are separated by the small (12 536 bp) and large (82 355 bp) single-copy regions (SSC and LSC, respectively). Genes on the outside of the circle are transcribed clockwise, those on the inside anticlockwise. Overall, the A + T content is 61.5%, although this varies from 71.2% in non-coding regions to 45.3% in rRNA genes.

(b) Length and feature analysis of some sequenced chloroplast genomes. (Redrawn with permission from Maier, R. M., *et al.* (1995).)

BOX 1.1 Continued

accumulation of large regions of non-coding DNA. Thus, the maize mitochondrial genome is much larger than the chloroplast genome but contains far fewer genes, most of the genes being involved in protein synthesis or respiratory electron transport. Plant mitochondrial DNA also contains chloroplast DNA (termed 'promiscuous DNA').

The plant mitochondrial genome can, surprisingly, exhibit changes in its structure (although this property is not universal). Complete, circular 'master' chromosomes (or master circles) are postulated to contain all the mitochondrial genes; these, however, have never been isolated or identified. Subgenomic, circular DNA circles lacking a complete set of mitochondrial genes have been identified. These are thought to be formed from the master chromosome by recombination events between repeated DNA sequences in the mitochondrial DNA.

Gene-product function	Acronym ^a	Number of genes in chloroplast genome
<i>Transcription/translation</i>		
rRNA	<i>rrn</i>	4
tRNA	<i>trn</i>	30
Ribosomal proteins		
— small subunit	<i>rps</i>	12
— large subunit	<i>rpl</i>	9
RNA polymerase	<i>rpo</i>	4
Translation initiation factor	<i>inf</i>	1
<i>Photosynthesis</i>		
Photosystem I	<i>psa</i>	5
Photosystem II	<i>psb</i>	14
Large subunit of rubisco	<i>rbcL</i>	1
Cytochrome	<i>pet</i>	4
ATP synthase subunits	<i>atp</i>	6
NADH dehydrogenase	<i>ndh</i>	11

^a The acronym can be used to identify the position of the gene in the genome.

Chromatin

The first-order structure of DNA packing consists of DNA wrapped round cores composed of histones, which are basic proteins (Figure 1.1). Each core has two turns of DNA wrapped round it (approximately 150 base pairs (bp)) and is termed a 'nucleosome'. Each nucleosome is separated from the next by spacer DNA, the length of which can vary. This continuous string of nucleosomes is termed the '10-nm fibre'. This 10-nm fibre is further coiled to produce the 30-nm fibre. The latter fibre has a coiled or solenoid structure, with six nucleosomes to every turn. These 30-nm fibres can be further organised by anchoring the fibre at various points to a protein scaffold and introducing 'loops'. During mitosis, when chromosomes

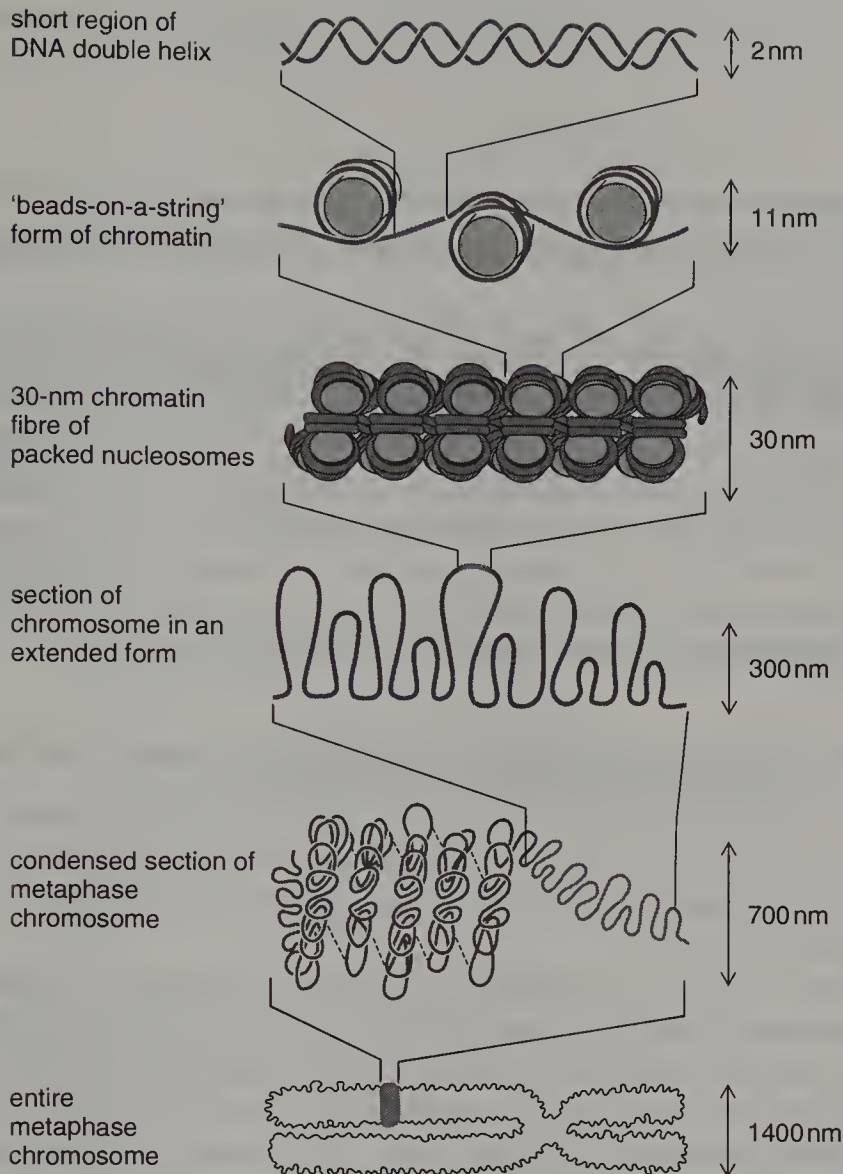


Figure 1.1 The structure of DNA, chromatin and the higher order packing of chromatin. DNA is a linear, double-stranded molecule. The nucleosome is the basic unit of chromatin organisation and consists of approximately 150 bp of DNA (two turns) wrapped round a core of histone proteins. This core is composed of two molecules each of histones H2A, H2B, H3 and H4. Nucleosomes are separated by spacer DNA to which histone H1 is attached. This structure of a string of nucleosomes, termed the '10-nm fibre', considerably reduces the length of the DNA. This structure is further coiled to produce the 30-nm fibre, which has a solenoid structure with six nucleosomes to every turn. This 30-nm fibre can be further organised into loops by anchoring the fibre at A/T-rich areas (scaffold-associated regions, or SARS) to a protein scaffold. In mitosis, when the transcription of the vast majority of genes is halted, the loops can be further coiled and the chromosomes become condensed. (Reproduced with permission from Alberts, B., *et al.* (1989).)

condense and become visible (with the aid of a microscope), these loops can be further coiled.

An introduction to gene structure and gene expression

Defined regions of the genome, termed 'transcription units', serve as templates for the synthesis of another nucleic acid, called 'ribonucleic acid' (RNA). There are three main types of RNA: ribosomal RNA (rRNA), which is a structural component of ribosomes; transfer RNA (tRNA), which is a class of small RNAs involved in protein synthesis (or translation); and messenger RNA (mRNA), which carries the protein-coding information from DNA and is subsequently translated during protein synthesis. Although rRNA and tRNA are not translated to give proteins, the DNA regions that code for these RNAs are still termed genes.

Gene structure and expression in a eukaryotic protein-coding gene

As plant biotechnology deals largely with genes that encode proteins, the structure of this type of gene can be looked at in more detail. Figure 1.2 gives a diagrammatic representation of the main features of a eukaryotic protein-coding gene.

Each transcription unit is associated with its own regulatory regions; the most important of which is considered to be a flanking region at the 5' end of a gene, i.e. the promoter. One important feature of the promoter is that it contains the binding site for the enzyme RNA polymerase II, this enzyme is responsible for making an RNA copy of the DNA in a process known as 'transcription'. Transcription is initiated from defined points (transcription start sites) close to the RNA polymerase binding site (see Figure 1.3).

The transcription unit is the region of the gene that is transcribed by RNA polymerase into pre-mRNA. This pre-mRNA is composed of a 5' untranslated region (A), the coding region and a 3' untranslated region (B). The coding region is composed of exons and introns. Exons are represented in the mature mRNA and carry the information necessary for protein synthesis (they are translated), and are therefore said to be 'coding'. Exons are separated by introns (**NB** not all genes contain introns) that are not present in the mature mRNA and are said to be 'non-coding'.

The 3' end of the gene also contains a poly(A) site (AATAAA) which acts as a signal for the adenylation of the RNA. During polyadenylation and processing of the pre-mRNA, the RNA is cleaved, resulting in the pre-mRNA being shorter than the transcription unit. The pre-mRNA is also modified at its 5' end by the addition of a 'cap' of 7-methylguanosine linked to the pre-mRNA by three phosphate groups. The cap and the poly(A) tail

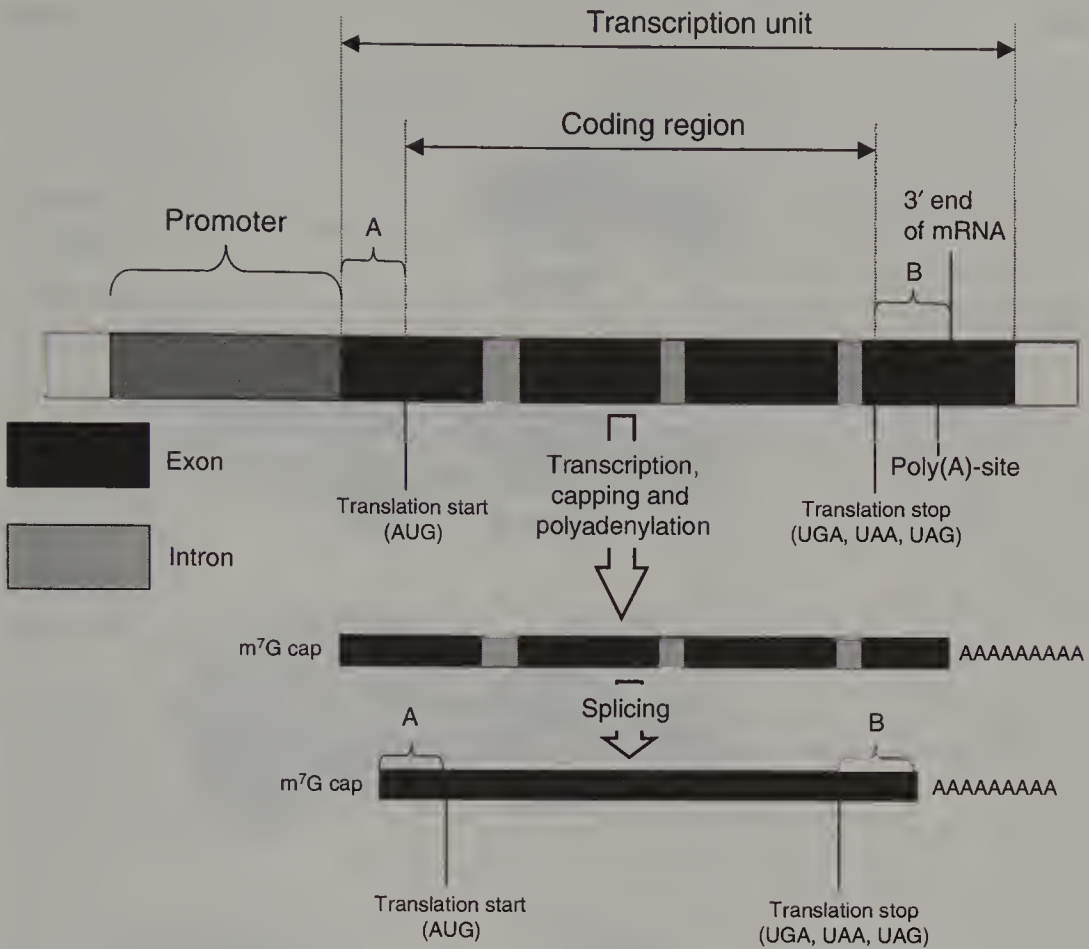


Figure 1.2 The structure and expression of a eukaryotic protein-coding gene. Details are explained in the text.

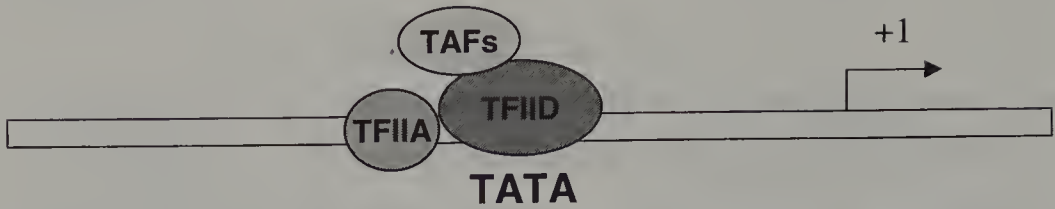
tend to make the RNA more stable by preventing nuclease-mediated destruction.

Introns are removed, or spliced out, during mRNA processing. Exon/intron boundaries are delineated by specific base sequences that act as signals for the cleavage and removal of the intron. The left-hand (5') border (AAG↓GUAAGU, where ↓ indicates the position of cleavage) of the intron is cleaved first, and the free end of the intron loops back and forms a bond with a specific residue in the right-hand border (GCAG↓GU). This structure is called a 'lariat'. This right-hand border is then cleaved and the exons joined together. Plant introns are also characterised by UA-rich regions that may be important for intron splicing in plants. The mRNA is then exported from the nucleus to the cytoplasm where it associates with ribosomes and is translated to produce the protein gene-product.

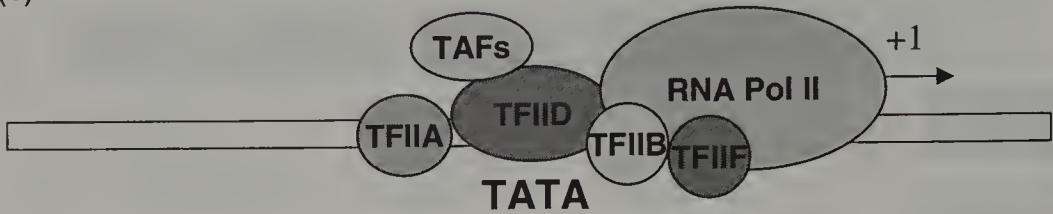
Translation

During translation the mRNA serves as a template for ordering the amino acids (Table 1.1) that make up the protein gene-product. The process is

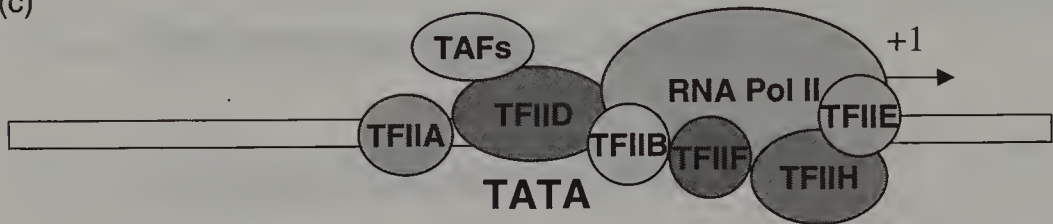
(a)



(b)



(c)



(d)

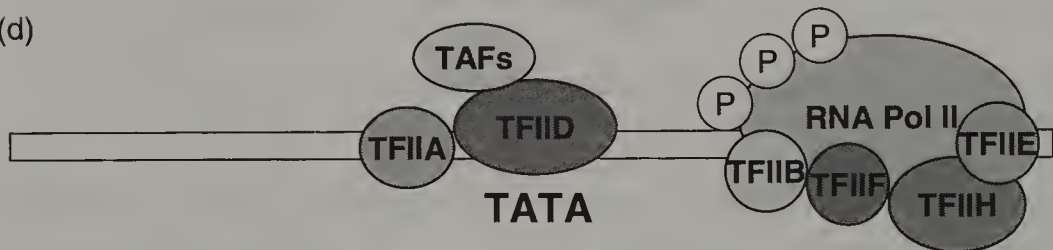


Figure 1.3 A simplified explanation of RNA polymerase II binding to a eukaryotic promoter.

(a) General transcription factors bind to the TATA box of the promoter. The first protein to bind, via its TBP (TATA-binding protein) subunit, is TFIID (transcription factor IID), which is itself a multiprotein complex. This interaction requires TBP-associated factors (TAFs) and is promoted by TFIIA.

(b) TFIIB binds to the complex and recruits TFIIF, which is bound to RNA polymerase (Pol) II.

(c) TFIIE and TFIIH subsequently bind to the complex. An ATP-dependent helicase activity associated with TFIIH unwinds the promoter DNA. In this state the RNA polymerase is said to be in the initiation mode.

(d) The C-terminal domain (CTD) of RNA polymerase is then phosphorylated, causing transition to the elongation mode, and transcription is initiated from the start site (+1). In animals, it is known that several cyclin/CDK (cyclin-dependent (protein) kinase) complexes can phosphorylate the CTD of RNA polymerase II; CDK7/cyclinH (a component of TFIIH), CDK8/cyclinC and CDK9/cyclinT.

(Data taken with permission from Buratowski, S. (2000).)

Table 1.1 Amino acids

Amino acid	Single-letter representation ^a	Three-letter representation ^a	Number of codons ^b
Alanine	A	Ala	4
Arginine	R	Arg	6
Asparagine	N	Asn	2
Aspartic acid	D	Asp	2
Cysteine	C	Cys	2
Glutamic acid	E	Glu	2
Glutamine	Q	Gln	2
Glycine	G	Gly	4
Histidine	H	His	2
Isoleucine	I	Ile	3
Leucine	L	Leu	6
Lysine	K	Lys	2
Methionine	M	Met	1
Phenylalanine	F	Phe	2
Proline	P	Pro	4
Serine	S	Ser	6
Threonine	T	Thr	4
Tryptophan	W	Trp	1
Tyrosine	Y	Tyr	2
Valine	V	Val	4
STOP	—	—	3

^a Each amino acid can be represented by either a one- or three-letter code.

^b Each amino acid can be 'coded' for by up to six codons.

known as 'translation' because the nucleotide 'language' of DNA and RNA is converted into the amino acid language of proteins (Figure 1.4 gives a simplified overview of translation).

The small subunit of a ribosome bound with a tRNA carrying the amino acid methionine associates with the 5' end of mRNA. The large subunit then binds to the small subunit and translation is initiated from the translation start site (AUG). The ribosome passes along the mRNA in a 5' to 3' direction. When a stop codon is reached the ribosome dissociates into its two subunits and the protein is released. Several ribosomes may translate an mRNA at once, forming a 'polysome'.

Each amino acid is 'coded' for by a group of three nucleotides called 'a codon' (see Table 1.2). These codons are translated one at a time. Some amino

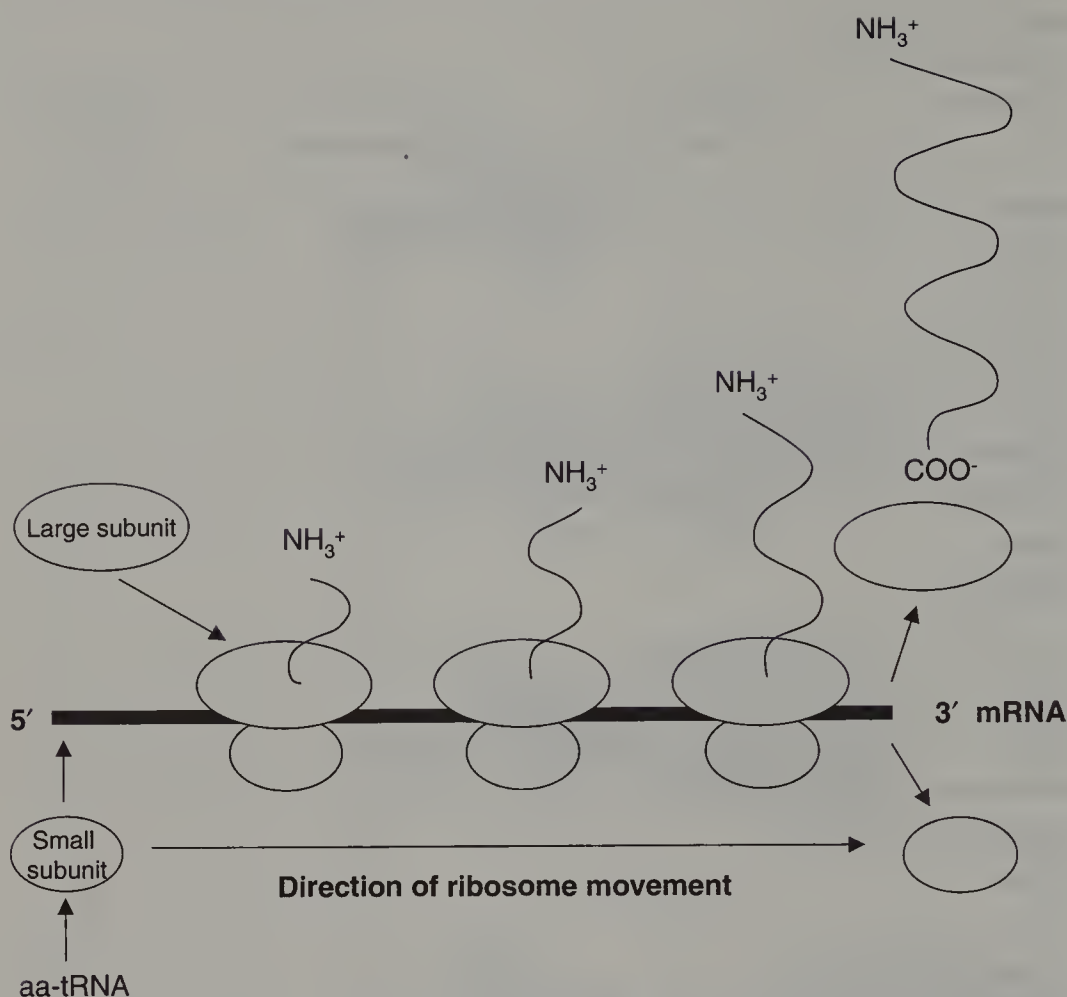


Figure 1.4 A simplified overview of translation. (Adapted with permission from Buchanan, B. B., *et al.* (2000).) Details are explained in the text.

acids are encoded by only one codon, some by multiple codons. Amino acids have to be attached to a tRNA molecule that acts as an ‘adapter’ to link each mRNA codon with the appropriate amino acid. In order to do this each ‘type’ of tRNA has to fulfil two tasks. An ‘anticodon’ (i.e. three complementary nucleotides) must base-pair with the appropriate codon on the mRNA (Figure 1.5). The 3’ end of the tRNA must also bind to a specific amino acid, whence the tRNA is said to be charged.

Protein synthesis actually occurs on specialised multiprotein complexes called ‘ribosomes’. Ribosomes are composed of both proteins and ribosomal RNA (rRNA). Ribosomes function to bring the tRNAs carrying amino acids into proximity with the mRNA template and to catalyse the formation of peptide bonds between the amino acids in order to form the protein.

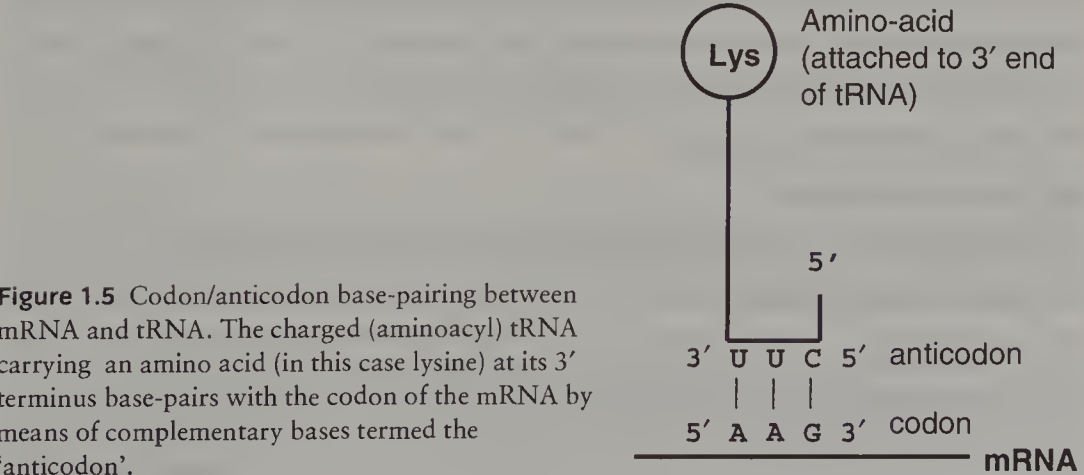
Initiation of translation

The small subunit of a ribosome, which already has a tRNA charged with methionine associated with it, binds to the 5’ cap structure of the mRNA. The

Table 1.2 The genetic code

		Second base						
		U ^a		C		A		G
First base	U ^a							
	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
	UUA	Leu	UCA	Ser	UAA	STOP	UGA	STOP
	UUG	Leu	UCG	Ser	UAG	STOP	UGG	Trp
	C							
	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
	A							
	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
	AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
	AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
	AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg
	G							
	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly
	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly

^a In RNA, T is replaced with U (uracil).
Notes: 61 of the possible 64 codons code for amino acids. Three (UAA, UAG, UGA) cause termination of translation and are termed STOP codons. AUG is the translation initiation codon, but also codes for internal methionines. One amino acid is inserted into the growing protein for every 3-nucleotide codon of the mRNA. For amino acids with multiple codons, some species will use particular codons more frequently than others, exhibiting so-called ‘codon-preference’. This point can be illustrated if we look at the codon usage for proline in alfalfa and barley. There are four codons for proline (CCU, CCC, CCA and CCG), which in alfalfa are used with the following frequencies: CCU 41%, CCC 10.4%, CCA 43.3% and CCG 5.3%. In barley, the codon usage is CCU 14.5%, CCC 31.1%, CCA 23.3% and CCG 32.1%. Thus, in alfalfa, CCG is the least commonly used codon for proline, whilst in barley it is the most commonly used.



small subunit then identifies the start AUG codon on the mRNA, establishing the start position and reading frame for translation. This is done in a process known as 'scanning' as the small subunit of the ribosome travels along the mRNA in a 5' to 3' direction. Translation is generally initiated from the first AUG codon on the mRNA, but the sequence surrounding this start codon influences the efficiency of initiation of translation. The large subunit of the ribosome then binds to the small subunit.

Elongation

The addition of amino acids to the growing proteins occurs one amino acid at a time and involves three sites of the ribosome, which are termed 'P', 'A' and 'E' (see Figure 1.6 for an overview of translation elongation). The P-site is where peptide bonds are formed between the growing protein and the charged tRNA that occupies the P-site. The A-site, which is empty, binds a charged tRNA if the anticodon of the tRNA matches the codon exposed in the A-site. Any tRNA (now uncharged) in the E-site is ejected. A peptide bond is formed between the amino acid on the tRNA in the A-site and the growing protein, effectively transferring the growing protein to the amino acid attached to the tRNA occupying the A-site. This process results in the growing protein being extended by one amino acid.

The uncharged tRNA now occupying the P-site is moved into the now-vacant E-site. The tRNA (with the growing protein) is moved from the A-site to the P-site and the ribosome moves along the mRNA by one codon. This movement exposes another codon in the A-site, which then binds another charged tRNA with the appropriate anticodon. This process is repeated until a stop codon is reached, at which point the protein is released and the ribosome dissociates from the mRNA.

Regulation of gene expression

It is obviously of vital importance that the expression of genes is carefully controlled to ensure they are expressed in the correct spatial and temporal fashion. In eukaryotes, gene expression can be controlled at many different 'levels' or stages. These can be broadly classified as follows:

- (1) chromatin conformation;
- (2) gene transcription;
- (3) nuclear RNA modification, splicing, turnover and transport;
- (4) cytoplasmic RNA turnover;
- (5) translation;
- (6) post-translational modification;
- (7) protein localisation;
- (8) protein turnover.

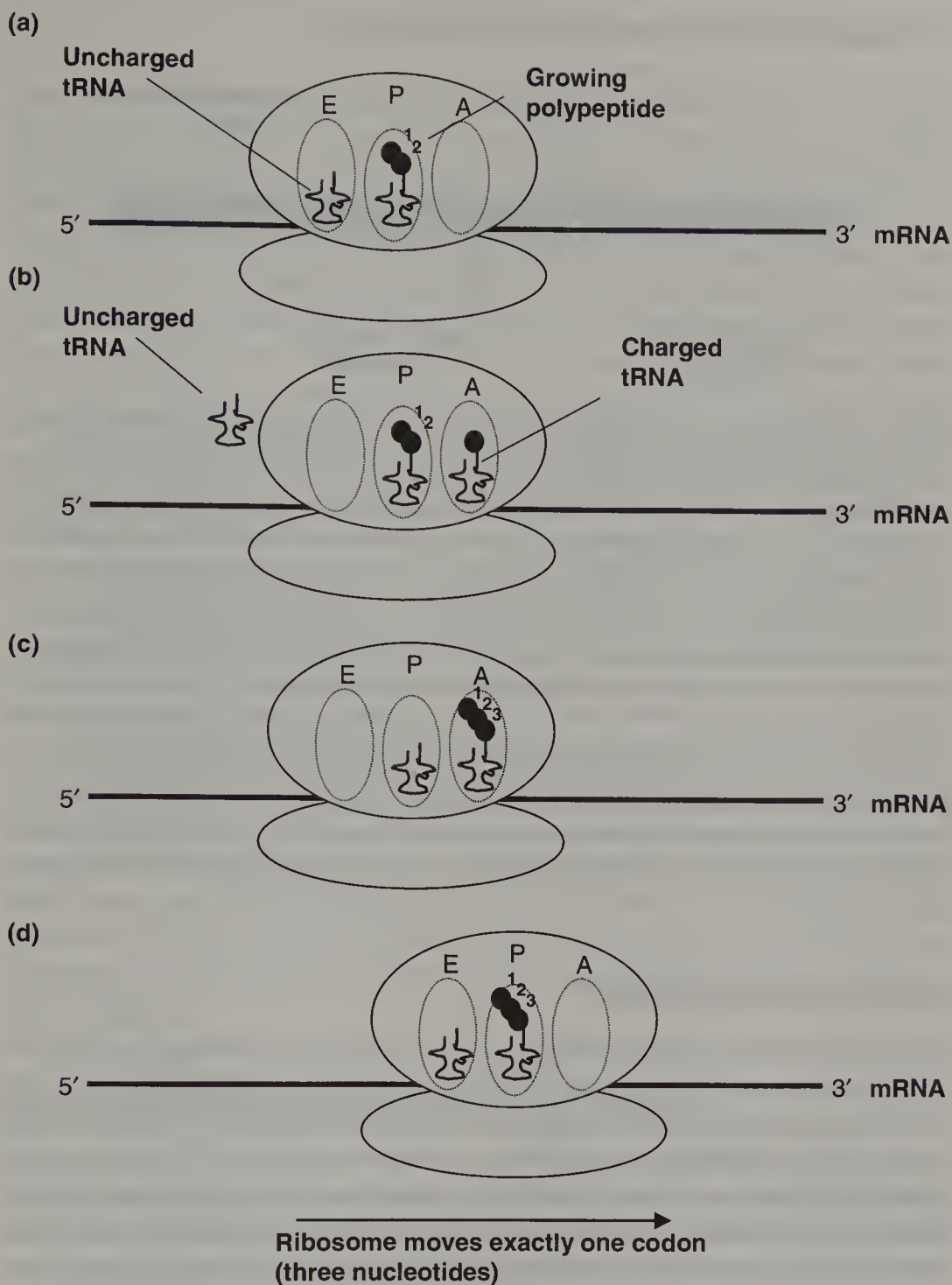


Figure 1.6 Elongation reactions of protein synthesis.

(a) A tRNA, to which is attached the growing protein, occupies the P-site of the ribosome. The E-site is occupied by an uncharged tRNA. The A-site is empty, thereby exposing a codon on the mRNA.

(b) A charged (aminoacyl) tRNA binds to the A-site if its anticodon is complementary to the codon of the mRNA exposed in the A-site. The uncharged tRNA that was occupying the E-site is ejected.

(c) A peptide bond is formed between the growing protein (on a peptidyl-tRNA) and the amino acid on the aminoacyl-tRNA occupying the A-site. The growing protein is transferred to the tRNA occupying the A-site. The protein is now one amino acid longer.

(d) The uncharged tRNA occupying the P-site moves to the E-site. The peptidyl-tRNA in the A-site moves to the P-site and the ribosome moves along the mRNA by exactly one codon. This process is called 'translocation'. This translocation exposes the next codon of the mRNA in the now-vacant A-site.

(Adapted with permission from Buchanan, B. B., *et al.* (2000).)

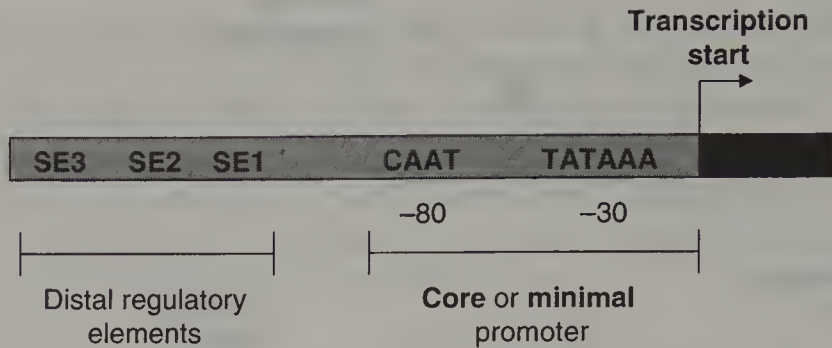


Figure 1.7 A simplified diagram of the promoter from a eukaryotic protein-coding gene. The promoter region contains several sequence elements (*cis*-elements). Some, such as the TATA box (consensus sequence T(C/G)TATA(T/A)A₁₋₃) and the CAAT box, comprise what is known as the 'minimal' or 'core' promoter. Much of the core promoter is involved in binding general transcription factors and RNA polymerase II. Core or minimal promoters are often incapable of initiating transcription by themselves. Promoters also contain distal regulatory elements (SEs, or sequence elements in this figure) that bind transcription factors (*trans*-activating factors) which regulate transcription from the gene. Transcription factors can either promote transcription or inhibit transcription. Several different transcription factors may bind to the promoter of a gene. The regulation of plant gene expression by internal and external stimuli and in particular tissues is mediated by the binding and action of a host of transcription factors. Many genes are capable of responding to more than one stimulus due to the presence of multiple *cis*-elements in their promoter. The presence of multiple *cis*-elements in a promoter and the large number of transcription factors means there is a huge number of potential control 'modules' that can be used to regulate gene expression. Gene expression may also be regulated by elements that are located a great distance from the core promoter. These elements, that may be located either 5' or 3' of the gene they regulate, are termed 'enhancers'.

Chromatin conformation

Chromatin conformation influences gene expression by affecting the ability of RNA polymerase to bind to the DNA. The accessibility of DNA in the nucleosome to RNA polymerase is regulated by acetylation of lysine residues in the core histones. Methylation of DNA can block transcription by altering chromatin conformation. Other chromatin remodelling systems are also found in plants. These systems can induce reversible changes in chromatin conformation that can control gene expression. Changes in chromatin conformation can, unlike regulation by transcription factors, affect gene expression over large stretches of the chromosome.

Gene transcription

Most of the controls present in a plant gene are located 5' of the transcription start site in the region known as 'the promoter'. The promoter contains various sequence elements (short stretches of DNA of a defined sequence) which function to recruit the proteins that regulate transcription of the gene (Figure 1.7). These sequence elements are termed '*cis*-elements' (because they are on the same DNA strand as the coding region of the gene). The most basic of

these sequence elements is the TATA box, which is found about 25–30 bp upstream of the transcription start site. The TATA box functions to direct RNA polymerase II to the correct position on the gene to initiate transcription. Some genes contain multiple TATA boxes, whilst others may contain no recognisable TATA box. The lack of a TATA box was previously thought to be a feature associated with ‘housekeeping’ genes (i.e. ones that are constitutively expressed) but this has proved not to be the case, with some inducible genes having no TATA box in their promoter. Other motifs can compensate for the lack of a TATA box, and in some classes of genes (such as nuclear genes involved in photosynthesis) a TATA-less promoter may be common. Other sequence elements are often found further upstream from the TATA box. In eukaryotic genes the most common of these are the CAAT and GC boxes which enhance the activity of RNA polymerase. Sequence elements such as the TATA box are often referred to as ‘core’ or ‘minimal’ promoter elements, as they contribute to the binding of RNA polymerase to the promoter and are responsible for the initiation of basal transcription. The core promoter, via interactions with other promoter elements, plays an important role in the regulation of plant gene expression.

Other sequence elements are also found in the promoters of genes. These are involved in binding specific proteins, termed ‘transcription factors’ (Box 1.2), which are required for transcription by RNA polymerase. These transcription factors (or *trans*-acting factors) are involved in linking various signals (both external and internal) to gene expression and are responsible for determining the level, place and timing of expression. Similar sequence elements are often found in genes with similar functions or that respond to similar signals.

RNA modification, splicing, turnover and transport

Many processes must be completed so that functional mRNA can enter the cytoplasm for subsequent translation. All these processes may be used to exert some control over gene expression. Different splice sites may be recognised during splicing of the pre-mRNA, thus generating different transcripts from the same gene. This alternative splicing may be used to generate different transcripts in different tissues, or in response to different stimuli, or to generate multiple transcripts from a single gene (Box 1.3 gives some examples of alternative splicing in plants). Alternative splicing is a common control mechanism in some organisms, and some well-documented examples of biological significance exist in plants.

The addition of a 5′ cap and a 3′ poly(A) tail is known to stabilise RNA (although the most profound effect of these modifications is to increase the efficiency of translation). The half-lives of individual mRNA species may still vary considerably (from minutes to days), with specific sequences or motifs thought to be responsible for rapid mRNA destabilisation. In plants, an

BOX 1.2 Plant transcription factors

Plant transcription factors contain several domains: a DNA-binding region, an oligomerisation region and a transcription regulation domain. Most transcription factors have only one type of DNA-binding region and one type of oligomerisation region, although some may contain two distinct types. The DNA-binding region contains amino acids (which are often basic) that interact with DNA at the *cis*-sequence elements. These determine the specificity of the transcription factor. The oligomerisation region is important in determining protein–protein interactions, as most transcription factors can form hetero- and/or homodimers. The formation of these dimers may affect sequence-element specificity, transactivation efficiency or DNA-binding affinity. Transcription regulation domains either repress or activate transcription.

Transcription factors are classified, or categorised, on the basis of their structural features. Plants appear to have largely evolved independently when it comes to transcriptional control. Some classes of eukaryotic transcription factors appear to be absent from the genome of *Arabidopsis*, whilst many classes are apparently unique to plants. Some conserved classes of transcription factors fulfil different functions in plants as opposed to animals.

The main classes of transcription factors are:

1. Helix–turn–helix (HTH). This type of transcription factor binds DNA as a dimer.
2. Leucine zipper. This type also binds DNA as a dimer. A variation is the basic zipper (bZIP).
3. Zinc finger. Bind to DNA as either dimers or monomers.
4. Helix–loop–helix (HLH). Bind to DNA as dimers.
5. High-mobility group (HMG)-box motif.

element termed ‘DST’ (downstream element) has been identified in the 3′ untranslated region (UTR) of short-lived mRNAs that are induced by the plant hormone auxin (small auxin-up RNA or *SAUR* transcripts). Two conserved subdomains (ATAGAT and GTA) appear to be critical for DST function. In common with mammals, AU-rich areas (which usually contain multiple copies of the sequence motif AUUUA) destabilise plant mRNAs if they are present in the 3′-UTR (these sequence motifs are discussed in Chapter 4). A variety of environmental and hormonal signals can also influence the half-life of specific mRNAs. For example, the half-life of the *Fed1* gene (encoding the photosynthetic electron carrier ferredoxin) transcript is much increased in plants in the light compared with those in the dark and α -amylase transcripts are destabilised by the presence of sucrose.

These ‘post-transcriptional’ processes can have a profound effect on the amount of mRNA present in the cytoplasm of a cell. A long-studied example of this is in the accumulation of legumin mRNA (legumin is a storage protein in seeds) following the recovery from sulphur starvation—here the levels of

BOX 1.3 Alternative splicing

Alternative splicing is an extremely important control mechanism in humans, where over one-third of the genes may be alternatively spliced. Although predicted to be not as widespread in plants, it is becoming apparent that alternative splicing is used in plants as a control mechanism. In this box we will look at just two examples in a little detail to illustrate the complexity that may be introduced to gene expression by alternative splicing (and processing).

Rubisco activase

The alternative splicing of ribulose-1,5-bisphosphate carboxylase/oxygenase activase, better known as 'Rubisco activase', is one of the best-studied examples of alternative splicing in plants. Rubisco activase is a nuclear-encoded gene for a chloroplast protein that regulates Rubisco activity in response to light-induced changes in the redox potential and changes in the ADP/ATP ratio.

Alternative splicing results in the production of two polypeptides of different size, the larger isoform having additional amino acids at the C-terminus. Both isoforms can regulate Rubisco activity, but the larger isoform is the most sensitive to ADP inhibition and is the only isoform that is redox-regulated. This sensitivity to ADP depends on the presence of two cysteine residues in the additional C-terminal amino acids, which, via a disulphide bond, are involved in the redox regulation of Rubisco activase. The larger isoform can also alter the activity of the smaller isoform, which is proposed to create a 'fine-tuning' mechanism for the regulation of Rubisco activase.

FCA

The *FCA* gene, which encodes an RNA-binding protein, promotes photoperiod-independent flowering in *Arabidopsis*. The *FCA* gene is alternatively processed at two positions to yield four different transcripts: α , β , γ , and δ . Transcript α constitutes less than 1% of the *FCA* mRNA, β about 55%, γ about 35% and δ about 10%. Transcript γ is thought to be the only transcript that encodes the full-length protein. Transcript α results from the inclusion of intron 3 in the transcript and results in the premature termination of translation. Transcript β results from cleavage and polyadenylation within intron 3, generating a protein identical to that from transcript α . Transcript γ results from the correct splicing of intron 3. Transcript δ results from alternative splicing in intron 13. The purpose of the alternative splicing/processing appears to be to regulate the spatial and temporal accumulation of the functional FCA protein.

legumin mRNA were increased by more than 20-fold, but transcription was only increased by about twofold.

Introns from plant genes are known to be capable of increasing gene expression from both homologous and heterogenous promoters. First intron sequences are often used in monocot transformation experiments to increase transgene expression. The precise mechanism by which introns increase gene expression is unclear, but it is thought to involve an effect on RNA processing/splicing, as mutation of splice sites results in lower expression levels.

BOX 1.4 Repression of translation by upstream open reading frames (ORFs)

Short open reading frames are present in the 5' UTR of several plant transcription-factor gene families. The translation initiation codon of these short open reading frames does not facilitate efficient translation initiation. These upstream short ORFs repress translation of downstream ORFs, with the effect that the transcription factor itself (the downstream ORF) is only translated when the competition for ribosomes is not intense. Upstream ORFs (uORFs) only occur in some members of each family and thus provide a mechanism for the differential regulation of transcription factor activity. The plant *S*-adenosyl-L-methionine decarboxylase (AdoMetDC) transcript also contains uORFs, both a 'tiny' uORF (2–3 codons) and a 'small', highly conserved uORF (50–54 codons). The product of the small uORF, possibly an RNA-binding protein, represses translation of the AdoMetDC ORF.

Upstream ORFs may also be involved in regulating gene expression in response to specific signals. For example, the 5' UTR of an *Arabidopsis* bZIP protein mRNA (*ATB2*), which contains four overlapping uORFs, represses translation in response to sucrose, but not other sugars.

Translation

It is becoming increasingly apparent that translational regulation in response to various signals is an extremely important step in regulating gene expression (Box 1.4 gives some more specific examples of translational regulation of gene expression). The efficiency of translation initiation and elongation can be influenced by a variety of factors.

The presence of the 5' cap and 3' poly(A)-tail in mRNA act synergistically to enhance translation. The efficiency of translation initiation depends on several factors, the most important ones being the distance of the AUG translation initiation codon from the 5' end of the mRNA and the secondary structure around the initiation codon. Computer analysis has allowed consensus sequences for translation initiation codons to be identified in both monocot and dicot plants:

monocots	(A/C)	(A/G)	(A/C)	C <u>AUG</u> GC
dicots	A	A	(A/C)	A <u>AUG</u> GC

(where the translation initiation codon, AUG, is underlined).

Further analysis has demonstrated that the nucleotides in positions +4 to +11 are also involved in the regulation of gene expression.

Translation may also be promoted by the binding of specific proteins to the mRNA.

Post-translational modification

Many proteins have to be modified after translation in order to become functional. Proteins can undergo a great variety of possible post-translation

Table 1.3 Protein localisation signals

Organelle/subcellular compartment	Signal
Endoplasmic reticulum (ER)	HDEL/KDEL/RDEL motif ^a
Vacuole	LQRD (vacuolar targeting sequence) ^b
Nucleus	Nuclear localisation signal (NLS) ^c
Peroxisomes/glyoxisomes	SKL motif (S/A/C) (K/R/H)L ^d
Chloroplasts	N-terminal transit peptide ^e
Mitochondria	N-terminal transit peptide ^f

^a Usually present in the C-terminus. Addition of this motif to other proteins may not be sufficient for retention in the ER but will slow transport.

^b May not be sufficient for targeting. Proper targeting may require protein folding to bring 'signal patches' to the surface of the protein.

^c The NLS usually consists of a small region (located anywhere in the protein) rich in basic amino acids (Arg and Lys). Other types of NLS are also found. NLSs are not absolutely conserved in either size or composition.

^d The addition of a C-terminus SKLL motif to reporter genes targets the protein to the peroxisome in rice.

^e The composition of the transit peptide varies, although they usually have no (or few) acidic residues but are rich in serine and threonine. Various computer programmes are available to predict chloroplast transit peptides in a protein sequence. The targeting peptide is cleaved from the mature protein. Although the chloroplast has a genome, the majority of chloroplastic proteins are encoded in the nucleus, necessitating transport into the chloroplast.

^f The composition of transit peptides directing proteins to the mitochondrion is variable, although there may be some regions of homology. They generally lack acidic residues (as is the case with chloroplast transit peptides) and are enriched in alanine, leucine, arginine and serine. As with chloroplast transit peptides, computer programmes are available to predict mitochondrial transit peptides. The transit peptide is cleaved from the mature protein.

(Data taken with permission from Wallace, T. P. and Howe, C. J. (1993). Plant organellar targeting sequences. In *Plant Molecular Biology Labfax* (ed. R. R. D. Croy), pp. 287–92. BIOS Scientific, Oxford.)

modifications. Signal peptides (which direct proteins to certain subcellular structures or organelles) may be proteolytically cleaved from the precursor protein. Specific residues may be modified by phosphorylation, glycosylation or acetylation. Reversible changes in these modifications (particularly phosphorylation and acetylation) are used to regulate the activity and location of a large number of proteins.

Localisation

In order to function properly, proteins have to be localised in the correct subcellular compartment. Specific signals (Table 1.3) have been identified that are responsible for targeting proteins to the nucleus, chloroplasts, mitochondria, vacuoles, peroxisomes and for retention in the endoplasmic reticulum (ER).

Protein turnover

Most proteins present in the cytosol of cells are stable and have relatively long half-lives (days). However, some proteins (which are usually coded for

by short-lived mRNAs) are rapidly turned-over due to the presence of particular amino-acid residues or motifs, which target these proteins for rapid proteolysis.

The N-terminal amino acid is known to influence protein stability. Although all proteins are, initially, synthesised with a methionine as the N-terminal amino acid, this is often subsequently removed. A class of enzymes called aminoacyl-tRNA-protein transferases can also add other amino acids to the N-terminus of certain proteins. Some amino acids (such as Met, Gly, Ser, Thr and Val) are found to stabilise proteins, whilst others (such as Arg, Asn, Asp, Gln, Glu, His, Ile, Leu, Lys, Phe, Trp and Tyr) result in rapid degradation. The N-terminus amino acid may also be modified by acetylation, which increases the stability of the protein.

Conclusions

In conclusion, gene expression can be controlled at many points. The expression of particular genes may be controlled at multiple steps to ensure the correct regulation. Specific stimuli that influence gene expression, such as environmental signals and plant hormones, may also exert their effects at more than one stage in the gene expression process. This multiplicity of controls allows gene expression to be finely regulated and to respond in a flexible way to the needs of the plant cell. Box 1.5 uses the example of methyl jasmonate to illustrate how one compound can control gene expression in different ways.

Implications for plant transformation

The introduction and stable integration of transgenes into the genome of the host plant is only the first step in successfully manipulating a plant. The transgenes have to be expressed in the appropriate fashion, both spatially and temporally. The transcript has to be properly processed and the protein product has to be modified appropriately and targeted to the correct cellular compartment. In order for these prerequisites to be achieved, considerable effort may need to be put into the design of a transgene before it is introduced into the plant.

In some cases the decisions are fairly simple. If a high level of expression of a cytoplasmic protein is required, and the timing and location of expression are not important, well-characterised promoters such as the 35S promoter from cauliflower mosaic virus could be used (see also Chapter 4). However, in many cases such a simple constitutive pattern of expression is not what is required. Alternative control sequences, either in the form of promoter fragments from other genes or as specific sequence elements, can be used to drive transgene expression in a characteristic fashion (Box 1.6 gives a simple introduction to how suitable promoter fragments or sequence elements can be identified).

BOX 1.5 The control of gene expression by methyl jasmonate

Methyl jasmonate (MeJA) is a cyclopentanone compound synthesised from linolenic acid. It is one of the main fragrances of jasmine oil, and a flavour component of black tea. However, MeJA is also involved in a variety of physiological processes, such as embryogenesis, defence against wounding and pathogen attack as well as osmotic stress. It induces the expression of a large number of sequences, termed 'JIPs' (jasmonate-induced protein). As well as inducing the expression of some sequences, MeJA has a profound negative effect on the expression of other genes, particularly those involved in photosynthesis. This may be part of the plant's defence mechanism to wounding, as photosynthesis imposes a heavy energetic burden on the wounded plant.

Transcriptional control

MeJA negatively regulates the transcription of a range of nuclear-encoded genes for photosynthetic proteins, such as the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (*rbcS* gene) and several light-harvesting chlorophyll-protein complex apoproteins. This negative effect on transcription is not universal, as *JIP* genes continue to be transcribed.

Post-transcriptional control

MeJA can, by alternative processing, cause modification of the 5' end of the transcript of the plastid-encoded Rubisco large-subunit gene. MeJA-induced alternative processing of the primary transcript results in a transcript with a 94-base untranslated region as opposed to a 'normal' 59-base untranslated region. This extension of the untranslated region results in an inhibition of translation initiation, due to the presence of a region with homology to the 3' terminus of both 16S and 18S rRNA. MeJA also seems to selectively destabilise some transcripts, such as those for *rbcS*.

Translational control

MeJA represses the synthesis of most pre-existing proteins. In the case of proteins involved in photosynthesis this is because their transcripts are associated with small polysomes (or the non-polysomal fraction) and is indicative of an inhibition of translation initiation. This displacement from polysomes appears to be a specific rather than a general effect, as JIP transcripts remain associated with all polysome fractions and are translated normally.

For example, the transgene may be required to be expressed in response to some stimulus, either external or internal, or may need to be expressed in a particular tissue. A few examples of the findings from promoter analyses are given below to illustrate how different promoter elements can be used to drive transgene expression, and Table 1.4 lists some *cis*-elements involved in the control of gene expression in response to various environmental and hormonal signals.

BOX 1.6 Promoter analysis

Much of plant science depends on understanding promoter function. This is particularly important in plant biotechnology when appropriate promoters have to be added to transgenes in order to drive expression in the correct temporal and spatial fashion.

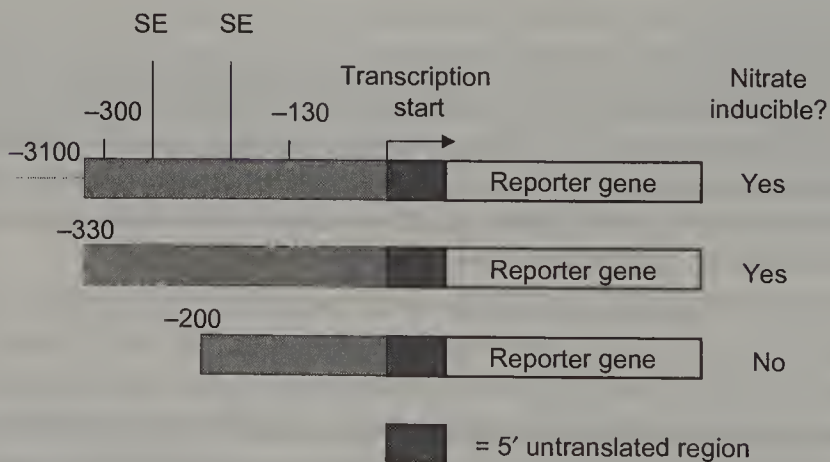
Promoters are often studied using a combination of molecular biology and plant transformation methods. A cloned DNA sequence is analysed by computer and the various features (promoter, coding region, poly(A)-site, etc.) provisionally identified. The promoter is then fused to a reporter gene (reporter genes are discussed in more detail in Chapter 4), which can be used to monitor the pattern of expression in transgenic plants. In subsequent analyses, the size of the promoter fragment used can be reduced and/or various areas can be mutated and the effect of these changes on the pattern of gene expression monitored. Various molecular techniques, such as DNA footprinting, can be used to determine which areas of the promoter binds proteins (possible transcription factors) and these can be compared with the results from the reporter gene experiments. Depending on the precise experiments used, sequence elements can be mapped on to the promoter fragments down to a resolution of 1 base pair.

Increasingly, potential *cis*-elements in cloned, sequenced DNA are identified by computer, using programs that screen libraries of known *cis*-element sequences.

Once specific sequence elements have been identified they can be fused to a minimal promoter, and assays performed to investigate whether they confer the expected pattern of expression on the minimal promoter.

The spinach nitrite reductase gene

As an example of promoter analysis we can look at a study of the promoter of the spinach nitrite reductase gene. Nitrite reductase catalyses the reduction of nitrite to



Spinach nitrite reductase gene promoter analysis. Various promoter::reporter gene fusions are shown. The promoter region used is represented by the black line, and the figures represent the lengths of promoter fragment used in each construct. The constructs used also contained some of the 5' untranslated region. The positions of the two sequence elements (SEs) identified by DNA footprinting are indicated.

BOX 1.6 Continued

ammonia in the chloroplast (see Box 5.1). The transcript is 2.3 kbp long and transcription is induced by nitrate. An analysis of the promoter region has identified several sequence motifs that are responsible for this nitrate inducibility. Various promoter::reporter gene fusions were used to identify which regions of the promoter were responsible for nitrate induction of transcription (see the figure). A region between 330 and 200 base pairs upstream of the beginning of the gene was identified by comparing the nitrate inducibility of reporter gene expression of various constructs. DNA footprinting assays of the -300 to -130 region identified two 'GATA' sequence elements (SEs) in this region with the sequence 5'-TAGATA-3' and 5'-TATCTA-3', which bound proteins in response to nitrate treatment. Additional sequences involved in nitrate inducibility were also found to be located in the 5' untranslated region of the gene.

Table 1.4 Examples of plant *cis*-regulatory elements

Stimulus and name ^a	Sequence ^b	Source ^c
<i>Light-response elements (LREs)</i>		
Box I	ACCTAACT	Parsley <i>chalcone synthase</i>
Box II	TGTGTGGTTAATATG	Pea <i>rbcS</i> -3A
Box III	ACTATTTTCACTATC	Pea <i>rbcS</i> -3A
<i>Auxin (AuxREs)</i>		
AuxRE	TGTCTC	Soybean SAUR
AuxRE	TGACCTAA	Soybean GH3
AuxRE	GTCCCAT	Pea IAA4/5
<i>Abscisic acid (ABREs)</i>		
ABRE A	GCC <u>ACGT</u> GGG ^d	Maize <i>rab28</i>
ABRE B	TCC <u>ACGT</u> CTC	Maize <i>rab28</i>
ABRE	GT <u>ACGT</u> GGCGC	Rice α -amylase
CE	ACGCGCCTCCTC	Maize <i>rab28</i>
<i>Gibberellic acid</i>	GGCCGATAACAAACTCCGGCC	Barley α -amylase
<i>Jasmonic acid</i>		
JASE 1	CGTCAATGAA	<i>Arabidopsis</i> spp. OPR1
JASE 2	CATACGTCGTCGTCAA	<i>Arabidopsis</i> spp. OPR1

^a The name of the *cis*-element is given and/or the stimulus, the transduction of which involves the *cis*-element.

^b The DNA sequence of the element is written 5' to 3'.

^c The plant species and gene name are given. Gene names are in italics for clarity.

^d The conserved ACGT core is underlined.

Abbreviations: *rbcS*, ribulose biphosphate 1,5-carboxylase/oxygenase small subunit; AuxRE, auxin-response element; SAUR, small auxin-up RNA; ABRE, abscisic acid (ABA)-response element; CE, coupling element.

Examples of promoter elements used to drive transgene expression

Light-regulated gene expression

Analysis of the promoters from light-responsive genes has identified the regions required for light-induced expression. Several consensus sequence elements have been identified in those regions shown to be necessary for promoter activity. These sequence elements are termed 'light-responsive elements' (LREs). The action of LREs is complex, as no one LRE is found in all light-regulated genes and two or more LREs must be present for light-induced expression. Some LREs may also be involved in repression of expression as opposed to induction. It has been demonstrated that pairwise combinations of any LREs are sufficient to confer light-induced expression to minimal promoters.

Abscisic acid (ABA)-induced gene expression

Abscisic acid is a plant hormone that controls the expression of a large number of genes. Promoter analysis has demonstrated that a small fragment from the promoter of some abscisic acid-induced genes is sufficient to confer ABA-inducibility to minimal promoters. ABA induction mediated by the fragment was found to depend on the presence of two sequence elements: one termed 'ABRE' (ABA response element, which in common with some other *cis*-elements has an ACGT core); and another termed the 'CE' (coupling element).

Tissue-specific expression

Analysis of the promoters of various genes expressed in specific tissues has allowed the identification of those sequence elements responsible for the pattern of expression. Seed storage proteins have been extensively studied in this respect, and sequence elements (such as the legumin box, 5'-TCCATAGCCATGCAAGCTGCA-3') responsible for localising expression have been identified.

Protein targeting

Proteins may also have to be targeted to a particular subcellular location in order to have the desired effect. As will be seen in Chapter 5, the production of herbicide-resistant plants often requires the transgene product to be targeted to the chloroplast, which can be done by adding the transit peptide from a chloroplast-targeted protein. Deliberate 'mis-targeting' of proteins to alternative subcellular locations may be of some benefit in some areas of plant biotechnology, as normal controls that limit accumulation may be circumvented.

Heterologous promoters

Several assumptions are made when considering the use of heterologous promoters to drive transgene expression. The first is that the *trans*-activating factors from one plant will recognise the *cis*-sequence elements of the heterol-

ogenous promoter. In many cases, this appears to happen, but in some cases (particularly if monocot promoter elements are used to drive transgene expression in a dicot or vice versa) it does not. It is also usually assumed, sometimes incorrectly, that the transgene will be properly spliced and processed and the protein correctly folded and modified. As the controls of gene expression become better understood it is becoming apparent that what the plant biotechnologist can achieve with the use of heterologous promoters and other tools is an approximation of the 'natural' expression profile. Fortunately, in many cases this approximation is close enough for its use in a biotechnology programme.

Genome size and organisation

It is perhaps fitting, especially in the light of recent progress in genome sequencing, to ask what we know about plant genomes at the molecular level and to consider what uses this information can be put to.

The size of the nuclear genome varies among organisms, with eukaryotic cells having an unreplicated DNA content (the C value) varying from $\sim 10^7$ to 10^{11} bp. In the broadest terms, the size of the nuclear genome reflects the complexity of the organism, so that the genome of bacteria is smaller than that of fungi, which in turn is smaller than that of animals and plants. However, this simple relationship does not always hold true, a situation known as 'the C-value paradox'. Higher plants for example, which can be assumed to have a similar degree of complexity and a similar number of genes, exhibit genome sizes that vary by several orders of magnitude (Figure 1.8), and many amphibia have C-values much larger than that of humans. It is also known that in many organisms only a small percentage of the genome actually encodes proteins. Even accounting for other regions, such as introns and the 5' and 3' regions of genes, this means that the vast majority of the DNA in some organisms is non-coding and apparently functionless.

Much of this non-coding DNA in the nuclear genome comprises repetitive DNA. Repetitive DNA consists of groups or families of similar repeated sequences and can be divided into two types: tandem repeats (or satellite DNA) and dispersed repeats. Tandem repeats are associated with particular locations, and primarily compose the centromeric (where the fibres of the mitotic spindle attach themselves) and telomeric (the 'ends' of chromosomes) regions of chromosomes. Dispersed repeats, which can in some plant species compose a significant proportion of the genome, tend to be found throughout the genome and are often related to transposable elements.

Arabidopsis and the new technologies

Recent years have seen a revolution take place in biology. Perhaps the most publicised example of this revolution is the completion of the sequencing of

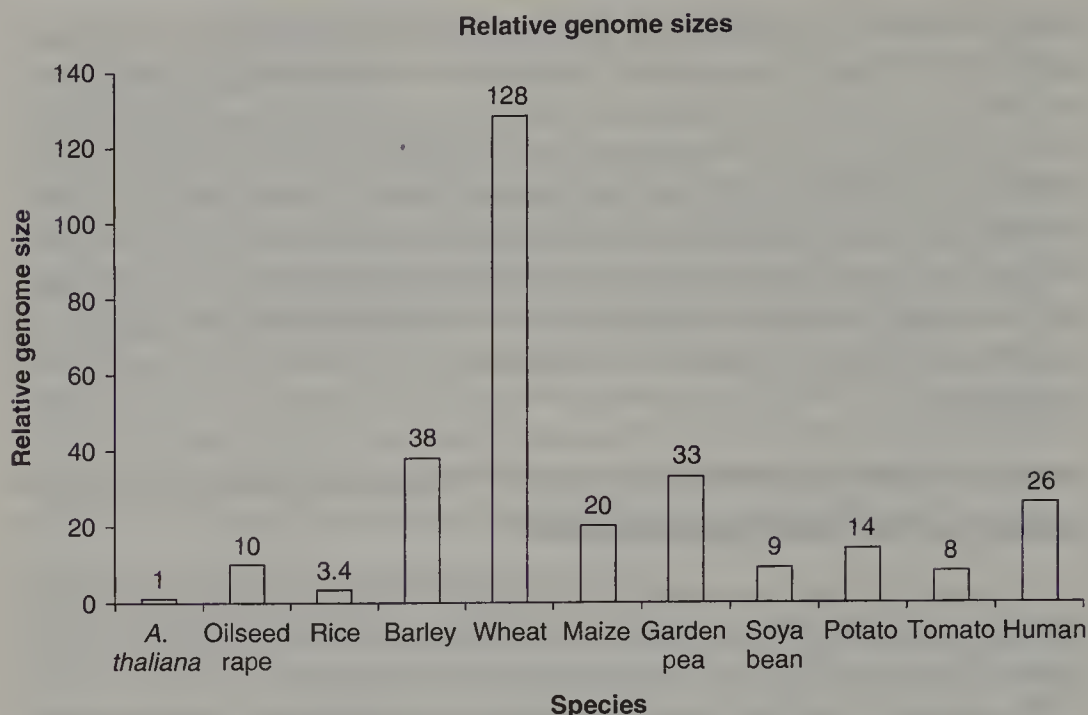


Figure 1.8 Relative genome sizes in a range of crop plants (both monocot and dicot) compared to the human genome.

the human genome, a marvellous technical achievement that should prove invaluable in improving the quality of life. Plant biology has also seen its own revolutions, some of which have admittedly followed from advances in other areas, but some of which are unique to plants (see Box 1.7). These advances in plant science may also yield tangible benefits to humanity very rapidly, given the applied focus of much plant research and the smaller number of ethical/moral issues associated with plant science.

In the light of these advances, it seems appropriate to look at the findings of these new technologies and the uses to which they can be put.

Genome-sequencing projects—technology, findings and applications

Arabidopsis: a model plant

Arabidopsis thaliana is a small, dicotyledonous cruciferous plant that has become the model species for most aspects of plant biology. It has achieved this status despite being a weed with no commercial value, although it is a member of the mustard family and is related to crops such as cabbage, cauliflower, sprouts and oilseed rape (canola).

A. thaliana does, though, have a range of attributes that make it the ideal experimental plant. It has a short life cycle, small stature and produces large

Box 1.7 The new technologies

Genomics is only one of a range of recently developed technologies that can be of use to the plant biotechnologist. Indeed, the true potential of the various genome-sequencing efforts will only be realised when they are seen as part of a co-ordinated effort involving all these new technologies.

Functional genomics

One of the most direct ways of investigating gene function is to mutate the gene so that it no longer produces a functional protein product. In plants, this can be done in a variety of ways, usually involving the introduction of a piece of 'foreign' DNA into the gene. The advantage of this type of 'insertional mutagenesis' is that it is relatively straightforward to locate the point of insertion. As this process is more or less random, large 'libraries' of mutated plants, each of which has a mutation in a particular gene, have been established. These mutants can be used to investigate the function of particular genes in a process known as 'reverse genetics', where the genotype is known and the resultant phenotype identified.

Transcriptomics

Transcriptomics is a technology developed to enable the genome-wide analysis of gene expression patterns. This is a highly automated process and allows experimenters to identify all the changes in gene expression that occur, for example, in response to a particular treatment. Identifying genes that show altered expression patterns due to a particular treatment, or at a particular developmental stage, is one way of identifying genes that are potentially useful to the plant biotechnologist. For instance, genes that are found to be expressed in high amounts following an attack by a pest may well be involved in the plant's defence against that pest.

Proteomics

Attempts to study the proteome (the total protein complement encoded by the genome) are in their infancy, and, at present, analysis is restricted to the relatively abundant proteins. The advantage of this technique is that it actually looks at the 'functional' outcome of gene expression (i.e. proteins) and can be used to study post-translational modifications. This technique attempts to analyse the next level of gene expression after transcriptomics.

Metabolomics

This allows the study of the chemical constituents of the cell using analytical chemistry. Different classes of metabolites (such as plant hormones) can be studied, and differences and similarities correlated with desirable traits. One interesting outcome of metabolomics is the demonstration that genetic engineering really is a quite precise science, with far fewer differences between a transformed plant and a control plant than between different varieties of the same plant.

numbers of offspring, making it ideally suited for genetic and mutational analysis. It is also easily transformed and, most importantly for genome sequencing initiatives, it has one of the smallest genomes in higher plants (125 Mb). *A. thaliana* was therefore chosen to be the first plant to have its complete genome sequenced

'The Arabidopsis Genome Sequencing Initiative'—successful, international, publicly funded, freely available

December 2000 saw the publication of the complete genome sequence of *A. thaliana*, the culmination of years of effort by a multinational team. The international collaborative 'Arabidopsis Genome Initiative' (AGI) began sequencing the genome in 1996. A total of 115.4 million bases were sequenced with a previously unmatched accuracy of between 99.99 and 99.999%, thanks to advances in technology. The remaining 10 Mb represents repeats and/or areas that are difficult to sequence. It is surely one of the best examples of what can be achieved by publicly funded science given the focus and resources.

The statistics of the *A. thaliana* genome are summarised in Table 1.5. There are predicted to be about 25 500 genes in the *Arabidopsis* genome. This is more than for some multicellular eukaryotes (*Caenorhabditis elegans*, a nematode, has about 19 000 genes and the fruit fly *Drosophila melanogaster* has about 13 600 genes) but less than the 30 000–40 000 predicted for humans.

It is important to remember that this analysis is based on computer predictions. The figures in Table 1.5 and comparison with other eukaryotes perhaps give a misleading impression about the relative 'complexities' of these organisms. There is extensive gene duplication in *Arabidopsis*, so that the total number of distinct protein types is only about 11 600, a similar figure to that estimated for both *C. elegans* and *Drosophila melanogaster*.

Functional analysis

About 70% of the predicted genes in *Arabidopsis* can be assigned a function based on their sequence similarity to proteins of known function in all organisms (Figure 1.9). However, it is important to note that only about 9% of the predicted genes have been characterised experimentally.

The 30% of genes that cannot be assigned a function comprise both plant-specific proteins and proteins with a high degree of similarity to genes of unknown function from other organisms.

Arabidopsis has a relatively high proportion of genes involved in metabolism and defence. This is likely to reflect the fact that *Arabidopsis* is a sessile autotroph. Within each class of protein, the proportion of *Arabidopsis* proteins having identified equivalents in other eukaryotes varies. Only 8–23% of the proteins involved in transcriptional control have counterparts in other eukaryotes, whereas the figure is 48–60% for proteins involved in protein synthesis.

Table 1.5 A synopsis of the findings from the Arabidopsis genome sequencing initiative

	Chr1 ^a	Chr2	Chr3	Chr4	Chr5	Total
Length (kbp)	29,105	19,647	23,173	17,550	25,953	115,410
Top arm (kbp)	14,449	3,697	13,590	3,052	11,132	
Bottom arm (kbp)	14,656	16,040	9,582	14,498	14,803	
Number of genes	6,543	4,036	5,220	3,825	5,874	25,498
Exons:						
Number	35,482	19,631	26,570	20,073	31,226	132,982
Total length (kbp)	8,773	5,100	6,655	5,151	7,571	33,249
% of total length	30	26	29	29	29	29
Average per gene	5.4	4.9	5.1	5.2	5.3	
Average size (bp)	247	259	250	256	242	
Introns:						
Number	28,939	15,595	21,350	16,248	25,352	107,484
Total length (kbp)	4,829	2,768	3,398	3,031	4,030	18,056
% of total	17	14	15	17	16	
Average size (bp)	168	177	159	186	159	
Gene density (kbp per gene)	4.0	4.9	4.5	4.6	4.4	

^aChr = chromosome.
Data from The Arabidopsis Genome Initiative (2000). *Nature* 408:796–815.

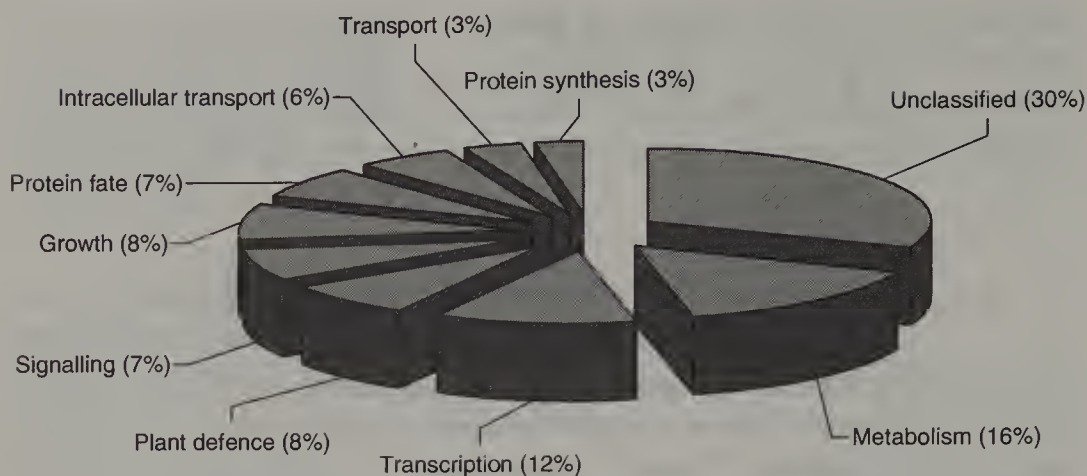


Figure 1.9 Functional analysis of the *Arabidopsis* genome. The proportion of genes assigned to different functional classes is shown. (Data from The *Arabidopsis* Genome Initiative (2000).)

Biotechnological implications of the 'Arabidopsis Genome Sequencing Initiative'

The data that has emerged from the 'Arabidopsis Genome Sequencing Initiative' has proved to be extremely useful in many areas of plant biology. Many biotechnology companies are already taking advantage of the results of the *Arabidopsis* genome sequencing effort. Although *Arabidopsis* is not a crop plant, many of the processes that are of interest to the plant biotechnologist (e.g. disease and pest resistance, and stress tolerance) share common features in both *Arabidopsis* and crop plants. *Arabidopsis* provides the ideal tool for plant scientists to understand these processes, an essential step before they can be successfully manipulated using biotechnology. This aspect is becoming increasingly important as public awareness and concern about plant biotechnology increases and efforts are made to reduce the number of proteins from other, completely different, organisms used in plant biotechnology.

Crop plant genome sequencing

The completion of the 'Arabidopsis Genome Sequencing Initiative' is without doubt a major milestone in plant biology, and has already demonstrated the usefulness of *Arabidopsis* as a model system. However, it is also beyond doubt that *Arabidopsis* will not feed the world. Most of the world's population depend on cereals, and rice in particular, to supply their calorific and nutritional needs. Cereals are monocotyledonous plants and therefore exhibit important differences from *Arabidopsis* (a dicot) in some aspects of their development. Particularly important in this respect are the differences in seed/endosperm development between monocots and dicots. Therefore, although *Arabidopsis* is, in many respects, the ideal model system, the elucidation of its genome may not provide clues to some specific problems. Rice also has some major advantages that make it an ideal choice for the next plant to have its genome se-

Box 1.8 Synteny and colinearity

Plant genomes show a huge variation in chromosome size and number, which is reflected by a size range from ~100 Mb found in some Cruciferae (such as *Arabidopsis*) to ~100 000 Mb in some members of the lily family. However, the gene number (and type) does not show such marked differences. Much of the difference in genome size between different plants is due to the amount of repetitive DNA present in the genome (although polyploidy may also account for some of the difference). Changes in genome size can be a (relatively) frequent event during evolution, but these changes do not appear to be associated with large changes in gene content and gene order. This conservation of gene content and order is termed 'colinearity', although recently the use of the term 'synteny' (derived from the Greek *syn*, meaning together and *taenia*, meaning ribbon) has increased to cover both colinearity and homeology (the residual homology of previously completely homologous chromosomes). Colinearity or synteny is a startling phenomenon. *Arabidopsis* (dicots) and rice (a monocot) diverged from a common ancestor about 200 million years ago, and yet there is low, but detectable, synteny between the two species. Where synteny becomes much more obvious, and useful, is when the genomes of various cereals are aligned and compared. Agronomically important traits (grain yield, pest resistance, etc.) identified in other cereals, such as maize, can be associated with the rice genome sequence, allowing identification of the genes responsible for the trait. Many cereal quantitative-trait loci (QTLs) have already been placed on the rice genome map.

quenced. Perhaps most importantly, the rice genome is only about four times larger than that of *Arabidopsis*, making it one of the smallest cereal crop genomes. Good genetic maps also exist for rice, and genetic mapping experiments had already shown that the gene content and order (gene order is explained in Box 1.8) were conserved in cereals.

The rice genome

There is a major multinational effort (the IRGSP—the International Rice Genome Sequencing Project) to sequence the genome of rice, which is expected to be completed in draft form by the end of 2002. This initiative has recently been given a tremendous boost by the publication in April of 2002 of draft sequences of the two major subspecies of rice. These draft genome sequences have been produced in a relatively short time using new technologies, but do not have the accuracy of sequences based on more conventional technology. Nevertheless, the data will prove extremely useful in not only speeding up the multinational rice genome effort but also in applying genome sequencing results to real problems in plant biotechnology.

The usefulness of the rice genome sequence can be illustrated simply by looking at some statistics. Rice is predicted to have 33 000–55 000 genes (depending on the sequence and gene prediction method used); so, whilst over 80% of the predicted proteins in *Arabidopsis* are also present in rice, more than 45% of the predicted rice genes are not present in *Arabidopsis* (although

this figure may prove to be an overestimate). Therefore, there are significant differences in not only gene number but also gene type between rice and *Arabidopsis* that may reflect the different development and physiologies of monocots and dicots. However, as was expected, there has proved to be considerable conservation of gene type and order in the cereals, with 98% of the known maize, wheat and barley proteins being present in rice.

Other crops

There is considerable debate as to which other, if any, crop plants should have their genome sequenced. Many consider that the data from *Arabidopsis* and rice genome sequencing initiatives should be sufficient, especially given the considerable synteny, or co-linearity, that exists within monocots and dicots. Others disagree, and argue that the genomes of other important crops (such as wheat and maize, and also, for example, trees), should also be sequenced. Technological advances are making genome sequencing projects cheaper and quicker, but the big disadvantage with these crops is that their genomes are extremely large and complicated (see Figure 1.8).

Findings from genome sequencing initiatives are not only applicable to projects involving genetic manipulation. The findings should prove invaluable in conventional breeding programmes allowing, for example, specific traits to be easily followed during breeding programmes.

Summary

This chapter has briefly looked at various aspects of what might called 'the genetic systems' of plants. It is not intended to be an all-encompassing look at plant genomes and gene regulation, whole volumes being required to cover such complex subjects. It should, however, provide a framework sufficient to enable the more molecular aspects of the remainder of the book to be understood without reference to other books.

Understanding how gene expression is controlled is key to any strategy to genetically manipulate plants. Transgenes (those 'foreign' genes introduced into the plant genome by plant transformation) are subject to control in the same way as the endogenous genes of the plant. It is interesting to observe the synergy that has developed between the need to understand how transgenes are regulated and the use of plant transformation strategies as a tool to elucidate the controls of gene expression.

As the controls of gene expression become better understood, this information is fed through into improvements to plant transformation strategies. Some of the most obvious developments are being made in the design of plant transformation vectors and the modification of the transgene itself so as to ensure optimal expression. Some of these improvements are discussed in subsequent chapters, particularly Chapter 4.

The rice and *Arabidopsis* genome sequencing projects discussed in this chapter have generated a huge amount of data. Much of this data is of interest because of the light it sheds on fundamental processes, such as evolution. The data is, however, also contributing hugely to our understanding of how plant gene expression is controlled and integrated. Coupled with information from fundamental studies on gene function, these advances are making possible the development of the next generation of genetically mod-

ified (GM) crops. These new-generation crops, which to a large extent will be based on modifying the regulation or complement of endogenous genes, bring about the possibility of manipulating complex agronomic traits like yield. Such crops are discussed briefly in Chapter 12. In Chapter 9 we will see how the manipulation of transcription factor expression, in order to influence the expression of a number of genes, can be used in plant biotechnology. This type of approach will become more common in plant biotechnology as gene expression controls, genome structures and gene function become better understood.

Further reading

Gene structure and expression

- Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J. D. (1989). *Molecular biology of the cell*. Garland, New York.
- Bailey-Serres, J. (1999). Selective translation of cytoplasmic mRNAs in plants. *Trends in Plant Science*, 4, 142–8.
- Bailey-Serres, J., Rochaix, J.-D., Wassenegger, M. and Filipowicz, W. (1999). Plants, their organelles, viruses and transgenes reveal the mechanisms and relevance of post-transcriptional processes. *EMBO Journal*, 18, 5153–8.
- Berk, A. J. (1999). Activation of RNA polymerase II transcription. *Current Opinion in Cell Biology*, 11, 330–5.
- Buchanan, B. B., Gruissem, W. and Jones, R. L. (2000). *Biochemistry and molecular biology of plants*. American Society of Plant Physiologists, Rockville MD.
- Buratowski, S. (2000). Snapshots of RNA polymerase II transcription initiation. *Current Opinion in Cell Biology*, 12, 320–5.
- Eckardt, N. A. (2002). Alternative splicing and the control of flowering time. *Plant Cell*, 14, 743–7.
- Gutierrez, R. A., MacIntosh, G. C. and Green, P. J. (1999). Current perspectives on mRNA stability in plants: multiple levels and mechanisms of control. *Trends in Plant Science*, 4, 429–38.
- Kornberg, R. D. (1999). Eukaryotic transcriptional control. *Trends in Cell Biology*, 9, M46–M49.
- Li, G., Hall, T. C. and Holme-Davis, R. (2002). Plant chromatin: development and gene control. *BioEssays*, 24, 234–43.
- Lorkovic, Z. J., Wieczorek Kirk, D. A. W., Lambermon, M. H. L. and Filipowicz, W. (2000). Pre-mRNA splicing in higher plants. *Trends in Plant Science*, 5, 160–7.
- Lukaszewicz, M., Feuermann, M., Jerouville, B., Stas, A. and Boutry, M. (2000). *In vivo* evaluation of the context sequence of the translation initiation codon in plants. *Plant Science* 154, 89–98.
- Maier, R. M., Neckermann, K., Igloi, G. L. and Kossel, H. (1995). Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *Journal of Molecular Biology* 251, 614–28.
- Miller, W. A., Waterhouse, P. M., Brown, J. W. S. and Browning, K. S. (2001). The RNA world in plants: post-transcriptional control III. *Plant Cell*, 13, 1710–17.
- Reinbothe, S., Mollenhauer, B. and Reinbothe, C. (1994). JIPS and RIPS: the regulation of gene expression by jasmonates in response to environmental cues and pathogens. *Plant Cell*, 6, 1197–209.

Sawant, S. V., Kiran, K., Singh, P. K. and Tuli, R. (2001). Sequence architecture downstream of the initiator codon enhances gene expression and protein stability in plants. *Plant Physiology* 126, 1630–6.

Biome (searchable links)—web-link 1.1: [http:// bioresearch.ac.uk](http://bioresearch.ac.uk)

ExPASy—web-link 1.2: <http://www.expasy.ch>

NCBI on-line books—web-link 1.3:

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Books>

Genome sequencing projects

Bevan, M., Mayer, K., White, O., Eisen, J. A., Preuss, D., Bureau, T., Salzberg, S. L. and Mewes, H.-W. (2001). Sequence and analysis of the *Arabidopsis* genome. *Current Opinion in Plant Biology* 4, 105–10.

Goff, S. A., Ricke, D., Lan, T.-H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B. M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.-L., Chen, L., Cooper, B., Park, S., Wood, T. C., Mao, L., Quail, P., Wing, R., Dean, P., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller, R. M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A. and Briggs, S. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp *japonica*). *Science*, 296, 92–100.

The *Arabidopsis* genome initiative: sequence and analysis of the flowering plant *Arabidopsis thaliana*. (2000). *Nature* 408, 796–815.

Yu, J., Hu, S., Wang, J., Wong, G. K.-S., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, C., Tong, W., Cong, L., Geng, J., Han, Y., Li, L., Li, W., Hu, G., Huang, X., Li, W., Li, J., Liu, Z., Li, L., Liu, J., Qi, Q., Liu, J., Li, L., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Zhang, J., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Ren, X., Chen, X., He, J., Liu, D., Tan, W., Tioan, C., Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W., Li, P., Chen, W., Wang, X., Zhang, Y., Hu, J., Wang, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Li, G., Liu, S., Tao, M., Wang, J., Zhu, L., Yuan, L. and Yang, H. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp *indica*). *Science*, 296, 79–92.

Nature genome gateway—web-link 1.4: <http://www.nature.com/genomics>

The Arabidopsis Information Resource—web-link 1.5: <http://www.arabidopsis.org>

The Institute for Genomic Research—web-link 1.6: <http://www.tigr.org>

2

Plant tissue culture

Introduction

Most methods of plant transformation applied to GM crops require that a whole plant is regenerated from isolated plant cells or tissue which have been genetically transformed. This regeneration is conducted *in vitro* so that the environment and growth medium can be manipulated to ensure a high frequency of regeneration. In addition to a high frequency of regeneration, the regenerable cells must be accessible to gene transfer by whatever technique is chosen (gene transfer methods are described in Chapter 3). The primary aim is therefore to produce, as easily and as quickly as possible, a large number of regenerable cells that are accessible to gene transfer. The subsequent regeneration step is often the most difficult step in plant transformation studies. However, it is important to remember that a high frequency of regeneration does not necessarily correlate with high transformation efficiency.

This chapter will consider some basic issues concerned with plant tissue culture *in vitro*, particularly as applied to plant transformation. It will also look at the basic culture types used for plant transformation and cover some of the techniques that can be used to regenerate whole transformed plants from transformed cells or tissue.

Plant tissue culture

Practically any plant transformation experiment relies at some point on tissue culture. There are some exceptions to this generalisation (Chapter 3 will look at some), but the ability to regenerate plants from isolated cells or tissues *in vitro* underpins most plant transformation systems.

Plasticity and totipotency

Two concepts, plasticity and totipotency, are central to understanding plant cell culture and regeneration.

Plants, due to their sessile nature and long life span, have developed a greater ability to endure extreme conditions and predation than have animals. Many of the processes involved in plant growth and development adapt to environmental conditions. This plasticity allows plants to alter their metabolism, growth and development to best suit their environment. Particularly important aspects of this adaptation, as far as plant tissue culture and regeneration are concerned, are the abilities to initiate cell division from almost any tissue of the plant and to regenerate lost organs or undergo different developmental pathways in response to particular stimuli. When plant cells and tissues are cultured *in vitro* they generally exhibit a very high degree of plasticity, which allows one type of tissue or organ to be initiated from another type. In this way, whole plants can be subsequently regenerated.

This regeneration of whole organisms depends upon the concept that all plant cells can, given the correct stimuli, express the total genetic potential of the parent plant. This maintenance of genetic potential is called 'totipotency'. Plant cell culture and regeneration do, in fact, provide the most compelling evidence for totipotency.

In practical terms though, identifying the culture conditions and stimuli required to manifest this totipotency can be extremely difficult and it is still a largely empirical process.

The culture environment

When cultured *in vitro*, all the needs, both chemical (see Table 2.1) and physical, of the plant cells have to met by the culture vessel, the growth medium and the external environment (light, temperature, etc.). The growth medium has to supply all the essential mineral ions required for growth and development. In many cases (as the biosynthetic capability of cells cultured *in vitro* may not replicate that of the parent plant), it must also supply additional organic supplements such as amino acids and vitamins. Many plant cell cultures, as they are not photosynthetic, also require the addition of a fixed carbon source in the form of a sugar (most often sucrose). One other vital component that must also be supplied is water, the principal biological solvent. Physical factors, such as temperature, pH, the gaseous environment, light (quality and duration) and osmotic pressure, also have to be maintained within acceptable limits.

Plant cell culture media

Culture media used for the *in vitro* cultivation of plant cells are composed of three basic components:

- (1) essential elements, or mineral ions, supplied as a complex mixture of salts;
- (2) an organic supplement supplying vitamins and/or amino acids; and
- (3) a source of fixed carbon; usually supplied as the sugar sucrose.

Table 2.1 Some of the elements important for plant nutrition and their physiological function. These elements have to be supplied by the culture medium in order to support the growth of healthy cultures *in vitro*

Element	Function
Nitrogen	Component of proteins, nucleic acids and some coenzymes Element required in greatest amount
Potassium	Regulates osmotic potential, principal inorganic cation
Calcium	Cell wall synthesis, membrane function, cell signalling
Magnesium	Enzyme cofactor, component of chlorophyll
Phosphorus	Component of nucleic acids, energy transfer, component of intermediates in respiration and photosynthesis
Sulphur	Component of some amino acids (methionine, cysteine) and some cofactors
Chlorine	Required for photosynthesis
Iron	Electron transfer as a component of cytochromes
Manganese	Enzyme cofactor
Cobalt	Component of some vitamins
Copper	Enzyme cofactor, electron-transfer reactions
Zinc	Enzyme cofactor, chlorophyll biosynthesis
Molybdenum	Enzyme cofactor, component of nitrate reductase

For practical purposes, the essential elements are further divided into the following categories:

- (1) macroelements (or macronutrients);
- (2) microelements (or micronutrients); and
- (3) an iron source.

Complete, plant cell culture medium is usually made by combining several different components, as outlined in Table 2.2.

Media components

It is useful to briefly consider some of the individual components of the stock solutions.

Macroelements

As is implied by the name, the stock solution supplies those elements required in large amounts for plant growth and development. Nitrogen, phosphorus, potassium, magnesium, calcium and sulphur (and carbon, which is added separately) are usually regarded as macroelements. These elements usually comprise at least 0.1% of the dry weight of plants.

Table 2.2 Composition of a typical plant culture medium. The medium described here is that of Murashige and Skoog (MS)^a

Essential element	Concentration in stock solution (mg/l)	Concentration in medium (mg/l)
<i>Macroelements^b</i>		
NH ₄ NO ₃	33 000	1 650
KNO ₃	38 000	1 900
CaCl ₂ .2H ₂ O	8 800	440
MgSO ₄ .7H ₂ O	7 400	370
KH ₂ PO ₄	3 400	170
<i>Microelements^c</i>		
KI	166	0.83
H ₃ BO ₃	1 240	6.2
MnSO ₄ .4H ₂ O	4 460	22.3
ZnSO ₄ .7H ₂ O	1 720	8.6
Na ₂ MoO ₄ .2H ₂ O	50	0.25
CuSO ₄ .5H ₂ O	5	0.025
CoCl ₂ .6H ₂ O	5	0.025
<i>Iron source^c</i>		
FeSO ₄ .7H ₂ O	5 560	27.8
Na ₂ EDTA.2H ₂ O	7 460	37.3
<i>Organic supplement^c</i>		
Myoinositol	20 000	100
Nicotinic acid	100	0.5
Pyridoxine-HCl	100	0.5
Thiamine-HCl	100	0.5
Glycine	400	2
<i>Carbon source^d</i>		
Sucrose	Added as solid	30 000

^a Many other commonly used plant culture media (such as Gamborg's B5 and Schenk and Hildebrandt (SH) medium) are similar in composition to MS medium and can be thought of as 'high-salt' media. MS is an extremely widely used medium and forms the basis for many other media formulations.

^b 50 ml of stock solution used per litre of medium.

^c 5 ml of stock solution used per litre of medium.

^d Added as solid.

Nitrogen is most commonly supplied as a mixture of nitrate ions (from the KNO₃) and ammonium ions (from the NH₄NO₃). Theoretically, there is an advantage in supplying nitrogen in the form of ammonium ions, as nitrogen must be in the reduced form to be incorporated into macromolecules. Nitrate ions therefore need to be reduced before incorporation. However, at high concentrations, ammonium ions can be toxic to plant cell cultures and uptake of ammonium ions from the medium causes acidification of the medium. In order to use ammonium ions as the sole nitrogen source, the medium needs to be buffered. High concentrations of ammonium ions can also cause culture

problems by increasing the frequency of vitrification (the culture appears pale and 'glassy' and is usually unsuitable for further culture). Using a mixture of nitrate and ammonium ions has the advantage of weakly buffering the medium as the uptake of nitrate ions causes OH^- ions to be excreted.

Phosphorus is usually supplied as the phosphate ion of ammonium, sodium or potassium salts. High concentrations of phosphate can lead to the precipitation of medium elements as insoluble phosphates.

Microelements

These elements are required in trace amounts for plant growth and development, and have many and diverse roles. Manganese, iodine, copper, cobalt, boron, molybdenum, iron and zinc usually comprise the microelements, although other elements such as nickel and aluminium are frequently found in some formulations.

Iron is usually added as iron sulphate, although iron citrate can also be used. Ethylenediaminetetraacetic acid (EDTA) is usually used in conjunction with the iron sulphate. The EDTA complexes with the iron so as to allow the slow and continuous release of iron into the medium. Uncomplexed iron can precipitate out of the medium as ferric oxide.

Organic supplements

Only two vitamins, thiamine (vitamin B_1) and myoinositol (considered a B vitamin) are considered essential for the culture of plant cells *in vitro*. However, other vitamins are often added to plant cell culture media for historical reasons.

Amino acids are also commonly included in the organic supplement. The most frequently used is glycine (arginine, asparagine, aspartic acid, alanine, glutamic acid, glutamine and proline are also used), but in many cases its inclusion is not essential. Amino acids provide a source of reduced nitrogen and, like ammonium ions, uptake causes acidification of the medium. Casein hydrolysate can be used as a relatively cheap source of a mix of amino acids.

Carbon source

Sucrose is cheap, easily available, readily assimilated and relatively stable and is therefore the most commonly used carbon source. Other carbohydrates (such as glucose, maltose, galactose and sorbitol) can also be used (see Chapter 3), and in specialised circumstances may prove superior to sucrose.

Gelling agents

Media for plant cell culture *in vitro* can be used in either liquid or 'solid' forms, depending on the type of culture being grown. For any culture types that require the plant cells or tissues to be grown on the surface of the medium, it must be solidified (more correctly termed 'gelled'). Agar, produced from seaweed, is the most common type of gelling agent, and is ideal for routine applications. However, because it is a natural product, the agar quality can vary from supplier to supplier and from batch to batch. For more

demanding applications (see, for instance, the section on microspore culture below and Chapter 3), a range of purer (and in some cases, considerably more expensive) gelling agents are available. Purified agar or agarose can be used, as can a variety of gellan gums.

Summary

These components, then, are the basic ‘chemical’ necessities for plant cell culture media. However, other additions are made in order to manipulate the pattern of growth and development of the plant cell culture.

Plant growth regulators

We have already briefly considered the concepts of plasticity and totipotency. The essential point as far as plant cell culture is concerned is that, due to this plasticity and totipotency, specific media manipulations can be used to direct the development of plant cells in culture.

Plant growth regulators are the critical media components in determining the developmental pathway of the plant cells. The plant growth regulators used most commonly are plant hormones or their synthetic analogues.

Classes of plant growth regulators

There are five main classes of plant growth regulator used in plant cell culture, namely:

- (1) auxins;
- (2) cytokinins;
- (3) gibberellins;
- (4) abscisic acid;
- (5) ethylene.

Each class of plant growth regulator will be briefly looked at.

Auxins

Auxins promote both cell division and cell growth. The most important naturally occurring auxin is IAA (indole-3-acetic acid), but its use in plant cell culture media is limited because it is unstable to both heat and light. Occasionally, amino acid conjugates of IAA (such as indole-acetyl-L-alanine and indole-acetyl-L-glycine), which are more stable, are used to partially alleviate the problems associated with the use of IAA. It is more common, though, to use stable chemical analogues of IAA as a source of auxin in plant cell culture media. 2,4-Dichlorophenoxyacetic acid (2,4-D) is the most commonly used auxin and is extremely effective in most circumstances. Other auxins are available (see Table 2.3), and some may be more effective or ‘potent’ than 2,4-D in some instances.

Table 2.3 Commonly used auxins, their abbreviation and chemical name

Abbreviation/name	Chemical name
2,4-D	2,4-dichlorophenoxyacetic acid
2,4,5-T	2,4,5-trichlorophenoxyacetic acid
Dicamba	2-methoxy-3,6-dichlorobenzoic acid
IAA	Indole-3-acetic acid
IBA	Indole-3-butyric acid
MCPA	2-methyl-4-chlorophenoxyacetic acid
NAA	1-naphthylacetic acid
NOA	2-naphthylacetic acid
Picloram	4-amino-2,5,6-trichloropicolinic acid

Table 2.4 Commonly used cytokinins, their abbreviation and chemical name

Abbreviation/name	Chemical name
BAP ^a	6-benzylaminopurine
2iP (IPA) ^b	[N ⁶ -(2-isopentyl)adenine]
Kinetin ^a	6-furfurylaminopurine
Thidiazuron ^c	1-phenyl-3-(1,2,3-thiadiazol-5-yl)urea
Zeatin ^b	4-hydroxy-3-methyl-trans-2-butenylaminopurine

^a Synthetic analogues.^b Naturally occurring cytokinins.^c A substituted phenylurea-type cytokinin.

Cytokinins

Cytokinins promote cell division. Naturally occurring cytokinins are a large group of structurally related (they are purine derivatives) compounds. Of the naturally occurring cytokinins, two have some use in plant tissue culture media (see Table 2.4). These are zeatin and 2iP (2-isopentyl adenine). Their use is not widespread as they are expensive (particularly zeatin) and relatively unstable. The synthetic analogues, kinetin and BAP (benzylaminopurine), are therefore used more frequently. Non-purine-based chemicals, such as substituted phenylureas, are also used as cytokinins in plant cell culture media. These substituted phenylureas can also substitute for auxin in some culture systems.

Gibberellins

There are numerous, naturally occurring, structurally related compounds termed 'gibberellins'. They are involved in regulating cell elongation, and are agronomically important in determining plant height and fruit-set. Only a few

of the gibberellins are used in plant tissue culture media, GA₃ being the most common.

Abscisic acid

Abscicic acid (ABA) inhibits cell division. It is most commonly used in plant tissue culture to promote distinct developmental pathways such as somatic embryogenesis (see also Box 2.1).

Ethylene

Ethylene is a gaseous, naturally occurring, plant growth regulator most commonly associated with controlling fruit ripening in climacteric fruits, and its use in plant tissue culture is not widespread. It does, though, present a particular problem for plant tissue culture. Some plant cell cultures produce ethylene, which, if it builds up sufficiently, can inhibit the growth and development of the culture. The type of culture vessel used and its means of closure affect the gaseous exchange between the culture vessel and the outside atmosphere and thus the levels of ethylene present in the culture.

Plant growth regulators and tissue culture

Generalisations about plant growth regulators and their use in plant cell culture media have been developed from initial observations made in the 1950s. There is, however, some considerable difficulty in predicting the effects of plant growth regulators: this is because of the great differences in culture response between species, cultivars and even plants of the same cultivar grown under different conditions.

However, some principles do hold true and have become the paradigm on which most plant tissue culture regimes are based.

Auxins and cytokinins are the most widely used plant growth regulators in plant tissue culture and are usually used together, the ratio of the auxin to the cytokinin determining the type of culture established or regenerated (see Figure 2.1). A high auxin to cytokinin ratio generally favours root formation, whereas a high cytokinin to auxin ratio favours shoot formation. An intermediate ratio favours callus production.

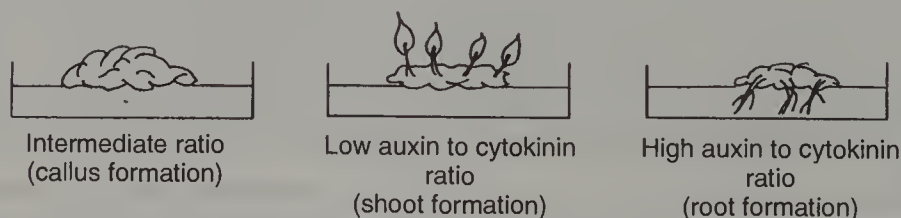


Figure 2.1 The effect of different ratios of auxin to cytokinin on the growth and morphogenesis of callus. High auxin to cytokinin ratios promote root development, low ratios promote shoot development. Intermediate ratios promote continued growth of the callus without differentiation.

Culture types

Cultures are generally initiated from sterile pieces of a whole plant. These pieces are termed 'explants', and may consist of pieces of organs, such as leaves or roots, or may be specific cell types, such as pollen or endosperm. Many features of the explant are known to affect the efficiency of culture initiation. Generally, younger, more rapidly growing tissue (or tissue at an early stage of development) is most effective.

Several different culture types most commonly used in plant transformation studies will now be examined in more detail.

Callus

Explants, when cultured on the appropriate medium, usually with both an auxin and a cytokinin, can give rise to an unorganised, growing and dividing mass of cells. It is thought that any plant tissue can be used as an explant, if the correct conditions are found. In culture, this proliferation can be maintained more or less indefinitely, provided that the callus is subcultured on to fresh medium periodically. During callus formation there is some degree of dedifferentiation (i.e. the changes that occur during development and specialisation are, to some extent, reversed), both in morphology (callus is usually composed of unspecialised parenchyma cells) and metabolism. One major consequence of this dedifferentiation is that most plant cultures lose the ability to photosynthesise. This has important consequences for the culture of callus tissue, as the metabolic profile will probably not match that of the donor plant. This necessitates the addition of other components—such as vitamins and, most importantly, a carbon source—to the culture medium, in addition to the usual mineral nutrients.

Callus culture is often performed in the dark (the lack of photosynthetic capability being no drawback) as light can encourage differentiation of the callus.

During long-term culture, the culture may lose the requirement for auxin and/or cytokinin. This process, known as 'habituation', is common in callus cultures from some plant species (such as sugar beet).

Callus cultures are extremely important in plant biotechnology. Manipulation of the auxin to cytokinin ratio in the medium can lead to the development of shoots, roots or somatic embryos from which whole plants can subsequently be produced. Callus cultures can also be used to initiate cell suspensions, which are used in a variety of ways in plant transformation studies.

Cell-suspension cultures

Callus cultures, broadly speaking, fall into one of two categories: compact or friable. In compact callus the cells are densely aggregated, whereas in friable

callus the cells are only loosely associated with each other and the callus becomes soft and breaks apart easily. Friable callus provides the inoculum to form cell-suspension cultures. Explants from some plant species or particular cell types tend not to form friable callus, making cell-suspension initiation a difficult task. The friability of callus can sometimes be improved by manipulating the medium components or by repeated subculturing. The friability of the callus can also sometimes be improved by culturing it on 'semi-solid' medium (medium with a low concentration of gelling agent).

When friable callus is placed into a liquid medium (usually the same composition as the solid medium used for the callus culture) and then agitated, single cells and/or small clumps of cells are released into the medium. Under the correct conditions, these released cells continue to grow and divide, eventually producing a cell-suspension culture. A relatively large inoculum should be used when initiating cell suspensions so that the released cell numbers build up quickly. The inoculum should not be too large though, as toxic products released from damaged or stressed cells can build up to lethal levels. Large cell clumps can be removed during subculture of the cell suspension.

Cell suspensions can be maintained relatively simply as batch cultures in conical flasks. They are continually cultured by repeated subculturing into fresh medium. This results in dilution of the suspension and the initiation of another batch growth cycle. The degree of dilution during subculture should be determined empirically for each culture. Too great a degree of dilution will result in a greatly extended lag period or, in extreme cases, death of the transferred cells.

After subculture, the cells divide and the biomass of the culture increases in a characteristic fashion, until nutrients in the medium are exhausted and/or toxic by-products build up to inhibitory levels—this is called the 'stationary phase'. If cells are left in the stationary phase for too long, they will die and the culture will be lost. Therefore, cells should be transferred as they enter the stationary phase. It is therefore important that the batch growth-cycle parameters are determined for each cell-suspension culture.

Protoplasts

Protoplasts are plant cells with the cell wall removed. Protoplasts are most commonly isolated from either leaf mesophyll cells or cell suspensions, although other sources can be used to advantage. Two general approaches to removing the cell wall (a difficult task without damaging the protoplast) can be taken—mechanical or enzymatic isolation.

Mechanical isolation, although possible, often results in low yields, poor quality and poor performance in culture due to substances released from damaged cells.

Enzymatic isolation is usually carried out in a simple salt solution with a high osmoticum, plus the cell wall degrading enzymes. It is usual to use a mix

of both cellulase and pectinase enzymes, which must be of high quality and purity.

Protoplasts are fragile and easily damaged, and therefore must be cultured carefully. Liquid medium is not agitated and a high osmotic potential is maintained, at least in the initial stages. The liquid medium must be shallow enough to allow aeration in the absence of agitation. Protoplasts can be plated out on to solid medium and callus produced. Whole plants can be regenerated by organogenesis or somatic embryogenesis from this callus.

Protoplasts are ideal targets for transformation by a variety of means.

Root cultures

Root cultures can be established *in vitro* from explants of the root tip of either primary or lateral roots and can be cultured on fairly simple media. The growth of roots *in vitro* is potentially unlimited, as roots are indeterminate organs. Although the establishment of root cultures was one of the first achievements of modern plant tissue culture, they are not widely used in plant transformation studies.

Shoot tip and meristem culture

The tips of shoots (which contain the shoot apical meristem) can be cultured *in vitro*, producing clumps of shoots from either axillary or adventitious buds. This method can be used for clonal propagation.

Shoot meristem cultures are potential alternatives to the more commonly used methods for cereal regeneration (see the Case study below) as they are less genotype-dependent and more efficient (seedlings can be used as donor material).

Embryo culture

Embryos can be used as explants to generate callus cultures or somatic embryos. Both immature and mature embryos can be used as explants. Immature, embryo-derived embryogenic callus is the most popular method of monocot plant regeneration.

Microspore culture

Haploid tissue can be cultured *in vitro* by using pollen or anthers as an explant. Pollen contains the male gametophyte, which is termed the 'microspore'. Both callus and embryos can be produced from pollen. Two main approaches can be taken to produce *in vitro* cultures from haploid tissue.

The first method depends on using the anther as the explant. Anthers (somatic tissue that surrounds and contains the pollen) can be cultured on solid medium (agar should not be used to solidify the medium as it contains

inhibitory substances). Pollen-derived embryos are subsequently produced via dehiscence of the mature anthers. The dehiscence of the anther depends both on its isolation at the correct stage and on the correct culture conditions. In some species, the reliance on natural dehiscence can be circumvented by cutting the wall of the anther, although this does, of course, take a considerable amount of time. Anthers can also be cultured in liquid medium, and pollen released from the anthers can be induced to form embryos, although the efficiency of plant regeneration is often very low. Immature pollen can also be extracted from developing anthers and cultured directly, although this is a very time-consuming process.

Both methods have advantages and disadvantages. Some beneficial effects to the culture are observed when anthers are used as the explant material. There is, however, the danger that some of the embryos produced from anther culture will originate from the somatic anther tissue rather than the haploid microspore cells. If isolated pollen is used there is no danger of mixed embryo formation, but the efficiency is low and the process is time-consuming.

In microspore culture, the condition of the donor plant is of critical importance, as is the timing of isolation. Pretreatments, such as a cold treatment, are often found to increase the efficiency. These pretreatments can be applied before culture, or, in some species, after placing the anthers in culture.

Plant species can be divided into two groups, depending on whether they require the addition of plant growth regulators to the medium for pollen/anther culture; those that do also often require organic supplements, e.g. amino acids. Many of the cereals (rice, wheat, barley and maize) require medium supplemented with plant growth regulators for pollen/anther culture.

Regeneration from microspore explants can be obtained by direct embryogenesis, or via a callus stage and subsequent embryogenesis.

Haploid tissue cultures can also be initiated from the female gametophyte (the ovule). In some cases, this is a more efficient method than using pollen or anthers.

The ploidy of the plants obtained from haploid cultures may not be haploid. This can be a consequence of chromosome doubling during the culture period. Chromosome doubling (which often has to be induced by treatment with chemicals such as colchicine) may be an advantage, as in many cases haploid plants are not the desired outcome of regeneration from haploid tissues. Such plants are often referred to as 'di-haploids', because they contain two copies of the same haploid genome.

Plant regeneration

Having looked at the main types of plant culture that can be established *in vitro*, we can now look at how whole plants can be regenerated from these cultures.

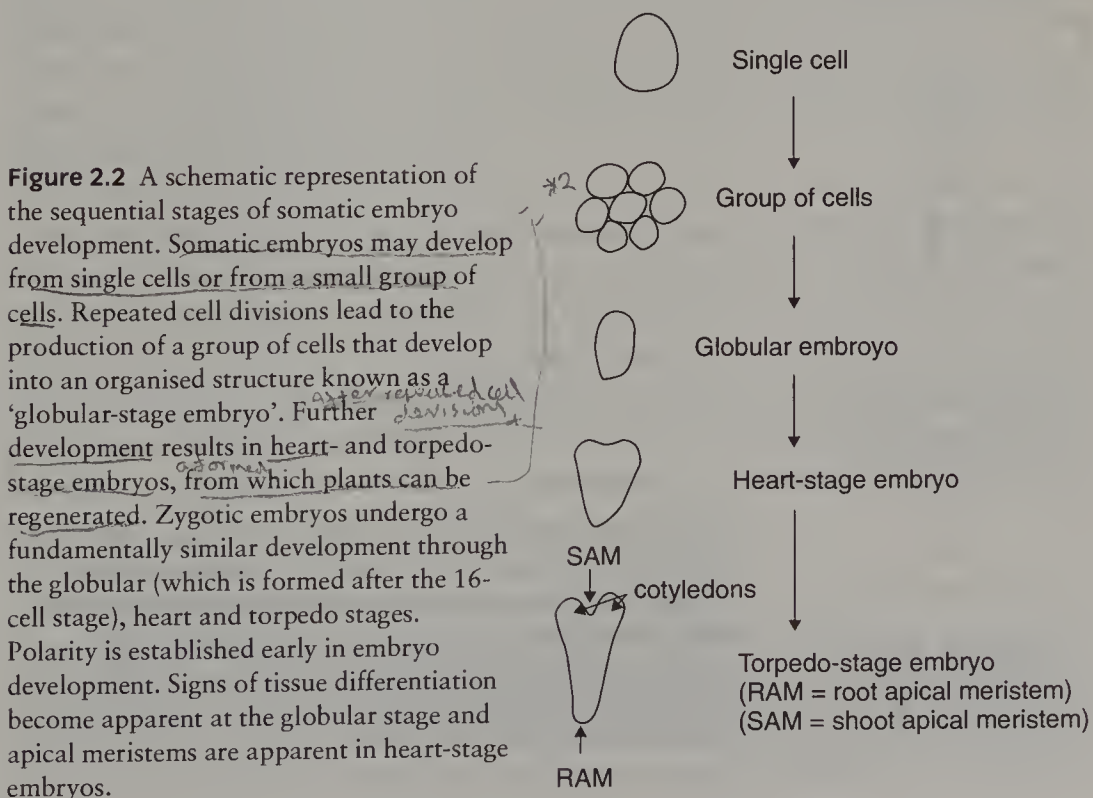
In broad terms, two methods of plant regeneration are widely used in plant transformation studies, i.e. somatic embryogenesis and organogenesis.

Somatic embryogenesis

In somatic (asexual) embryogenesis, embryo-like structures, which (can develop into whole plants) in a way analogous to zygotic embryos, are formed from somatic tissues (Figure 2.2). These somatic embryos can be produced either directly or indirectly. In direct somatic embryogenesis, the embryo is formed directly from a cell or small group of cells without the production of an intervening callus. Though common from some tissues (usually reproductive tissues such as the nucellus, styles or pollen), direct somatic embryogenesis is generally rare in comparison with indirect somatic embryogenesis. An example of direct somatic embryogenesis is given in Box 2.1.

In indirect somatic embryogenesis, callus is first produced from the explant. Embryos can then be produced from the callus tissue or from a cell suspension produced from that callus. Somatic embryogenesis from carrot is the classical example of indirect somatic embryogenesis and is explained in more detail in Box 2.1

Somatic embryogenesis usually proceeds in two distinct stages. In the initial stage (embryo initiation), a high concentration of 2,4-D is used. In the second stage (embryo production) embryos are produced in a medium with no or very low levels of 2,4-D.

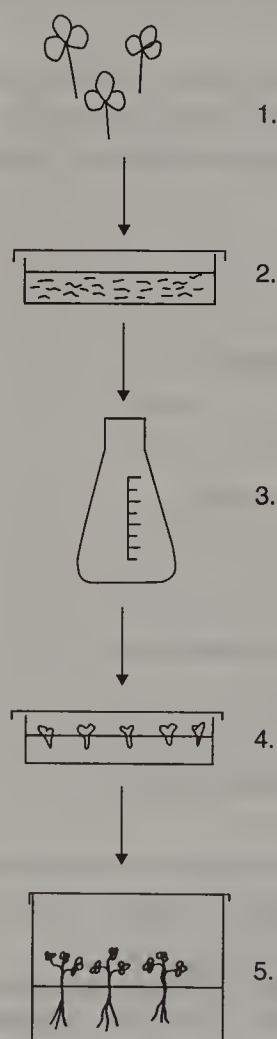


BOX 2.1 Somatic embryogenesis**Indirect somatic embryogenesis in carrot (*Daucus carota*)**

A callus can be established from explants from a wide range of carrot tissues by placing the explant on solid medium (e.g. Murashige and Skoog (MS)) containing 2,4-D (1 mg l^{-1}). This callus can be used to produce a cell suspension by placing it in agitated liquid MS medium containing 2,4-D (1 mg l^{-1}). This cell suspension can be maintained by repeated subculturing into 2,4-D-containing medium. Removal of the old 2,4-D-containing medium and replacement with fresh medium containing abscisic acid (0.025 mg l^{-1}) results in the production of embryos.

Direct somatic embryogenesis from alfalfa (*Medicago falcata*)

Young trifoliate leaves are used as the explant (see Figure). These are removed from the plant and chopped into small pieces. The pieces are washed in a plant growth regulator-free medium and placed in liquid medium (B5) supplemented with 2,4-D (4 mg l^{-1}), kinetin (0.2 mg l^{-1}), adenine (1 mg l^{-1}) and glutathione



Direct somatic embryogenesis in alfalfa. (1) Explants are removed from plants grown *in vitro*. (2) Explants are placed in liquid medium for embryo induction. (3) Embryos develop to the globular stage in liquid medium supplemented with maltose and polyethylene glycol. (4) Embryos mature on gelled medium containing abscisic acid (ABA). (5) Embryos develop into plants on solid medium.

BOX 2.1 Continued

(10 mg l⁻¹). The cultures are maintained in agitated liquid medium for about 10–15 days. Washing the explants and replacing the old medium with B5 medium supplemented with maltose and polyethylene glycol results in the development of the somatic embryos. These somatic embryos can be matured on solid medium containing abscisic acid.

Note that in both cases, although the production of somatic embryos from alfalfa necessitates the use of more complicated media, the production of embryos is fundamentally a two-step process. The initial medium, which contains 2,4-D, is replaced with a medium that does not contain 2,4-D.

In many systems it has been found that somatic embryogenesis is improved by supplying a source of reduced nitrogen, such as specific amino acids or casein hydrolysate.

CASE STUDY Cereal regeneration via somatic embryogenesis from immature or mature embryos

The principal method adopted for the tissue culture and regeneration of a wide range of cereal species is somatic embryogenesis, using cultures initiated from immature zygotic embryos. Embryogenic callus is normally initiated by placing the immature embryo on to a medium containing 2,4-D. Shoot regeneration is initiated by placing the embryogenic callus on a medium with BAP (with or without 2,4-D). These shoots can be subsequently rooted. The medium used for induction of embryogenic callus is usually a modified MS (for Triticeae) or N6 (for rice and maize). Maltose is often used as the carbon source in preference to sucrose, and additional organic supplements (such as specific amino acids, yeast extract and/or casein hydrolysate) are common.

The isolation and culture of immature embryos is, however, a labour-intensive and relatively expensive procedure. An additional problem is the small target size if the immature embryos are to be used for biolistic transformation (biolistic transformation is explained in Chapter 3). Alternatives are therefore being sought. One alternative is to use mature embryos (or seeds) as the explant to initiate embryogenic callus. This approach has been successfully applied to several cereal species such as rice and oats. The culture techniques and media used for culture establishment and regeneration from mature embryo/seed-derived cultures are fundamentally the same as those used for cultures initiated from immature embryos.

The importance of genotype

The major influence on tissue-culture response appears to be genetic, with culture requirements varying between species and cultivars. Model genotypes that responded well to culture *in vitro* were initially used in plant transformation studies. However, most of the model genotypes used were not elite, commercial cultivars. The commercial

cultivars tended to respond poorly to culture *in vitro*. One of the main aims is therefore to identify the components that make up a widely applicable, optimal culture regime.

Many factors have been investigated for their ability to improve the culture response from elite cultivars, including media components (such as alternative carbon sources, macro- and microelement concentrations and composition), media preparation method and donor plant condition and growth conditions.

Organogenesis

Somatic embryogenesis relies on plant regeneration through a process analogous to zygotic embryo germination. Organogenesis (Box 2.2) relies on the production of organs, either directly from an explant or from a callus culture. There are three methods of plant regeneration via organogenesis.

The first two methods depend on adventitious organs arising either from a callus culture or directly from an explant. Alternatively, axillary bud formation and growth can also be used to regenerate whole plants from some types of tissue culture.

Organogenesis relies on the inherent plasticity of plant tissues, and is regulated by altering the components of the medium. In particular, it is the auxin

BOX 2.2 Organogenesis in tobacco (*Nicotiana tabacum*)

Organogenesis from tobacco pith callus is the classical example of how varying plant growth regulator regimes can be used to manipulate the pattern of regeneration from plant tissue cultures.

When cultured on a medium containing both auxin and cytokinin, callus will proliferate. If the auxin to cytokinin ratio is increased, adventitious roots will form from the callus by organogenesis. If the auxin to cytokinin ratio is decreased adventitious shoots will be formed.

If the explants are cultured on medium containing only a cytokinin shoots can be produced directly.

Tobacco plants can also be easily regenerated from tobacco leaf pieces. Leaves are cut into approximately 1 cm squares with a sterile scalpel (avoiding large leaf veins and any damaged areas). The leaf pieces are then transferred (right side up) to gelled MS medium supplemented with 1 mg l^{-1} BAP (a cytokinin) and 0.1 mg l^{-1} NAA (an auxin). Over the next few weeks, callus forms on the explants, particularly around the cut surfaces. After 3 to 5 weeks shoots emerge directly from the explants or from callus derived from the explants. When these shoots are about 1 cm long they can be cut at the base and placed on to solid MS medium without any plant growth regulators. The shoots will form roots and form plantlets that will grow in this medium and can subsequently be transferred to soil.

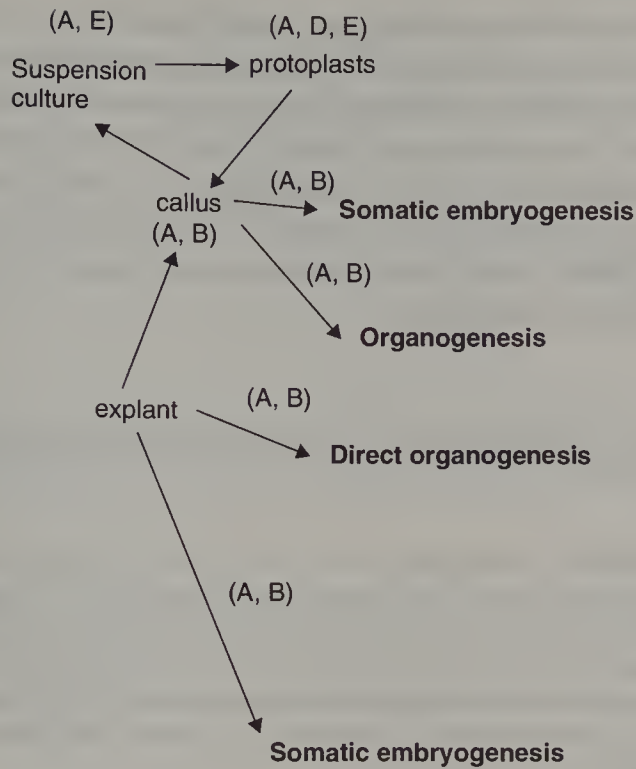


Figure 2.3 A simplified scheme for the integration of plant tissue culture into plant transformation protocols. An explant can be a variety of tissues, depending on the particular plant species being cultured. The explant can be used to initiate a variety of culture types, depending on the explant used. Regeneration by either organogenesis or somatic embryogenesis results in the production of whole plants. Different culture types and regeneration methods are amenable to different transformation protocols. The transformation protocols (see Chapter 3) highlighted in this figure are: (A) *Agrobacterium*-mediated; (B) biolistic transformation; (D) direct DNA uptake and (E) electroporation. Different combinations of culture type and transformation protocol are used depending on the plant species and cultivar being used. In some species a variety of culture types and regeneration methods can be used, which enables a wide variety of transformation protocols to be utilised. In other species there is effectively no choice over culture type and/or regeneration method, which can limit the transformation protocols that are applicable. (Redrawn with permission from Walden R. and Wingender R. (1995).)

to cytokinin ratio of the medium that determines which developmental pathway the regenerating tissue will take.

It is usual to induce shoot formation by increasing the cytokinin to auxin ratio of the culture medium. These shoots can then be rooted relatively simply.

Integration of plant tissue culture into plant transformation protocols

Various methods of plant regeneration are available to the plant biotechnologist. Some plant species may be amenable to regeneration by a variety of

methods, but some may only be regenerated by one method. In Chapter 3 the various methods that can be used to transform plants will be considered, but it is worthwhile briefly considering the interaction of plant regeneration methodology and transformation methodology here (Figure 2.3). Not all plant tissue is suited to every plant transformation method, and not all plant species can be regenerated by every method. There is therefore a need to find both a suitable plant tissue culture/regeneration regime and a compatible plant transformation methodology.

Summary

Tissue culture and plant regeneration are an integral part of most plant transformation strategies, and can often prove to be the most challenging aspect of a plant transformation protocol. Key to success in integrating plant tissue culture into plant transformation strategies is the realisation that a quick (to avoid too many deleterious effects from somaclonal variation) and efficient regeneration system must be developed. However, this system must also allow high transformation efficiencies from whichever transformation technique is adopted.

Not all regeneration protocols are compatible with all transformation techniques. Some crops may be amenable to a variety of regeneration and transformation strategies, others may currently only be amenable to one particular protocol. Advances are being made all the time, so it is impossible to say that a particular crop will never be regenerated by a particular protocol. However, some protocols, at least at the moment, are clearly more efficient than others. Regeneration from immature embryo-derived somatic embryos is, for example, the favoured method for regenerating monocot species.

This chapter has only scratched the surface of the problems and potentials of plant tissue culture. General rules are sometimes difficult to make because of the variability of response to particular protocols of different plant species or even cultivars. However, as will be seen in Chapter 3 and subsequent chapters, plant tissue culture has been successfully integrated into plant transformation strategies and the list of plant species that can be routinely transformed continues to grow.

Further reading

- Barcelo, P., Rasco-Gaunt, S., Thorpe, C. and Lazzeri, P. (2001). Transformation and gene expression. In *Advances in botanical research*, volume 34: *Biotechnology of cereals* (ed. P. R. Shewry, P. A. Lazzeri and K. J. Edwards), pp. 59–126. Academic Press, London.
- Dahleen, L. S. (1999). Donor-plant environment effects on regeneration from barley embryo-derived callus. *Crop Science*, 39, 682–5.
- Dahleen, L. S. and Bregitzer, P. (2002). An improved media system for high regeneration rates from barley immature embryo-derived callus cultures of commercial cultivars. *Crop Science*, 42, 934–8.
- Dodds, J. H. and Roberts, L. W. (1995). *Experiments in plant tissue culture*. Cambridge University Press, Cambridge.

- Fowler, M. R. (2000). Plant cell culture, laboratory techniques. In *Encyclopedia of cell technology* (ed. R. E. Spier), pp. 994–1004. Wiley, New York.
- Gamborg, O. L. (2002). Plant tissue culture. Biotechnology. Milestones. *In vitro Cellular and Developmental Biology—Plant*, 38, 84–92.
- Ramage, C. M. and Williams, R. R. (2002). Mineral nutrition and plant morphogenesis. *In vitro Cellular and Developmental Biology—Plant*, 38, 116–24.
- Sugiyama, M. (1999). Organogenesis *in vitro*. *Current Opinion in Plant Biology*, 2, 61–4.
- Torbert, K. A., Rines, H. W. and Somers, D. A. (1998). Transformation of oat using mature embryo-derived tissue cultures. *Crop Science*, 38, 226–31.
- Torbert, K. A., Rines, H. W., Kaeppler, H. F., Menon, G. K. and Somers, D. A. (1998). Genetically engineering elite oat cultivars. *Crop Science*, 38, 1685–7.
- Walden, R. and Wingender, R. (1995). Gene-transfer and plant-regeneration techniques. *Trends in Biotechnology*, 13, 324–31.

3

Techniques for plant transformation

Introduction

Plant transformation has become widely adopted as a method to both understand how plants work and to improve crop plant characteristics. Plant transformation depends on the stable introduction of transgene(s) into the genome of the plant. Various methods have been developed to achieve this and many plant species have been successfully transformed. Some transformation methods are based on utilising *Agrobacterium*, a pathogen of dicotyledonous (broad-leafed) plants that transfer genes into the plant genome. Problems with using *Agrobacterium* to transform monocotyledonous plants (grasses) spurred on the development of other methods, the so-called 'direct gene transfer' methods. The major direct gene transfer method, particle bombardment (or biolistics), is the method of choice in many laboratories for the transformation of monocotyledonous plants, despite *Agrobacterium*-based protocols having subsequently been developed for the transformation of monocotyledonous plants.

In this chapter we will look at both the biology of *Agrobacterium tumefaciens* and how it is used in plant transformation studies. Direct gene transfer methods for plant transformation, and in particular biolistics, will also be considered and examples of their use for the transformation of monocotyledonous and dicotyledonous crops will be given.

Agrobacterium-mediated gene transfer

The biology of *Agrobacterium*

Agrobacterium tumefaciens is a soil-borne, Gram-negative bacterium, much beloved of plant biologists and biotechnologists alike. The explosion of interest in *A. tumefaciens* has been driven by several observations that have both biotechnological and purely scientific ramifications. *A. tumefaciens* is the causative agent of 'crown gall' disease, an economically important disease of many plants (particularly grapes, walnuts, apples and roses). The ability to cause

crown galls (tumorous tissue growths) depends on the ability of *Agrobacterium* spp. to transfer bacterial genes into the plant genome. This startling feature is, to date, a unique example of inter-kingdom gene transfer, which has obvious scientific interest. More important to the biotechnologist is the development of plant transformation methods that make use of this unique capability.

Crown-gall disease

In order to fully appreciate and understand how *A. tumefaciens* became such a widely used tool in plant biotechnology it is worth looking at the biology of crown-gall disease. *A. tumefaciens*, the causative agent of crown-gall disease, is a Gram-negative, rod-shaped, motile bacterium found in the rhizosphere (the region around the roots of plants) where it normally survives on nutrients released from plant roots. However, if a plant is wounded or damaged, *A. tumefaciens* can infect the plant at the wound site and cause disease symptoms. *A. tumefaciens* is attracted to the wound site via chemotaxis, in response to chemicals (sugars and phenolic molecules) released from the damaged plant cells.

Crown-gall formation depends on the presence of a plasmid in *A. tumefaciens* known as the 'Ti (tumour-inducing) plasmid'. Part of this plasmid (the T-DNA region) is actually transferred from the bacterium and into the plant cell, where it becomes integrated into the genome of the host plant. The T-DNA carries genes that encode proteins involved in both hormone (auxin and cytokinin) biosynthesis and the biosynthesis of novel plant metabolites called 'opines' and 'agropines'. The production of auxin and cytokinin causes the plant cells to proliferate and so form the gall. These proliferating cells also produce opines (which are amino acid derivatives) and agropines (sugar derivatives) which are used by *A. tumefaciens* as its sole carbon and energy source (Figure 3.1 shows the structure of the common opines, octopine and nopaline). Different strains of *A. tumefaciens* contain different Ti plasmids that code for the production of different opines. Opines and agropines are not normally part of plant metabolism and are very stable chemicals, which provide a carbon and energy source that only *A. tumefaciens* can use (genes on the Ti plasmid that are not transferred to the plant encode proteins involved in opine uptake and catabolism). *A. tumefaciens* has therefore developed the ability to genetically transform plant cells in order to usurp the plant's biosynthetic machinery and produce nutrients that only it can utilise.

The Ti plasmid

As the Ti plasmid has a central role in crown-gall formation, and it is a portion of the Ti plasmid (the T-DNA region) that is actually integrated into the genome of the host plant, we shall look at the Ti plasmid in more detail.

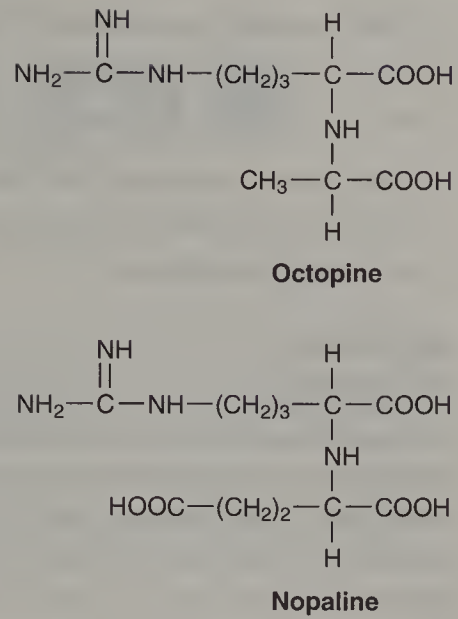


Figure 3.1 The structures of the opines octopine and nopaline are shown. Both octopine and nopaline are amino acid derivatives. Octopine is derived from the amino acid arginine and pyruvate. Nopaline is derived from arginine and α -ketoglutarate.

As we have seen, the ability of *A. tumefaciens* to cause crown-gall disease was found to depend on the presence in the bacteria of a large (~200 kb) plasmid termed the Ti plasmid. Analysis of the nuclear DNA from plant tumours showed that a portion of the Ti plasmid was integrated into the genome of the host plant. This portion was termed 'transfer DNA' (T-DNA) and was found to be responsible for the tumorous phenotype. Analysis showed that one or more copies of the T-DNA could be integrated into the genome, but, in general, the T-DNA insertions in the plant genome were bordered by small (24 bp), nearly perfect, direct repeats, which also border the T-DNA in the Ti plasmid (Figure 3.2).

Nopaline strains

TG(G/A)CAGGATATAT(-/T)G(T/G)(G/C)G(T/G)GTAAAC

Octopine strains

(C/T)GGCAGGATATA(T/A)C(A/C)(A/G)TTGTAA(A/T)T

TAAGTCGCTGTGTATGTTTGTGTTG (Enhancer or 'overdrive' sequence)

Figure 3.2 The consensus sequence of the border sequences from nopaline- and octopine-strain Ti plasmids. The nucleotide sequence of the enhancer sequence is also shown.

Ti-plasmid features

Ti plasmids from different strains of *A. tumefaciens* generally have several features in common:

1. They contain one (or more) T-DNA regions.
2. They contain a *vir* region.
3. They contain an origin of replication.
4. They contain a region enabling conjugative transfer.
5. They contain genes for the catabolism of opines (a class of amino acid/sugar conjugates).

The T-DNA

The T-DNA region of any Ti plasmid is defined by the presence of the right- and left-border sequences (see Figure 3.2). These border sequences are 24-bp imperfect repeats. Any DNA between the borders will be transferred into the genome of the host plant. Although not part of the T-DNA, octopine-strain Ti plasmids contain an 'overdrive' or enhancer sequence (see Figure 3.2) associated with the right-border sequence, which is required for optimal T-DNA transfer.

Nopaline strains of Ti plasmid have one T-DNA of approximately 20 kb, whereas octopine strains have two T-DNA regions, termed 'T_L' and 'T_R', that are approximately 14 and 7 kb, respectively. However, only the T_L region is oncogenic (the T_R region carries genes for opine biosynthesis).

The T-DNA of nopaline Ti plasmids contains 13 open reading frames (ORFs), whereas the T_L T-DNA of octopine plasmids contains 8 ORFs. These ORFs have the features of eukaryotic, rather than prokaryotic, genes. The nopaline T-DNA and the T_L octopine T-DNA show extensive similarity in an area known as the 'core' region. This region contains the genes that code for proteins involved in hormone biosynthesis (the oncogenes), opine synthesis and for determining tumour size. The organisation of the genes in the T_L T-DNA is shown in Figure 3.3.

The oncogenes

Two genes *auxA* (or *tms1* or *iaaM*) and *auxB* (or *tms2* or *iaaH*) encode proteins involved in the production of the auxin indole acetic acid (IAA). *auxA* encodes tryptophan monooxygenase and *auxB* encodes indole acetamidehydrolase.

Another gene (*cyt* (or *tmr* or *ipt*)) encodes an isopentenyl transferase that catalyses the most important step in cytokinin production.

These genes are the prime determinants of tumour phenotype and are therefore often referred to as 'oncogenes'. The T_R region of octopine T-DNA contains no oncogenes, explaining its lack of oncogenicity.

Other genes present on the T-DNA

Genes for the production of opines (either octopine or nopaline) are present on the T-DNA (the octopine T_R T-DNA possesses genes encoding proteins involved in the biosynthesis of other opines such as mannopine, agropine and fructopine). The *tml* gene, which is involved in determining tumour size in some species, is also found in the T-DNA (see Figure 3.3).

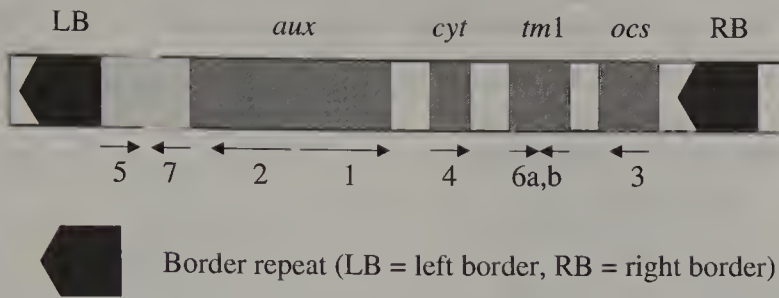


Figure 3.3 The genetic organisation of the T_L T-DNA of an octopine-type Ti plasmid. Only the T_L region is shown as this has homology with the T-DNA of nopaline-type Ti plasmids. Eight open reading frames (ORFs) are indicated (1–7), although ORFs 5 and 7 are not discussed in this text. Regions of import are shaded light grey, and include the *aux* genes (which encode enzymes involved in auxin biosynthesis), *cyt* which encodes isopentyl transferase (an enzyme involved in cytokinin production), *tm1* which is involved in regulating tumour size and *ocs* (octopine synthase) which encodes opine synthesis. (Redrawn with permission from Hughes, M. A. (1996).)

The *vir* region

The genes responsible for the transfer of the T-DNA region into the host plant are also situated on the Ti plasmid, in an ~ 40-kb region outside the T-DNA known as the *vir* (virulence) region.

There are at least nine *vir*-gene operons (Table 3.1 summarises *vir*-gene function). The functions of the *vir* genes present in these operons have, in many cases, been determined. In the following section we will look in more detail at *vir*-gene function and the roles they play in T-DNA transfer from *A. tumefaciens* to plant cells.

The process of T-DNA transfer and integration

T-DNA transfer and integration into the plant genome can be divided into the following steps (see Figure 3.4 for a simplified representation of this process).

1. Signal recognition by *Agrobacterium*

The *Agrobacterium* perceives signals, such as phenolics and sugars, which are released from wounded plant cells. Normally these substances are probably part of the plant's defence mechanism, being involved in phytoalexin and lignin synthesis. The substances released from wounded cells effectively signal the presence of plant cells that are competent for transformation.

2. Attachment to plant cells

Attachment of *Agrobacterium* to plant cells is a two-step process, involving an initial attachment via a polysaccharide (the product of the *attR* locus).

Table 3.1 *Agrobacterium* virulence protein function

Virulence protein	Function in <i>Agrobacterium</i> spp.	Function in plant	Plant proteins that interact
VirA	Phenolic sensor kinase Part of two-component system with VirG; phosphorylates and activates VirG		
VirG	Transcription factor Responsible for induction of <i>vir</i> gene expression		
VirB1–B11	Components of membrane structure (transfer apparatus) for transfer of T-DNA		
VirC1	‘Overdrive’ binding protein Enhances efficiency of T-DNA transfer		
VirD1	Required for T-DNA processing Modulates VirD2 activity		
VirD2	Nicks the T-DNA and directs T-DNA through the VirB/VirD4 transfer apparatus	Nuclear targeting of T-DNA Protection of 5' end of T-DNA from nucleases Possibly involved in integration	Importin- α^a Cyclophilins ^b pp2C ^c Histone H2A
VirD4	Component of transfer apparatus		
VirE1	VirE2 chaperone Required for VirE2 export from <i>Agrobacterium</i> spp.		
VirE2		Single-stranded DNA-binding protein Prevents T-DNA degradation by nucleases Involved in nuclear targeting and passage through NPC	VIP1 ^d VIP2 ^e
VirF		Cell-cycle regulation Elongation of S-phase	Plant Skp1 ^f homologue
VirJ	T-DNA export		

^a Importin (enables transport through the plant nuclear-pore complex (NPC)).

^b Cyclophilins may cause conformational changes in VirD2 or may be involved in targeting to chromatin and integration of the T-DNA.

^c Serine/threonine protein phosphatase type 2C (pp2C) negatively affects nuclear import of the T-DNA.

^d VIP1 facilitates nuclear import and may be involved in directing the T-DNA to chromatin.

^e VIP2 may target the T-DNA to transcriptionally active chromatin.

^f Skp1 (an F-box protein) homologue involved in protein degradation and cell-cycle control. May prolong the S-phase and improve the efficiency of integration.

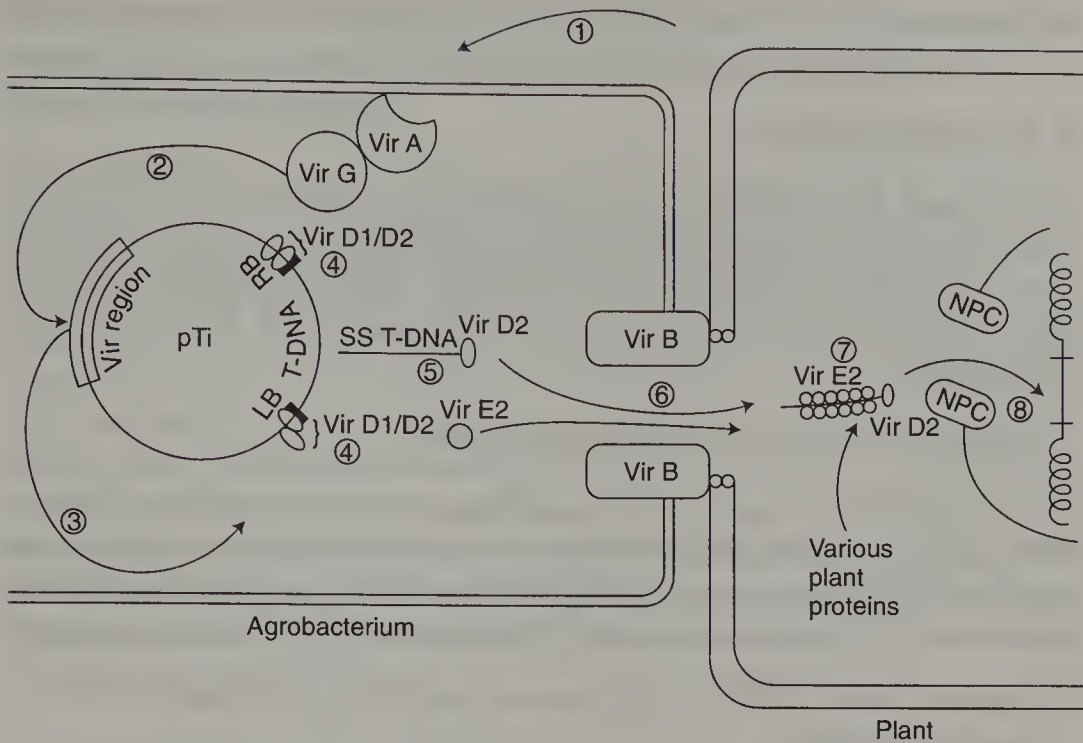


Figure 3.4 A simplified representation of the T-DNA transfer and integration process. Wounded plant cells release phenolic substances and sugars (1) that are sensed by VirA, which activates VirG by phosphorylation. VirG induces the expression of all the genes in the *vir* region of the Ti plasmid (2). Gene products of the *vir* genes (3) are involved in a variety of processes. VirD1 and VirD2 are involved in single-stranded T-DNA production, protection and export (4) and (5). VirB products form the transfer apparatus. The single-stranded T-DNA (associated with VirD2) and VirE2 are exported through the transfer apparatus (VirF may also be exported) (6). In the plant cell the T-DNA becomes coated with VirE2 (7). Various plant proteins interact with either the VirD2 or VirE2, which are attached to the T-DNA and influence transport and integration. The T-DNA/VirD2/VirE2/plant protein complex enters the nucleus through the nuclear pore complex. Integration into the plant chromosome (8) occurs *via* illegitimate recombination. (LB, left border; RB, right border; pTi, Ti plasmid; ss, single stranded; NPC, nuclear pore complex).

Subsequently, a mesh of cellulose fibres is produced by the bacterium. Several chromosomal virulence genes (*chv* genes) are involved in the attachment of the bacterial cells to the plant cells. The *chvA* and *chvB* genes are involved in the production and secretion of cyclic β -1,2-glycans and the *pscA* gene (*psc*, polysaccharide) is involved in secreting succinoglycan. *Agrobacterium* strains carrying mutated *chvA* or *chvB* genes are avirulent or extremely attenuated.

3. *vir* Gene induction

VirA (a membrane-linked sensor kinase) senses phenolics (such as acetosyringone) and autophosphorylates, subsequently phosphorylating and thereby activating VirG. VirG then induces expression of all the *vir* genes (including *virA* and *virG*). Many sugars, but in particular glucose, galactose and xylose, enhance *vir* gene induction. This enhancement requires another chromosomal

vir gene termed '*chvE*', which encodes a glucose/galactose transporter that interacts with VirA.

4. T-strand production

The left and right borders are recognised by a VirD1/VirD2 complex and VirD2 produces single-stranded nicks in the DNA. After nicking, VirD2 becomes covalently attached to the 5' end of the displaced single-stranded T-DNA strand. Repair synthesis replaces the displaced strand. VirC1 may assist in the process.

5. Transfer of T-DNA out of the bacterial cell

The T-DNA/VirD2 complex is exported from the bacterial cell by a 'T-pilus' (effectively a membrane-channel secretory system) composed of proteins encoded by the *virB* operon and VirD4. VirE2 and (VirF) are also exported from the bacterial cell (although it is possible that some other system is responsible).

6. Transfer of the T-DNA and Vir proteins into the plant cell and nuclear localisation

The T-DNA/VirD2 complex and other Vir proteins cross the plant plasma membrane, possibly through channels formed from VirE2. Once inside the plant cytoplasm the T-DNA strand becomes covered with VirE2 proteins, which have been postulated to protect the T-DNA from nucleases, facilitate nuclear localisation and confer the correct conformation to the T-DNA/VirD2 complex for passage through the nuclear-pore complex (NPC).

Both VirD2 and VirE2 have been shown to interact with a variety of plant proteins that are important in ensuring nuclear localisation of the T-DNA complex. VirD2 contains a nuclear localisation signal (NLS) that facilitates its interaction with a plant protein. This plant protein belongs to a group of proteins (called 'importins') involved in the recognition and transport through the NPC of proteins that possess a NLS.

VirE2 possesses two NLSs, but its nuclear localisation is mediated by another plant protein, termed 'VIP1', which functions to facilitate VirE2 NLS recognition by importins. This contribution of VirE2 to the nuclear localisation of T-DNA complexes may be particularly important for large T-DNAs.

The T-DNA and associated proteins pass through the nuclear pore, with the bound VirE2 proteins also giving the correct conformation to the T-DNA strand. VIP1 and another VirE2 interacting protein, VIP2, are then thought to direct the T-DNA strand to chromatin and possibly promote integration. It is possible that other plant proteins called cyclophilins—which bind VirD2—might also aid integration.

The T-DNA strand is integrated into the host plant genome by a process referred to as 'illegitimate recombination'. This process, unlike homologous recombination, does not depend on extensive regions of sequence similarity.

Practical applications of *Agrobacterium*-mediated plant transformation

We have looked at the basic biology of *A. tumefaciens* and how the T-DNA is transferred into the plant genome in some detail. We can now look at how *A. tumefaciens* is incorporated into plant transformation experiments. It is important to remember that the successful production of transgenic plants is reliant on combining a suitable transformation protocol with a robust regeneration protocol (see Chapter 2). *Agrobacterium*-mediated transformation may not prove suitable for all types of explant, and may not be the most efficient transformation method for some plant species. Nevertheless, *Agrobacterium*-mediated transformation is a widely used transformation method, partly for historical reasons and partly because of some advantages over other methods. *Agrobacterium*-mediated transformation methods are thought to induce less rearrangement of the transgene and result in a lower transgene copy number than direct DNA delivery methods.

CASE STUDY 1 *Agrobacterium*-mediated transformation of tobacco

Tobacco is a relatively easy plant to transform with *Agrobacterium* and provides a good introduction to the use of *Agrobacterium* in plant transformation. In this case study we will look at the processes used for *Agrobacterium*-mediated plant transformation.

Several factors have to be considered in the design and implementation of any plant transformation study:

1. **The plant tissue to be transformed.** The purpose of most plant transformation experiments in plant biotechnology is to produce whole, transgenic plants. The explant used in transformation experiments must therefore be capable of producing whole plants by regeneration, and should contain a high proportion of cells that are competent for transformation (explants and the regeneration of whole plants from them are discussed in detail in Chapter 2).

2. **The vector used to deliver the transgene into the genome of the plant.** The construction of vectors for use in *Agrobacterium*-mediated transformation is discussed in detail in Chapter 4. However, there are some key points, relevant to their use for plant transformation, which we can consider here (a simple vector for *Agrobacterium*-mediated plant transformation is shown in Figure 3.5). Vectors used in *Agrobacterium*-mediated transformation are derivatives of, or are based on, the naturally occurring Ti plasmid. They are, however, extensively modified so that most of the features of a natural Ti plasmid are removed, only the left- and right-border sequences being used to ensure transfer of the T-DNA region between them. The vector also contains a selectable marker on the T-DNA so that transformed plants can be identified, as well as a separate selectable marker (outside the T-DNA) to enable identification of the transformed bacteria. Virulence genes required for transfer of the T-DNA are often located on a separate plasmid in the bacterium.

3. **The strain of *Agrobacterium* used.** Several widely used strains of *Agrobacterium*

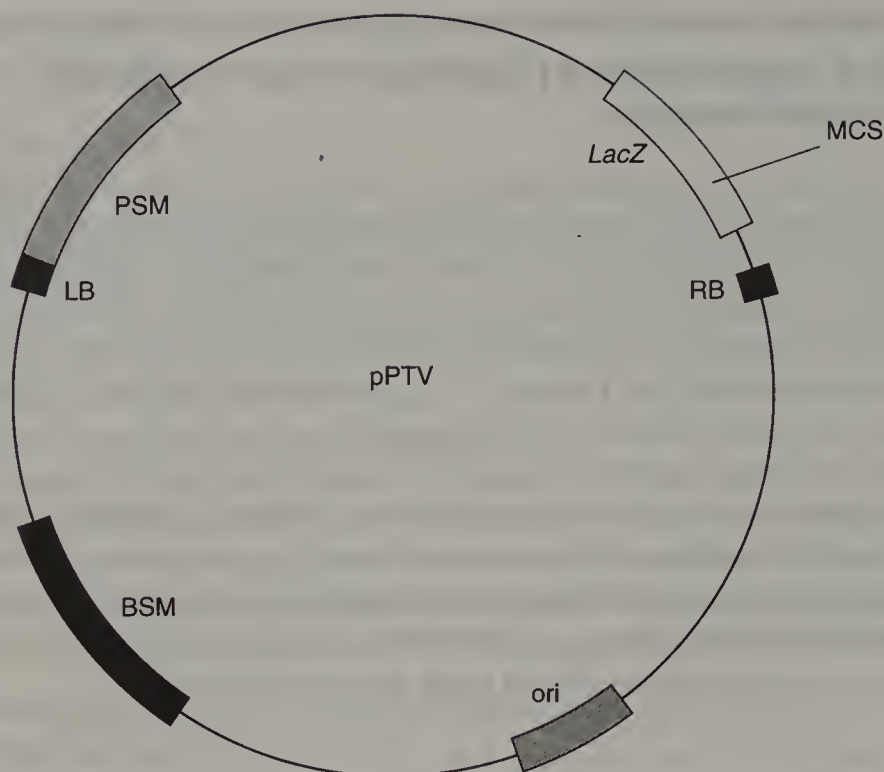


Figure 3.5 A simplified representation of a plasmid plant transformation vector (pPTV) showing the essential features of a binary-type plant transformation vector. The DNA that is transferred to the plant genome (the T-DNA) is situated between, and defined by, the left (LB) and right (RB) borders. It contains a multiple cloning site (MCS) in *LacZ* α to facilitate cloning of the transgene into the vector. The T-DNA also contains a selectable marker gene to enable selection of transformed plant cells (PSM). Outside the T-DNA is a bacterial selectable marker gene (BSM) to allow selection of transformed bacteria. There is also an origin of replication (*ori*) to allow replication of the plasmid. Binary vectors like pPTV are described fully in Chapter 4.

are available for plant transformation experiments. For the transformation of crops that, like tobacco, are amenable to *Agrobacterium*-mediated transformation, the choice of strain is not critical to the success of the experiment. However, for more recalcitrant plant species, the choice of strain is a major factor contributing to the success or failure of the experiment.

The basic protocol used for any *Agrobacterium*-mediated transformation experiment can be illustrated with the following example of tobacco transformation.

1. Suitable plant tissue, to be used as a source of explants, is removed from the donor plant and sterilised (if it is not from a sterile plant). For tobacco transformation, leaves are ideal.
2. The leaf tissue is cut into small pieces (using a scalpel or cork-borer) and placed into a culture of *Agrobacterium* (which contains the vector) for about 30 minutes, a process known as co-cultivation. During this incubation, the bacteria attach to the plant cells. The explants are subsequently removed from the bacterial culture, excess bacterial culture blotted off, and then placed on to solid MS medium that contains no selective agent.
3. The incubation of the explants with *Agrobacterium* is allowed to continue for 2 days to allow transfer of the T-DNA to the plant cells.

4. The explants are removed from the medium and washed in an antibiotic solution (such as cefotaxime) that kills *Agrobacterium* cells.
 5. The explants are transferred to fresh solid medium. This medium is supplemented with a selective agent (often kanamycin, but this will depend on which selectable marker gene is present in the T-DNA of the vector) to prevent the growth of non-transformed plant cells. It also contains cefotaxime (to prevent the growth of any *Agrobacterium* that were not killed by the initial treatment with cefotaxime). An auxin (naphthylacetic acid (NAA), 0.1 mg l^{-1}) and a cytokinin (benzylaminopurine (BAP), 1 mg l^{-1}) are also included to encourage regeneration by organogenesis (see Chapter 2). The relatively high cytokinin to auxin ratio promotes shoot formation from the explants. These shoots can be rooted by placing them on solid medium containing a high auxin to cytokinin ratio.
-

Although this is a specific example, most *Agrobacterium*-mediated transformation protocols (although see *in planta* methods described in a later section of this chapter) follow a similar pattern, which is summarised below:

1. Identify a suitable explant.
2. Co-cultivate with the *Agrobacterium*.
3. Kill the *Agrobacterium* with a suitable antibiotic (which does not harm the plant tissue).
4. Select for transformed plant cells.
5. Regenerate whole plants.

Experimental details vary depending on the plant species being transformed and the strain of *Agrobacterium* being used. There is a multitude of protocols available for the transformation of a wide range of crops, even ones such as cereals that were previously considered to be incapable of being transformed by *Agrobacterium*-mediated methods. These protocols are constantly being improved, which makes the identification of specific protocols redundant. However, some generalisations can be made.

Dicotyledonous plants, the natural target for *Agrobacterium* transformation, are, in general, easily transformed using standard vectors (see Chapter 4) and standard strains of *Agrobacterium*, such as LBA4404. Some crops, such as cereals (which are not naturally infected by *Agrobacterium*) are more difficult to transform and may require the use of modified vectors and/or so-called 'supervirulent' strains of *Agrobacterium* (such as EHA101 or EHA105). These supervirulence systems rely on extra copies of some of the virulence genes either being present on the vector itself or on a separate plasmid in the *Agrobacterium*. Different selectable markers (Chapter 4) to identify transformed plant cells may also be required for some plant species, particularly cereals. Antibiotics such as kanamycin are commonly used to select for transformed plant cells, but alternatives such as herbicides or more potent antibiotics are often required for cereal transformation.

Transformation *in planta*

Transformation protocols requiring extensive periods of tissue culture have several drawbacks, including time (typically many months), space and, importantly, the tendency to induce somaclonal variation (i.e. the variation in plants derived from tissue culture). There is therefore considerable interest in developing transformation methods that obviate the need for tissue culture steps. Recently, such methods have been demonstrated to work in a limited number of plant species, opening up the possibility of widespread application to crop improvement. Methods for transformation *in planta* utilise either *Agrobacterium*-mediated transformation or direct gene transfer methods (which are discussed later in this chapter). One of the most promising methods, that was developed using a member of the genus *Arabidopsis* as the target plant, is the floral dip. This is an extremely simple method in which plants with young flowers are dipped (with or without a vacuum being applied) into a culture of *Agrobacterium* which also contains a surfactant. The plant is subsequently allowed to set seed, whereupon a small proportion of the seeds produced are transgenic. Although the efficiency of this technique is very low (at present, at least) the vast number of seeds produced results in an acceptable overall transformation efficiency. This, or similar techniques have been applied successfully to other plant species, including alfalfa and some brassicas.

The first part of this chapter has dealt with *Agrobacterium*-mediated transformation. However, another species of the genus *Agrobacterium*, *A. rhizogenes*, is also capable of transferring genes to plants, and has been developed into a plant transformation system (Box 3.1) that is used in some specialised circumstances.

Direct gene transfer methods

The term ‘direct gene transfer’ (or direct transfer) is used to discriminate between methods of plant transformation that rely on the use of *Agrobacterium* (indirect methods) and those that do not (direct methods). Direct gene transfer methods all rely on the delivery of large amounts of ‘naked’ DNA whilst the plant cell is transiently permeabilised.

Several different methods or strategies for direct gene transfer have been developed over the years. Some of these, particularly ‘biolistics’ or gene bombardment, have become widely adopted by plant biotechnologists.

Direct gene transfer methods, which have found particularly widespread use in the transformation of cereal crops (that initially proved difficult to transform with *Agrobacterium*), have some advantages and disadvantages (Table 3.2) when compared with *Agrobacterium*-mediated transformation.

BOX 3.1 *Agrobacterium rhizogenes*

Although *Agrobacterium rhizogenes* also infects plants, it differs from *A. tumefaciens* in that the resulting pathology is not crown galls but a phenomenon known as 'hairy roots'. At the site of infection there is a proliferation of roots. Plasmids in *Agrobacterium rhizogenes* (Ri plasmids) strains have been characterised, and it has been shown that there are of a number of different types, which can be classified based upon opine usage. Hormone biosynthetic genes are also present on some Ri plasmids. Homologues of the Ti-plasmid auxin-biosynthetic genes *tms1* and *tms2* have been identified on the agropine-strain Ri plasmids. However, despite the well-established link between auxin and rooting, the genes are unnecessary for virulence. Rather, a series of other open reading frames have been identified within the T-DNA of the plasmids, including the *rolB* and *rolC* genes that are involved in the metabolism of plant growth regulators and lead to the plant being sensitised to endogenous auxin. It is this increased sensitivity that leads to the root formation.

As with *A. tumefaciens*, vectors have been constructed that can be used in *A. rhizogenes* as binary systems (binary vectors are described in Chapter 4).

Hairy roots are important in some areas of plant biotechnology as they can be cultured *in vitro*. For many years they have been used as a source of secondary metabolites, but more recently they have been used as a system for the production of pharmaceutical proteins.

A. rhizogenes transformation was, at one stage, considered an alternative strategy to *A. tumefaciens* for gene transfer as it led to the production of defined tissues (hairy roots) that could be regenerated into whole plants. This strategy seems to have been discarded, however, as more efficient *A. tumefaciens* systems have been developed.

One of the major disadvantages with direct gene transfer methods is that they tend to lead to a higher frequency of transgene rearrangement (see Box 3.2) and a higher transgene copy number. This can lead to high frequencies of gene silencing (see Box 3.4).

The main types of direct gene transfer will be considered in some detail below. Other, less-reproducible methods, which will not be considered further, such as laser-mediated uptake of DNA, microinjection, ultrasound and *in planta* exogenous application, have mainly been used for the analysis of transient gene expression, although stable transformation has been reported for some of these techniques on rare occasions.

Particle bombardment

Particle bombardment ('biolistics') is the most important and most effective direct gene transfer method in regular use. In this technique, tungsten or gold particles are coated with the DNA that is to be used to transform the plant

Table 3.2 Direct gene transfer methods

Direct gene transfer method	Comments
Particle bombardment	Very successful method. Risk of gene rearrangements and high copy number. Useful for transient expression assays
Electroporation	Transgenic plants obtained from a range of cereal crops. Low efficiency. Requires careful optimisation
DNA uptake into protoplasts	Used for all major cereal crops. Requires optimisation with a regenerable cell suspension that may not be available
Silicon carbide fibres	Requires regenerable cell suspensions. Transgenic plants obtained from a number of species

tissue. The particles are propelled at high speed into the target plant material, where the DNA is released within the cell and can integrate into the genome. The delivery of DNA using this technology has allowed transient gene expression (which does not depend on integration of the transgene into the plant genome) to be widely studied, but integration of the transgene occurs only infrequently. In order to generate transgenic plants, the plant material, the tissue culture regime and the transformation conditions have to be optimised quite carefully.

Practical bombardment systems were first developed in 1987 and used an explosive charge to propel DNA-coated tungsten particles. This technology was the key to cereal transformation. All the major cereals were able to be transformed, and the first commercial GM crops, such as maize containing the *Bt*-toxin gene (Chapter 6), were produced by this method. Developments to the technology led to the production of a number of systems, such as an electrostatic discharge device, and others based on gas flow. Of the latter type, a commercially produced, helium-driven, particle-bombardment apparatus (PDS-1000/He) has become the most widely used.

Attempts to optimise the system have focused on three aspects of the process: particle type and preparation; particle acceleration; and choice of target material. A balance has to be reached between the number and size of particles fired into the target cells, the damage they do and the amount of DNA they deliver. Too little DNA may lead to low transformation frequencies, but too much DNA may lead to a high copy number and rearrangements of the transgene constructs (Box 3.2). Recently, new methods of preparing the particles for biolistic transformation, which involve the use of aminosiloxanes (instead of calcium chloride and polyamines) to coat the particles with DNA, have led to higher transformation efficiencies at lower DNA concentrations.

BOX 3.2 Transgene integration and gene rearrangement during particle bombardment

The major complication found with particle-bombardment transformation methods is that the vector DNA is often rearranged and transgene copy number can be very high. Recent work on the mechanism of direct gene transfer to rice by biolistics suggests that the transfer is a two-stage process. During a preintegration phase, vector molecules (which can be either intact or partial) are spliced together. This gives rise to fragments carrying multiple gene copies, which subsequently integrate into the host plant genome. This integration event can act as a hotspot for further transgene integration at, or very close to, the initial point of integration. The implications of these observations are that particle bombardment generally results in the DNA being integrated in high copy number at a single locus (or a low number of loci). Despite this high copy number, the single locus may have benefits for subsequent breeding programmes.

Plant tissues used for bombardment are generally of two types: primary explants that are bombarded and then induced to become embryogenic; or proliferating embryogenic cultures that are bombarded and then allowed to proliferate further and subsequently to regenerate. To protect the plant tissue from the damage sustained during the bombardment procedure, treatments to induce limited plasmolysis or culture on high-osmoticum media have been used.

CASE STUDY 2 Biolistic transformation of rice

In this case study we will look at how the PDS-1000/He particle-bombardment system can be used to transform rice. The strategy chosen is one where two plasmids (see Figure 3.6) are introduced into the plant cell together. One plasmid (pOZ) carries the transgene of interest and the other (pHAG) carries a selectable marker (*hyg*) which confers resistance to the antibiotic hygromycin (selectable markers are described fully in Chapter 4). This plasmid also carries the *gusA* reporter gene (reporter genes are described in Chapter 4) which can be easily assayed histochemically. Both plasmids are coated on to gold particles (these particles are termed the 'microcarrier'), for bombardment into the plant material. The plant material used for transformation is embryogenic callus derived from mature seeds (see Chapter 2).

After bombardment, cells containing the antibiotic resistance gene are selected for on culture medium containing the antibiotic hygromycin (the selective agent). There is no direct selection for the other plasmid (pOZ), but it has been shown that in many cells both plasmids will integrate (dual transformation). In this system then, the aim is to select and regenerate plants in which both plasmids have stably integrated into the genome. This

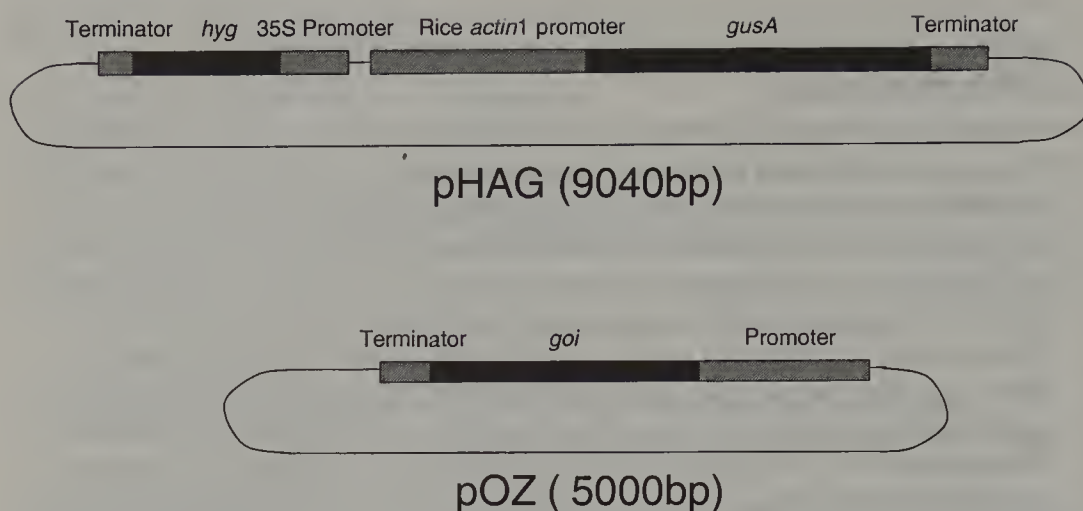


Figure 3.6 Plasmid maps of pHAG and pOZ. The plasmid pHAG contains a selectable marker gene (*hyg*) which is driven by a constitutive promoter (the 35S promoter from cauliflower mosaic virus, which is described in more detail in Chapter 4). The *hyg* gene confers resistance to the antibiotic hygromycin. pHAG also contains a reporter gene (*gusA*) which is driven by the rice *actin1* promoter, a strong promoter in cereals. The *gusA* gene product can be easily assayed and forms a convenient way of confirming that the hygromycin-resistant plant cells are transformed and not 'escapes'. It is the presence of this plasmid that is selected for after particle bombardment by growing the putatively transformed plant tissue on medium that contains hygromycin. The plasmid pOZ is used to deliver the gene of interest, *goi*. There is no selectable marker on pOZ, and therefore the presence of this plasmid cannot be directly selected for. However, dual transformation with both plasmids is a relatively common event. The *ipt* gene is driven by a suitable promoter, the characteristics of which will depend on the required pattern of expression. The promoter may be inducible, tissue specific or constitutive.

approach ultimately allows the selectable marker to be segregated out in subsequent generations. The resulting plants contain the gene of interest, but do not contain the hygromycin resistance gene, which is no longer necessary. The desirability of removing antibiotic resistance marker genes, in particular from crops, is discussed in Chapters 4 and 12.

Preparation of microcarriers

Gold or tungsten particles ('beads' or 'balls') are used as the microcarrier. These need to be pretreated by washing the microcarriers in ethanol and sterile distilled water. Suspensions of the microcarriers can then be stored as aliquots, either at 4°C (gold) or at -20°C (tungsten). Plasmid DNA (isolated from *Escherichia coli*) is attached to the gold microcarriers by mixing the DNA and the microcarriers in 2.5 mol l⁻¹ CaCl₂ and 0.1 mol l⁻¹ spermidine (a polyamine). The microcarriers are then mixed, washed with ethanol and finally resuspended in ethanol. The microcarriers are then applied to the macrocarrier membrane as an ethanol suspension and are allowed to dry on to the macrocarrier.

Preparation of plant material for transformation-callus induction

Dehulled, mature rice seeds are sterilised in ethanol and bleach, and subsequently washed in sterile distilled water to remove any traces of bleach. The dehulled seeds are then placed on MS1 medium (MS medium supplemented with 2.5 mg l⁻¹ 2,4-D, 30 g l⁻¹ maltose and so-

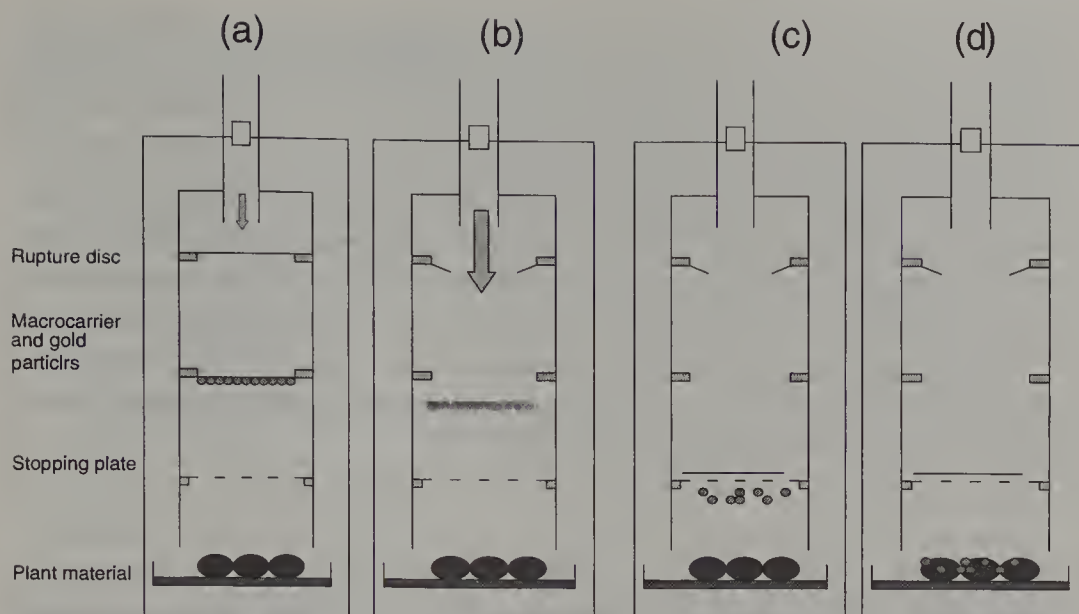


Figure 3.7 The PDS-1000/He particle-bombardment system. The plant tissue is placed into a vacuum chamber (chamber pressure 27 mmHg) 13 cm below the microcarrier stopping plate. The vector DNA-coated particles (the microcarriers) on the macrocarrier membrane are inserted into the apparatus (a). Once the vacuum in the lower part of the apparatus is established, the helium pressure above the rupture disc is increased until at 1100 psi (or whatever pressure the rupture disc is designed to rupture at) the rupture disc bursts (b). This propels fragments of the macrocarrier and the projectiles down the chamber. The macrocarrier is stopped at the stopping plate (c) allowing the microcarriers to pass through and hit the plant material (d). Various parameters can be optimised in this system: the distance between the stopping plate and the plant material can be varied; and the pressure at which the rupture disc bursts can also be selected. Varying these parameters allows the speed and pattern of the microcarriers to be adjusted to suit the needs of the plant material being transformed.

lidified with 2.5 g l⁻¹ Phytigel. Media composition is discussed in detail in Chapter 2). The cultures are then incubated in the dark at 25°C for 2 weeks. Callus tissue that develops from the seeds is transferred to fresh medium of the same composition and incubated for a further 2 weeks for embryogenic callus to develop. The embryogenic callus is transferred (50–70 calluses are arranged on one 9-cm diameter Petri dish) to a different medium (MS medium supplemented with 1 mg l⁻¹ 2,4-D, 14% (w/v) maltose and 2.5 g l⁻¹ Phytigel) and incubated for 4 hours. This medium contains a high concentration of maltose to generate a high osmoticum prior to particle bombardment of the plant tissue.

Bombardment conditions

The PDS-1000/He equipment is assembled as shown in Figure 3.7 and used to bombard the rice embryogenic callus. The operation of the PDS-1000/He equipment is also described in the figure.

Plant selection and regeneration

One day after bombardment, the embryogenic callus is transferred to MS1 medium. After 1 week, the embryogenic callus is transferred to selection medium (MS1 medium

supplemented with 40 mg l⁻¹ hygromycin) and incubated for 2 weeks in the dark. The callus is then transferred to fresh selection medium (of the same composition) and incubated for a further 2 weeks in the dark. Surviving callus is transferred to plant regeneration medium RM1 (MS medium supplemented with 3 mg l⁻¹ kinetin, 30 g l⁻¹ maltose, 40 mg l⁻¹ hygromycin and 6.25 g l⁻¹ Phytigel) and incubated for 3 weeks in the light to encourage shoot formation. Callus is then transferred to RM2 medium (as RM1 but with 2.5 g l⁻¹ Phytigel) for a further 3 weeks' incubation in the light. Green shoots that have formed on RM2 medium are rooted by transferring them on to R1 medium (half-strength MS medium supplemented with 1 mg l⁻¹ naphthalene acetic acid). Young plantlets are then subcultured on to R2 medium (half-strength MS basal medium). The plants are acclimatised in growth chambers and can then be transferred to soil and grown to maturity.

Bombardment has been used mainly for transformation of the nuclear genome, but more recently it has become the main strategy for introducing DNA into plastids (Box 3.3). Chloroplast transformation is a development for the twenty-first century, and its application to GM crops is discussed in detail in Chapters 11 and 12. Suffice it to say that chloroplast transformation offers opportunities for creating GM crops that are both environmentally friendly and more efficient.

Polyethylene glycol (PEG)-mediated transformation

Plant protoplasts (plant cells that have had their cell walls removed, see Chapter 2) can be transformed with naked DNA by treatment with polyethylene glycol (PEG) in the presence of divalent cations (usually calcium). The PEG and the divalent cations destabilise the plasma membrane of the plant protoplast and render it permeable to naked DNA. Once inside the protoplast the DNA enters the nucleus and integrates into the genome.

Plant protoplasts are not easy to work with, and the regeneration of fertile plants from protoplasts is problematical for some species, limiting the usefulness of the technique. The DNA used for transformation is also susceptible to degradation and rearrangement. Despite these limitations, the technique does have the advantages that protoplasts can be isolated and transformed in large numbers from a wide range of plant species.

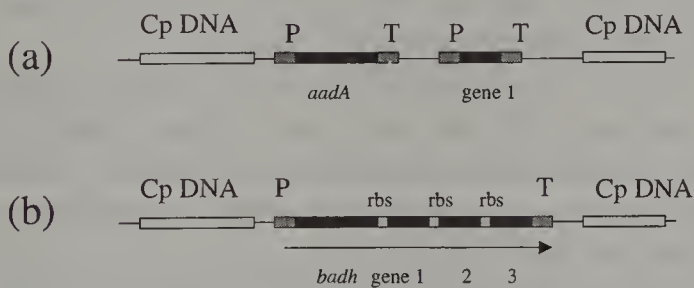
Electroporation

The electroporation of cells can be used to deliver DNA into plant cells and protoplasts. The vectors used can be simple plasmids; the genes of interest require plant regulatory sequences, but no specific sequences are required for integration. Material is incubated in a buffer solution containing DNA and subjected to high-voltage electrical pulses. The DNA then migrates through

BOX 3.3 Plastid transformation

There are a number of important reasons why chloroplast transformation has become an important technique in plant biotechnology. In many, but not all, crop species, chloroplasts display only maternal inheritance so there is no danger of any gene transfer to related weedy species through pollen. Another significant feature of this technology is the very large transgene copy number that may be obtained. Transgene insertion may occur into a single genome contained within an individual chloroplast. However, by using several rounds of selection during the plant regeneration process, it is possible to develop a homoplasmic population of chloroplasts (where all the chloroplasts are transformed). As there may be as many as 100 chloroplasts per cell, each containing up to 100 copies of the genome, there may be as many as 10,000 copies of the transgene per cell. An additional advantage is the absence of reports of gene silencing with chloroplast transformation. This fact, combined with the ability of the chloroplast to correctly fold and cross-link expressed proteins, means there is a tremendous potential in chloroplast transformation for very high-level gene expression and synthesis of active proteins.

This has led to the development of vectors designed specifically for chloroplast transformation. The Figure shows (A) a basic vector for chloroplast transformation in which selection is based on the antibiotic spectinomycin (the selectable marker gene is referred to as *aadA*). The vector is designed for the expression of a single transgene, which is fused to plastid regulatory sequences, such as the promoter from the *psbA* gene and ribosome-binding regions. The foreign genes are bordered by chloroplast DNA sequences. After introduction into the chloroplast by bombardment, homologous recombination occurs between the plastid sequences on the vector and those on the genome. This site-specific integration event avoids any problems associated with random insertion into the nuclear genome.



Vectors for chloroplast transformation. (A) A construct designed for the expression of one gene of interest (gene 1). There is also a selectable marker gene (*aadA*) which confers resistance to the antibiotic spectinomycin. Each gene (gene 1 and *aadA*) is driven by a promoter (P) and has a terminator (T). Flanking the two genes are regions of chloroplast DNA (Cp DNA) that direct integration of the DNA at a defined point in the chloroplast genome by homologous recombination. (B) A construct designed for the expression of multiple transgenes. One promoter (P) drives the expression of a polycistronic transcription unit, or operon. In this case the betaine-aldehyde dehydrogenase gene (*badh*) is used as the selectable marker. Each transgene is preceded by a ribosome-binding site (rbs) to ensure efficient translation. Other features are as in (A). (Redrawn with permission from Daniell, H., *et al.* (2002).)

BOX 3.3 Continued

More advanced vectors have been designed that take advantage of the 'prokaryotic' nature of some of the plastid genome (see Chapter 1). Multiple open reading frames can be introduced as a single transcription unit that is expressed as a polycistronic message (B). This offers a significant advantage over nuclear transformation, where individual genes are introduced separately and may have to be brought together by breeding strategies after plants with similar expression levels have been identified.

Further development is also being carried out on the selection systems used. Strategies have been developed to remove the selectable markers, by techniques such as reversal of the homologous recombination events; some vectors incorporate the use of the betaine-aldehyde selection system, which generates much higher transformation efficiencies (see Chapters 4 and 9).

high-voltage-induced pores in the plasma membrane and integrates into the genome. Electroporation has been successfully used to transform all the major cereals, particularly rice, wheat and maize. Initially, protoplasts were used for transformation, but one of the advantages of the system is that both intact cells and tissues (such as callus cultures, immature embryos and inflorescence material) can be used. This reduces some, but not all, of the tissue culture problems. However, the plant material used for electroporation may require specific treatments, such as pre- and postelectroporation incubations in high osmotic buffers. The efficiency of electroporation is also questionable, it is very dependent on the condition of the plant material used and the electroporation and tissue treatment conditions chosen.

Some studies have indicated that linear, rather than supercoiled, plasmid DNA and the addition of spermidine to the incubation buffer (which can induce condensation of the DNA) improves the efficiency of electroporation. Optimisation of what is known as the field strength (the voltage applied across the electrodes) is also important.

Even under optimal conditions the amount of DNA delivered to each cell is low, which has the benefit of producing transformants with a low transgene copy number. The delivery rate (the proportion of electroporated cells that actually receive DNA) can be high: in some experiments it has been shown that between 40 and 60% of the incubated cells receive DNA, and that as many as 50% of the cells survive the treatment. Electroporation also has the advantage that all the cells are in the same physiological state after transformation, unlike the situation with particle bombardment where transformed cells may be at a disadvantage due to damage from the transformation procedure.

Silicon carbide fibres—WHISKERS™

This is a simple technique for which no specialised equipment is required. Plant material (such as cells in suspension culture, embryos and embryo-derived calluses) is introduced into a buffer containing DNA and the silicon carbide fibres, which is then vortexed. The fibres, which are about 0.3–0.6 μm in diameter and 10–100 μm long, penetrate the cell wall and plasma membrane, allowing the DNA to gain access to the inside of the cell. The drawbacks of this technique relate to the availability of suitable plant material and the inherent dangers of the fibres, which require careful handling.

Although the procedure has been utilised with friable callus from maize, this type of friable callus is limited only to a few genotypes of maize and oats. Many cereals, the natural targets for the procedure, produce embryogenic callus that is hard and compact and not easily transformed, at present, with this technique. Recently though, some progress has been made in transforming such material, and procedures are being developed to allow transformation of cereals such as rice, wheat, barley and maize without the need to initiate cell suspensions.

Summary

A variety of techniques for plant transformation are available to the plant biotechnologist. These techniques can be split into two groups: *Agrobacterium*-mediated transformation; and direct gene transfer methods, of which the biolistics approach is probably the most widely used. These two groups of techniques are fundamentally different in mechanism, and are, in general, applied to different crops. *Agrobacterium*-mediated transformation is most widely used with dicotyledonous crops, which reflects the natural host range of members of the genus *Agrobacterium*. Direct gene transfer methods are most commonly used to transform monocotyledonous crops, such as cereals. In part, this reflects the initial difficulties with using *Agrobacterium* to transform monocotyledonous plants. Direct gene transfer methods and *Agrobacterium*-mediated methods have their own advantages and disadvantages. However, all plant transformation methods can suffer from a problem known as 'gene silencing' (Box 3.4), where transgene (and homologous endogenous genes) expression is actually repressed. Gene silencing is impossible to predict precisely, although some precautions in vector design and transformation protocol can be taken to reduce its frequency. Gene silencing can prove a major hurdle to the commercialisation of plant transformation products.

Despite any problems, improvements in plant transformation technologies, especially when coupled to an efficient plant regeneration protocol, have seen the list of crop species that can be routinely transformed grow. Crops that were once considered impossible to transform (cereals may well fall into this category) are now routinely transformed in many laboratories around the world.

These plant transformation technologies provide the basis for the advances in plant biotechnology discussed in the remainder of this book.

BOX 3.4 Gene silencing

Instability of transgene expression is still a problem encountered in many experiments involving transgenic plants, and is often referred to as 'gene silencing'. Gene silencing can involve a variety of methods and is still relatively poorly understood. Although a complete discussion of gene silencing is beyond the scope of this chapter, a brief description of some of the mechanisms thought to be involved will be given here. A more detailed description of gene silencing is given in Chapter 8.

Two types of gene silencing are of interest to the plant biotechnologist: transcriptional gene silencing (TGS) and post-transcriptional gene silencing (PTGS), both of which are homology-dependent, gene-silencing (HDGS) phenomena. The two types are distinguished according to the level of gene expression at which silencing occurs.

Transcriptional gene silencing

TGS generally occurs when genes share homology in their promoter regions. It usually results in altered methylation patterns and altered chromatin conformation, which result in gene silencing by repressing transcription.

Post-transcriptional gene silencing (co-suppression)

It has been demonstrated that the expression of a homologous transgene can lead to the inhibition of expression of both the transgene and the endogenous gene, primarily by decreasing RNA stability. Thus, instead of resulting in 'overexpression' of a particular gene, expression is actually reduced. The effect appears to increase with increases in the transcription of the transgene. Thus the use of very strong promoters (see Chapter 4) to drive transgene expression may actually lead to more co-suppression, and thus a lower level of expression, than the use of a weaker promoter. PTGS appears to result from the production, either in the nucleus or cytoplasm, of double-stranded RNA. This double-stranded RNA can result from the presence of 'cryptic' transcription start sites producing antisense RNA, transcription through inverted repeats and through the activity of an RNA-dependent RNA polymerase. This double-stranded RNA is cut by a double-stranded RNA endonuclease, resulting in the production of small (21–25 bases) stretches of sense and antisense RNA (termed 'siRNAs', i.e. small interfering RNAs) which can be transported throughout the plant. These siRNAs guide a different ribonuclease to the homologous single-stranded mRNA, which is cut and then further degraded.

One useful 'spin-off' of gene silencing is the development of a technique known as RNAi (RNA interference) where double-stranded RNA is deliberately made to silence gene expression in gene function studies.

Strategies to avoid gene silencing

As gene silencing is difficult to predict, no way of completely avoiding gene silencing during transformation of the nuclear genome has been described (Box 3.3 highlights how gene silencing is not observed with plastid transformation). However, some general recommendations can be made. Reducing the number of transgenes inserted (the copy number), avoiding the use of multiple copies of the same promoter or terminator (particularly in repeat structures) and avoiding using promoters and transgenes with a high degree of similarity to endogenous ones can also help to reduce the probability of gene silencing.

Further reading

General

Barcelo, P., Rasco-Gaunt, S., Thorpe, C. and Lazzeri, P. (2001). Transformation and gene expression. In *Advances in botanical research*, 34: *Biotechnology of cereals* (ed. P. R. Shewry, P. A. Lazzeri and K. J. Edwards), pp. 59–126. Academic Press, London.

Birch, R. G. (1997). Plant transformation: problems and strategies for practical application. *Annual Reviews in Plant Physiology and Plant Molecular Biology*, 48, 297–326.

Hansen, G. and Wright, M. S. (1999). Recent advances in the transformation of plants. *Trends in Plant Science*, 4, 226–31.

Hughes, M. A. (1996). *Plant Molecular Genetics*. Addison Wesley Longman Ltd Harlow, Essex.

Walden, R. and Wingender, R. (1995). Gene transfer and plant-regeneration techniques. *Trends in Biotechnology*, 13, 324–31.

Plant transformation—web-link 3.1: <http://www.uoguelph.ca/~jdberg/plantran.htm>

Transgenic crops: an introduction and resource guide—web-link 3.2:
<http://www.colostate.edu/programs/lifesciences/TransgenicCrops/how.html>

Agrobacterium-mediated transformation

Bent, A. F. (2000). *Arabidopsis in planta* transformation. Uses, mechanisms, and prospects for transformation of other species. *Plant Physiology*, 124, 1540–7.

De la Riva, G. A., Gonzalez-Cabrera, J., Vazquez-Padron, R. and Ayra-Paedo, C. (1998). *A. tumefaciens*: a natural tool for plant transformation. *Electronic Journal of Biotechnology*, 1, 1–16.

Ditt, R. F., Nester, E. W. and Comai, L. (2001). Plant gene expression response to *A. tumefaciens*. *Proceedings of the National Academy of Sciences USA*, 98, 10954–9.

Dumas, F., Duckely, M., Pelczar, P., Van Gelder, P. and Hohn, B. (2001). An *Agrobacterium* VirE2 channel for transferred-DNA transport into plant cells. *Proceedings of the National Academy of Sciences USA*, 98, 485–90.

Gelvin, S. B. (2000). *Agrobacterium* and plant genes involved in T-DNA transfer and integration. *Annual Reviews of Plant Physiology and Plant Molecular Biology*, 51, 223–56.

Ingram, J. (1998). Plant cyclophilins and *Agrobacterium*. *Trends in Plant Science*, 3, 292.

Schrammeijer, B., Risseuw, E., Pansegrau, W., Regensburg-Tuink, T. J. G., Crosby, W. L. and Hooykaas, P. J. J. (2001). Interaction of the virulence protein VirF of *A. tumefaciens* with plant homologs of the yeast Skp1 protein. *Current Biology*, 11, 258–62.

Ward, D. V., Zupan, J. R. and Zambryski, P. C. (2002). *Agrobacterium* VirE2 gets the VIP1 treatment in plant nuclear import. *Trends in Plant Science*, 7, 1–3.

Zupan, J., Muth, T. R., Draper, O. and Zambryski, P. (2000). The transfer of DNA from *A. tumefaciens* into plants: a feast of fundamental insights. *Plant Journal*, 23, 11–28.

Plant transformation—web-link 3.3:

<http://www-ceprap.ucdavis.edu/Transformation/transform1.htm>

The microbial world. Biology and control of crown gall (*A. tumefaciens*)—web-link 3.4:
<http://helios.bto.ed.ac.uk/bto/microbes/crown.htm>

Direct gene transfer

- Daniell, H., Khan, M. S. and Allison, L. (2002). Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology. *Trends in Plant Science*, 7, 84–91.
- Kohli, A., Leech, M., Vain, P., Laurie, D. A. and Christou, P. (1998). Transgene organisation in rice engineered through direct DNA transfer supports a two-phase integration mechanism mediated by the establishment of integration hot spots. *Proceedings of the National Academy of Sciences USA*, 95, 7203–8.
- Sorokin, A. P., Ke, X.-Y., Chen, D. F. and Elliott, M. C. (2000). Production of fertile transgenic wheat plants via tissue electroporation. *Plant Science*, 156, 227–33.

Gene silencing

- Chicas, A. and Macino, G. (2001). Characteristics of post-transcriptional gene silencing. *EMBO Reports*, 21, 992–6.
- Kooter, J. M., Matzke, M. A. and Meyer, P. (1999). Listening to the silent genes: transgene silencing, gene regulation and pathogen control. *Trends in Plant Science*, 4, 340–7.

4

Binary vectors for plant transformation

Introduction

Vectors for plant transformation have to be created in the same way as any other plasmid-derived cloning vector—by using molecular biology techniques. It is therefore useful to look at some of the basic features that are necessary or desirable in any vector. Such features are described here because the usefulness of vectors often equates with the ease with which they can be manipulated, both in standard laboratory hosts such as *Escherichia coli* and *in vitro*, before being used to transform plants. Later in this chapter we will see that many of the recent advances in the design of plant transformation vectors are aimed at improving the ease with which they can be manipulated in the laboratory, rather than their ease of use for plant transformation *per se*.

In addition, this chapter will describe the basic features of vectors used for plant transformation, their development into systems allowing the routine transformation of a range of plant species and some of the problems associated with vector design and transgene expression.

Desirable features of any plasmid vector

Ideally, as well as the ability to replicate independently from the chromosome (which depends on the presence of an origin of replication), cloning vectors should have the following properties:

1. Be of a small size (low molecular weight). Small size has several important advantages. Small plasmids are easier to manipulate *in vitro* as they are less liable to damage by shearing. Small plasmids are also usually present in higher copy number and therefore plasmid yields are higher. There is also less chance of the vector having other sites for restriction endonucleases, making the design and integration of a multiple cloning site (MCS) simpler.
2. Confer a selectable phenotype on the host cells so that transformed cells can be selected for. Most plasmid cloning vectors therefore carry genes that

confer resistance to an antibiotic, very often ampicillin (a penicillin derivative).

3. Contain single sites for a large number of restriction enzymes to enable the efficient production of recombinant vectors.
4. Enable the identification of bacterial colonies containing recombinant plasmids. This is usually achieved by clustering the unique restriction sites in one small area, the multiple cloning site, which is situated in the *lacZ* α sequence that encodes the N-terminal part of *E. coli* β -galactosidase. In suitable *E. coli* hosts, complementation results in the production of active β -galactosidase, which is detected by the bacterial colony turning blue due to hydrolysis of a chromogenic substrate. Cloning DNA into the MCS effectively results in the insertional mutagenesis of *lacZ* α . After transformation with these recombinant plasmids, *E. coli* will produce white colonies as no complementation can occur. Although not 100% reliable (both false-positives and false-negatives can be generated), such a system drastically reduces the number of colonies that need to be screened in order to identify those harbouring recombinant plasmids.

These features are illustrated in Figure 4.1, which shows a typical cloning vector widely used in molecular biology.

Development of plant transformation vectors

As the process of *Agrobacterium*-mediated plant transformation was elucidated (see Chapter 3) important information came to light that made the development of efficient plant transformation vectors possible. Chief amongst these were the following discoveries: that the only features of the T-DNA necessary for integration into the host plant genome were the short border sequences; the discovery that the removal of the oncogene sequences enabled plants to be regenerated from transformed plant tissue by manipulating the plant hormone composition of the medium; and the fact that the *vir* genes function in *trans*.

Basic features of vectors for plant transformation

Some of the basic factors that contribute to the design of a successful vector have already been considered. However, when considering the design of plant transformation vectors, several additional degrees of complication become apparent:

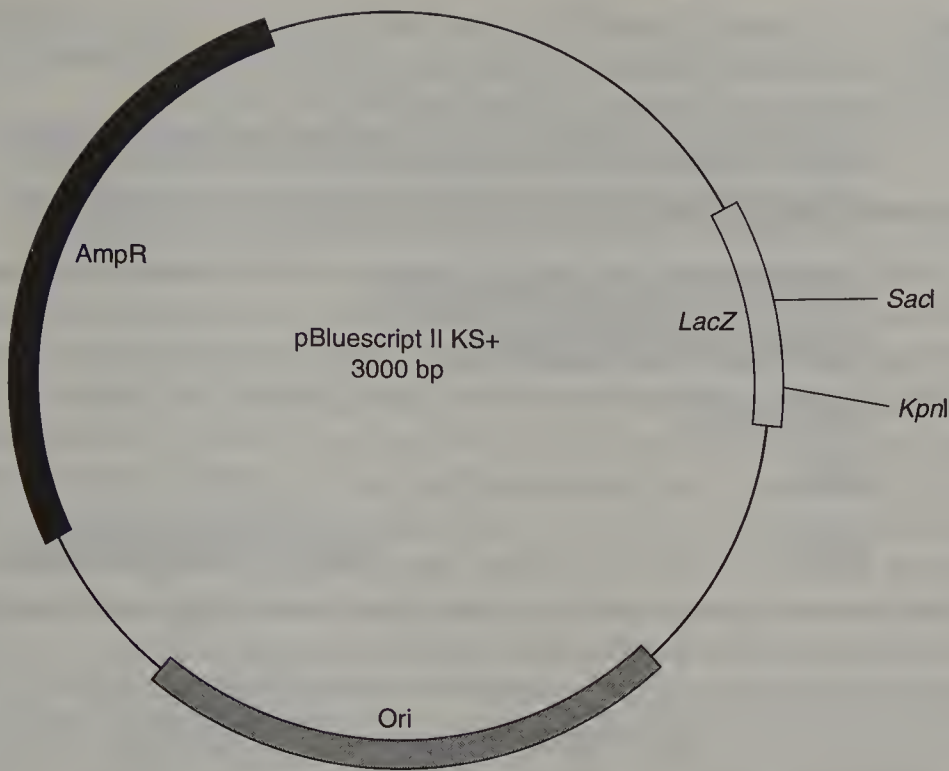


Figure 4.1 Diagram of a typical ‘standard’ cloning vector (pBluescript II KS+ in this example, although many fundamentally similar products exist) commonly used in molecular biology. The following features are worth noting:

1. It is of a small size (3 kb) to facilitate ease of handling and high transformation efficiencies.
2. It has an origin of replication (a pUC *ori*) to enable replication of the plasmid to high copy number in *E. coli*.
3. It has an antibiotic resistance marker to facilitate the recovery of transformed bacteria.
4. It has a multiple cloning site in the *lacZα* fragment to ease cloning and facilitate the identification of recombinant plasmid-containing colonies by blue/white screening. The multiple cloning site (MCS) is between the *SacI* and *KpnI* sites (in this particular plasmid there are restriction sites for over 20 other restriction enzymes in the MCS).

(Redrawn with permission from the pBluescript II KS+ figure on the Stratagene web site (<http://www.stratagene.com>).)

1. The plasmid (if it is to be used in *Agrobacterium*-mediated transformation) needs to be able to replicate not only in *E. coli* (so that routine manipulations can be carried out) but also in *Agrobacterium*.
2. Additional selectable markers need to be included so that the successfully transformed plants can be identified.
3. Border sequences need to be incorporated into the design of plasmid vectors for *Agrobacterium*-mediated transformation to ensure integration of the genes of interest into the host plant genome.
4. The genes (particularly if they are from prokaryotes or non-plant eukaryotes) that are to be integrated into the genome of the host plant may

need to be made ‘plant-like’. This includes the use of appropriate promoters and terminators to ensure that expression of the genes occurs. These features are often incorporated into the basic vector.

Promoters and terminators

An obvious requirement for any genes that are to be expressed as transgenes in plants is that they are expressed correctly (or at least in a predictable fashion). It is known that the major determinant of gene expression (level, location and timing) is the region upstream of the coding region. This region, termed ‘the promoter’, is therefore of vital importance (see Chapter 1). Any genes that are to be expressed in the transformed plant have to possess a promoter that will function in plants. This is an important consideration as many of the genes that are to be expressed in plants, particularly reporter genes and selectable marker genes, are bacterial in origin. They therefore have to be supplied with a promoter that will drive their expression in plants.

Transgenes also need to have suitable terminator sequences at their 3′ terminus to ensure that transcription ceases at the correct position. Failure to stop transcription can lead to the production of aberrant transcripts and can result in a range of deleterious effects, including inactivation of gene products and increased gene silencing.

In addition to the basic need for the promoter to be capable of driving expression of the gene in plants, there are other considerations that need to be taken into account, such as promoter strength, tissue specificity and developmental regulation.

Agrobacterium-derived promoter and terminator sequences

The genes from the Ti plasmid of *Agrobacterium* that code for opine synthesis, and in particular the nopaline synthase (*nos*) gene, are widely used as a source of both promoters and terminators in plant transformation vectors. Although derived from bacterial genes, their presence on the T-DNA means they are adapted to function in plants. The *nos* promoter is usually considered to be constitutive.

The 35S promoter

The most widely used promoter used to drive expression of genes in plant transformation vectors is the promoter of the cauliflower mosaic virus 35S RNA gene (35S promoter). This promoter is considered to be expressed in all tissues of transgenic plants (though not necessarily in all cell types). In dicots it drives expression at high levels, although in monocots the level of expression is not as high. This makes the 35S promoter ideal for driving the expression of selectable marker genes, and in some cases of reporter genes, as expression is more or less guaranteed. The activity of the 35S promoter can be further increased by the inclusion of one or more copies of the enhancer region (Figure 4.2 shows the structure of the 35S promoter). High-level constitutive

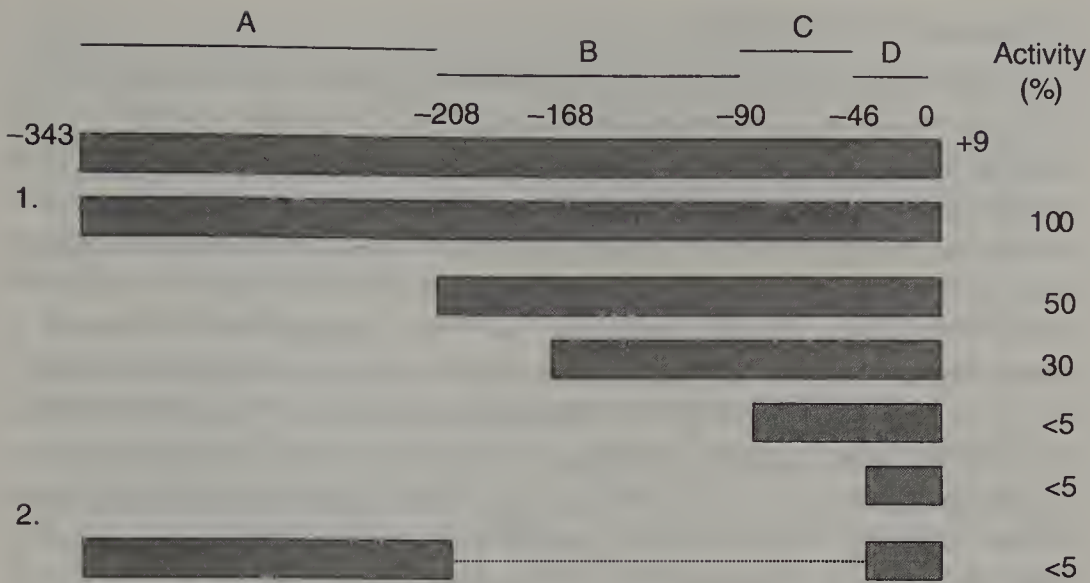


Figure 4.2 Structure/activity relationships in the 35S promoter. Deletion analysis of the 35S promoter. The upper most box represents the 35S promoter region as defined in Fang *et al.* (1989), including 343 base pairs of upstream 'promoter'. The letters above designate four functional regions that can be identified in the 35S promoter. The figures above the upper most box give the position of the end of the promoter fragments (5' deletions) used in the deletion analysis.

1. The boxes below represent the promoter fragments used to assess the activity (CAT activity of various promoter fragments).
2. The broken line represents a piece of the promoter 'removed' from a fragment, the two flanking pieces (-343 to -208 and -46 to 0) being ligated together to form a single promoter fragment.

Functional regions:

- A. (-343 to -208). Has no activity in itself (activity depends on the presence of the -90 to -46 region). Responsible for about 50% of the promoter activity.
- B. (-208 to -90). Functions as an enhancer and is responsible for most of the rest of the activity of the promoter.
- C. (-90 to -46). Has an accessory role in increasing the transcriptional activity of regions A and B.
- D. (-46 to 0). The core promoter region, containing the TATA box.

The -208 to -46 region of the promoter is often used in transformation vectors as an enhancer to increase the expression of the 35S promoter. Increasing the number of copies of the enhancer increase the transcriptional activity linearly, until a maximum stimulation of activity is reached with four copies. The enhancer also functions with heterologous promoters to increase transcriptional activity. The -343 to -90 region is primarily responsible for promoter activity in leaves, while the -90 to -46 region is required for expression in roots. (Data taken with permission from Fang, R.-X., *et al.* (1989).)

expression is not always a desirable feature though, particularly if the gene product is toxic in large amounts.

In monocots, alternatives, such as the maize ubiquitin I promoter or the rice actin promoter/first intron sequence, are often used to drive the high-level expression of transgenes.

Tissue-specific promoters

Considerable effort has been put into isolating promoters that can be used to drive expression in a tissue-specific manner. Such promoter systems are potentially useful in a number of situations. The expression of any potentially harmful substances can be limited to tissues that are not consumed by animals or humans, and genes involved in specific processes can be limited to tissues in which that process occurs. It is beyond the scope of this chapter to look at all the different tissue-specific promoters that have been isolated. However, it should be noted that several (some examples will be seen in later chapters) have been used to successfully drive transgene expression in the predicted pattern. It should also be noted, however, that there are examples where promoters have been found not to function, or not to drive expression in the predicted pattern, in heterologous systems. Considerable care has therefore to be taken with the use of promoters. Unless the promoter is well characterised and widely used in the plant species you intend to transform, characterisation of the expression pattern conferred by the promoter (promoter::reporter gene fusions could be used) should be carried out. This is particularly so as it becomes apparent that many *cis*-acting regulatory sequences are, in fact, found in regions other than the 'promoter'.

Inducible promoters

Considerable effort has also been put into isolating and developing inducible promoter systems for use in plants. Inducible expression systems can be divided into three categories: (1) non-plant-derived systems; (2) plant-derived systems that respond to environmental signals; and (3) plant-derived systems based on developmental control of gene expression (examples are given in Box 4.1). Theoretically, such systems allow the timing of transgene expression to be carefully controlled.

The advantage of non-plant-derived systems is that they are independent of normal plant processes, requiring the application of a specific substance to induce expression. This is, however, also their weakness, as the application of (in some cases expensive) inducers on an agricultural scale is, in most, but not all cases, not feasible or economically viable (see Box 4.1 for an example where it may prove to be both feasible and economically viable). Experimentally, however, some of these have proved to be quite useful.

Systems based on plant-derived components do not have the advantage of independence from normal plant processes. This does, however, make their use in agriculture potentially simpler as the application of an inducer is not required.

It should be noted that no one system is suited to all situations, and potentially more damaging is the fact that some of the systems developed have proved to be difficult to use or have proved to be 'leaky' (i.e. they direct expression even in the absence of the inducing agent).

BOX 4.1 Systems for inducible gene expression in plants

Inducible expression gives finer control over transgene expression than does a constitutive promoter such as the 35S promoter. Several systems have been developed with the aim of giving inducible transgene expression in plants. Some examples will be considered here.

Non-plant-derived systems

Some features are considered to be desirable in any inducible expression system that depends on the application of an exogenous chemical. These are:

1. There should be no expression of the transgene in the absence of the inducer.
2. The system should respond to only one inducer, or one class of inducer, so that the system is specific to the inducer.
3. Induction of gene expression should be rapid following the application of the inducer.
4. Gene expression should cease rapidly following withdrawal of the inducer.
5. The inducer should be non-toxic.
6. The inducer should not cause non-specific changes in gene expression.

Some examples of inducible expression systems are given below.

Tetracycline

The antibiotic tetracycline can be used to either de-repress or inactivate gene expression. In bacteria, the tetracycline repressor (TetR) binds to the *tet* operator and negatively regulates its expression. Upon binding to tetracycline the TetR is modified and released from the operator. This system has been adapted to function in plants. The TetR is constitutively overexpressed from a 35S promoter. The transgene is under the control of a chimeric promoter consisting of a core 35S promoter and several copies of the *tet* operator. The TetR normally represses expression of the transgene, but with the application of tetracycline it is released from the operators in the promoter of the transgene. This then results in transgene expression. This system was one of the first to be developed, but it does have some drawbacks. It does not appear to work in all plant species and tetracycline has to be applied constantly to ensure transgene expression, as tetracycline has a short half-life.

Tetracycline can also be used to inactivate transgene expression. A similar promoter is used to control the expression of the transgene, but the TetR is modified to convert it to an activator (tetracycline transactivator or tTA). The tTA binds to the operators and induces gene expression in the absence of tetracycline. When tetracycline is added to the system, the tTA is released and transgene expression ceases.

Alcohol (ethanol) inducible

In this system a chimeric promoter is again used to drive transgene expression. In addition to a core 35S promoter region, the promoter contains binding sites for the AlcR transcription factor (from *Aspergillus nidulans*), which is constitutively

BOX 4.1 Continued

expressed. Upon application of ethanol to the system, AlcR binds to the AlcR-binding sites and activates transcription.

Steroid inducible

A variety of systems conferring steroid inducibility on transgene expression exist. Most consist of a modified transcriptional activator capable of binding a steroid hormone or its analogue. Binding sites for the modified transcription factor are included in a chimeric promoter, which controls transgene expression. In the absence of an inducer the transcription factor is inactive, but on application of the inducer it binds to the promoter and activates expression. Systems that respond to different steroidal inducers have been developed. Amongst these are ones that respond to glucocorticoids (the synthetic analogue dexamethasone is used), oestrogen and ecdysone (an insect hormone). The dexamethasone system appears to have some serious drawbacks that may limit its usefulness. The development of the ecdysone-based system is, however, particularly interesting as it also responds to the non-steroidal ecdysone agonist RH5992 (tebufenozide). RH5992 is a widely used agrochemical, which opens up the possibility of using this system in field situations.

Copper inducible

This system is based on the yeast metallothionein regulatory system. A transcription factor is constitutively expressed. Upon binding copper, the transcription factor binds to elements in a chimeric promoter and activates expression.

Plant-derived systems based on response to environmental signals

Although not inducible in the same sense as the non-plant-derived systems, some systems have been developed that respond to a variety of environmental signals. These systems depend on either the construction of chimeric promoters or on the use of complete promoters from wound inducible genes to control the expression of transgenes. The chimeric promoters usually consist of a core promoter element (usually from the 35S promoter) and sequence elements that bind transactivating factors that respond to particular environmental signals. Several systems are briefly described here.

Wound inducible

Many wound inducible genes have been identified, and the promoters of some of these, particularly some proteinase inhibitors, characterised in some detail. The promoters of these genes have been found to confer wound inducibility on transgene expression. The wound inducibility may be mediated by other factors such as methyl-jasmonate (see also Chapter 1), which can be used to mimic the wound response in some cases. Wound inducible promoters can be used to drive the expression of pest-resistance genes after insect damage.

Heat-shock inducible

Sequence elements, termed 'HSEs' (heat-shock elements) mediate the heat-shock inducibility of genes (see Chapter 9). If included in chimeric promoters they confer heat-shock inducibility on transgene expression.

BOX 4.1 Continued**Plant-derived systems based on developmental control of gene expression**

Again, although not inducible in the sense of the non-plant-derived systems, many genes have been identified that are expressed at particular stages in plant development. The promoters from some of these genes have been studied in detail. As in the previous cases, either chimeric promoters or complete promoters are used to drive transgene expression.

Senescence-specific gene expression

The promoters from two *Arabidopsis* genes, termed 'SAG12' and 'SAG13' (senescence associated gene), which are induced during senescence, have been shown to confer senescence inducible expression on transgenes. For example, expressing the *ipt* gene from *Agrobacterium* (see Chapter 3), which is involved in cytokinin production, under the control of the SAG12 promoter delays senescence in tobacco.

Abscissic acid (ABA) inducible gene expression

Sequence elements responsible for ABA (see Chapter 2) induction of gene expression have been identified (see Chapter 1) and if combined with the core 35S promoter in a chimeric promoter confer ABA inducibility on transgene expression.

Auxin inducible gene expression

Similarly, several sequence elements responsible for auxin (see Chapter 2) induced gene expression have been identified. These auxin-response elements (AuxREs) when combined with a core 35S promoter confer auxin (and inactive auxin analogue) inducibility on transgene expression.

Selectable markers

Plant transformation is, in many cases, a very-low frequency event. It is therefore vital that some means for selecting the transformed plant tissue is provided by the plant transformation vector. In most cases this selection is based on the inclusion into the culture medium of a substance that is toxic to plants. The selectable marker on the vector confers resistance to the toxic substance when expressed in transformed plant tissue. In many cases, as in the situation when using *E. coli*, antibiotic resistance genes can be used as selectable markers in plants. Although plants are eukaryotic, antibiotics efficiently inhibit protein synthesis in the organelles, particularly the chloroplasts. Perhaps the most widely used selectable marker gene is the *nptII* gene that confers resistance to the antibiotic kanamycin. Other antibiotic resistance genes have also been used with some considerable success in plant transformation vectors. In part, this use of alternative selectable marker genes was driven by the observations that some plant species exhibited a very high degree of natural

resistance to kanamycin (such as cereals), and that some species (some soft fruits for example) were too sensitive to kanamycin for it to be used successfully. This made the selection of transformed tissue very difficult, leading to a high number of false-positives or the inability to recover transformed plants. Thus the use of any selectable marker has to be closely controlled, with appropriate concentrations being determined by kill-curves, on a case-by-case basis. The development of alternative selectable marker genes also allows for the re-transformation of plant tissue that already expresses one or more different selectable markers. Thus genes conferring resistance to antibiotics such as bleomycin, spectinomycin and hygromycin are used quite widely. These selective agents can be used at lower concentrations than kanamycin, and therefore usually result in a cleaner selection of transformed tissue.

Other resistance genes can also be used as selectable markers in plants. Amongst those widely used are genes that confer resistance to herbicides such as chlorsulphuron and bialaphos (see Chapter 5).

Public concern has, in recent years, questioned the general desirability of growing transformed crops, and, more specifically, the merits of using antibiotic or herbicide resistance genes as selectable markers during plant transformation. There is obvious concern about the creation of so-called 'super weeds' and the transfer of antibiotic resistance genes. Although in the long term it is undoubtedly better to remove these selectable marker genes from the transformed plants once their job has been done (this issue will be considered in more detail later in this chapter), other, more acceptable, selectable marker genes are also being introduced.

Some of the most interesting are based on the principle of facilitating alternative carbon-source utilisation. Thus genes from bacteria that allow the use of mannose or xylose (and, in fact, other carbohydrates) as carbon sources have been successfully used as selectable markers in plant transformation. These genes have also generally proved to be superior to standard antibiotic selection genes in some transformation protocols. These are examples of 'positive-selection' methods, so-called because untransformed tissue does not die (due to the presence of low concentrations of sucrose in the medium used for regeneration) but does not proliferate.

Other potential selectable markers are also being investigated, these include: the reporter gene *gfp*, which can be used for the visual screening of transformants; the *Agrobacterium* isopentenyl transferase (*ipt*) gene, which determines growth and morphology during regeneration; and the use of cytokinin glucuronides in combination with β -glucuronidase. Examples of various selectable markers are shown in Table 4.1.

Reporter genes

Reporter genes (see Table 4.2) are widely used in plant transformation vectors, both as a means of assessing gene expression by promoter analysis (see

Table 4.1 Selectable marker genes used in plant transformation, their source and mode of action

Selectable marker gene	Abbreviation	Source of gene	Selection mechanism	Selective agent
<i>Antibiotic resistance</i>				
Aminoglycoside adenyltransferase	<i>aadA</i>	<i>Shigella flexneri</i>	Antibiotic resistance	Streptomycin Spectinomycin
Bleomycin resistance	<i>ble</i>	<i>E. coli</i>	Antibiotic resistance	Bleomycin
Dihydropteroate synthase	<i>sul/dhps</i>	<i>E. coli</i>	Antibiotic resistance	Sulphonamides
Dihydrofolate reductase	<i>dhfr</i>	Mouse	Antibiotic resistance	Methotrexate
Hygromycin phosphotransferase	<i>hpt/aph IV/hyg</i>	<i>E. coli</i>	Antibiotic resistance	Hygromycin
Neomycin phosphotransferase II	<i>nptII/neo</i>	<i>E. coli</i>	Antibiotic resistance	Kanamycin
Neomycin phosphotransferase III	<i>nptIII</i>	<i>Streptococcus faecalis</i>	Antibiotic resistance	Geneticin (G418) Kanamycin Geneticin (G418)
<i>Herbicide resistance</i>				
Acetolactate synthase	<i>als</i>	<i>Arabidopsis</i> spp./maize/tobacco	Herbicide resistance	Sulphonylureas
Enolpyruvylshikimate phosphate synthase	<i>epsps/aroA</i>	Petunia hybrida/ <i>Agrobacterium</i> spp.	Herbicide resistance	Glyphosate
Glyphosate oxidoreductase	<i>gox</i>	<i>Achromobacter</i> LBAA	Herbicide resistance	Glyphosate
Phosphinothricin acetyltransferase	<i>bar/pat</i>	<i>Streptomyces hygroscopicus</i> / <i>S. viridochromogenes</i>	Herbicide resistance	Bialaphos Glufosinate L-phosphinothricin
Cyanamide hydratase	<i>cab</i>	<i>Myrothecium verrucaria</i>	Herbicide resistance	Cyanamide

Table 4.1 Continued

Selectable marker gene	Abbreviation	Source of gene	Selection mechanism	Selective agent
Bromoxynil nitrilase	<i>bxn</i>	<i>Klebsiella pneumoniae</i>	Herbicide resistance	Bromoxynil
Others				
Lysine-threonine aspartokinase	<i>lysC</i>	<i>E. coli</i>	Resistance to high levels of lysine or threonine	Lysine and threonine
Mannose-6-phosphate isomerase	<i>pmi/manA</i>	<i>E. coli</i>	Alternative carbon source	Mannose
Xylose isomerase	<i>xylA</i>	<i>Thermoanaerobacterium thermosulfurogenes</i> <i>Streptomyces rubiginosus</i> <i>E. coli</i>	Alternative carbon source	Xylose
β -Glucuronidase	<i>gus/hidA</i>		Complementation of missing medium component	Cytokinin glucuronide
Betaine aldehyde dehydrogenase ^a	<i>badh</i>	Spinach	Detoxification	Betaine aldehyde
Green fluorescent protein	<i>gfp</i>	<i>Aequorea victoria</i> (jellyfish)	Visual screening	None
Isopentenyl transferase ^b	<i>ipt</i>	<i>Agrobacterium</i> spp.	Growth in absence of exogenous cytokinin	None
Leafy cotyledon ^c	<i>lec</i>	Maize	Growth	None

^a See Chapters 9 and 12. Used as a selectable marker in chloroplast transformation.

^b If used with a constitutive promoter, plants have poor root formation and are often sterile, thus limiting the use of the *ipt* gene as a selectable marker. However, use with an inducible promoter has resulted in the development of an antibiotic-free selectable marker system.

^c Leafy cotyledon gives the transformed cells a competitive growth advantage, and when coupled with a visual marker to identify transformed cells can be used as an antibiotic-free selectable marker system.

Chapter 1) and as easily scored indicators of transformation (indeed, they have been used in place of selectable marker genes in some cases).

Ideally, reporter genes should be easy to assay, preferably with a non-destructive assay system, and there should be little or no endogenous activity in the plant to be transformed.

At present, only a small number of reporter genes are in widespread use in plant transformation vectors, these being β -glucuronidase (*uidA* or *gus*), green fluorescent protein (*gfp*), luciferase genes (*lux* and *luc*) and, to a lesser degree (although it is widely used in animal systems), the chloramphenicol acetyltransferase gene (*cat*). Each of these reporter genes will be looked at in more detail.

β -Glucuronidase

This is perhaps the most widely used reporter gene in plant transformation vectors. Its widespread acceptance is due to its many advantages over the previously existing marker genes. β -Glucuronidase can be assayed extremely sensitively using quick, easy and non-radioactive methods. It can be used to obtain both quantitative (i.e. the level of gene expression) and qualitative (i.e. localisation of gene expression) data. There is also little or no endogenous activity in most plant tissues (with the possible exception of reproductive tissues).

Quantitative data is obtained by assays utilising fluorogenic substrates such as 4-MUG (4-methylumbelliferyl- β -D-glucuronide) which is hydrolysed to 4-MU (methylumbelliferone). Standard enzyme assay protocols can be used and results compared with a standard curve of 4-MU fluorescence.

Qualitative data can be obtained from histochemical assays that allow tissue- and cell-specific localisation of β -glucuronidase (GUS) activity. These histochemical assays are conducted *in situ* with the chromogenic substrate X-gluc (5-bromo-4-chloro-3-indolyl β -D-glucuronide). GUS activity results in the deposition of an insoluble blue precipitate, effectively identifying the precise location of the expression.

In order to avoid problems associated with GUS expression in *Agrobacterium*, a GUS gene containing an intron (which can therefore only be expressed in eukaryotes) is often used in plant transformation vectors.

Green fluorescent protein (GFP)

Nowadays, *gfp* is rapidly becoming a very widely used reporter gene. GFP has the advantages that it is even easier to assay than GUS (given the correct equipment) and the assay is non-destructive (although non-destructive assays for GUS activity have been reported, they have not been widely adopted). This means that GFP can be used in situations where GUS cannot, for example in screening primary transformants, in time-course experiments, or analysing segregation in small seedlings.

The gene was isolated from the jellyfish *Aequorea victoria*, which are brightly luminescent organisms. In order to work efficiently in plants it was found that the *gfp* gene had to be significantly modified in order to: (1) remove a cryptic intron (i.e. *gfp* mRNA is efficiently *mis*-spliced in some plants, resulting in the removal of 84 nucleotides); (2) make the codon usage more ‘plant-like’; and (3) prevent accumulation in the nucleoplasm. (See Box 4.3.)

Luciferases

The firefly luciferase gene (*luc*) encodes an enzyme that catalyses the oxidation of D-luciferin in an ATP-dependent fashion. This oxidation results in the emission of light. Highly sensitive assays (based on photomultipliers, luminometers or film exposure) have been developed to detect the extremely rapid emission of light. Other assay systems are also commercially available. The firefly luciferase gene is not a widely used marker gene as assaying the gene product is difficult, but it is useful for the detection of low-level or highly localised expression.

The bacterial luciferase genes (*luxA* and *luxB* from *Vibrio harveyi*) are found in a few plant transformation vectors. They catalyse the oxidation of long-chain fatty aldehydes resulting in the emission of light.

Chloramphenicol acteyltransferase

Although widely used as a reporter gene in mammalian cells the availability of the GUS and GFP reporter systems has generally limited the use of the CAT system in plants, although it was the first bacterial gene to be expressed in plants. Some plant transformation vectors do, however, carry this reporter gene as it can be assayed very sensitively, but it requires a radioactive assay procedure.

Table 4.2 Reporter genes used in plant transformation

Reporter gene	Abbreviation	Source of gene	Detection/assay
β-Glucuronidase	<i>gus/luidA</i>	<i>E. coli</i>	Fluorimetric (quantitative) or histochemical (<i>in situ</i>), non-radioactive
Green fluorescent protein	<i>gfp</i>	<i>Aequorea victoria</i> (jellyfish)	Fluorescence, non-destructive
Chloramphenicol acetyltransferase	<i>cat</i>	<i>E. coli</i>	Radioactive assay of plant extract, sensitive, semi-quantitative
Luciferase	<i>luc</i>	<i>Photinus pyralis</i> (firefly)	Luminescence
Luciferase	<i>luxA, luxB</i>	<i>Vibrio harveyi</i>	Luminescence

Origins of replication

As both *E. coli* and *Agrobacterium* are Gram-negative bacteria, a single, broad host-range origin (*ori*) can be used to allow vector replication in both *E. coli* and *Agrobacterium*. However, the broad host-range origins used in many plant transformation vectors (e.g. the RK2 origin used in pBIN19 and derivatives) often results in a low plasmid copy number in *E. coli*, making manipulation of the DNA more difficult or time-consuming.

An alternative approach is to use two separate origins of replication, one to enable replication in *Agrobacterium* and a second to direct replication in *E. coli*. This second origin is usually derived from a high copy number cloning vector (such as the *ColE1 ori* from pBR322 or the pUC *ori*). This results in a higher copy number in *E. coli*, therefore facilitating manipulation, but it also results in a larger vector (as two separate origins are present) with the associated difficulties.

Co-integrative and binary vectors

Initially, vectors used for *Agrobacterium*-mediated plant transformation were deletion derivatives of Ti plasmids that were termed ‘co-integrative vectors’. The DNA to be introduced into the plant transformation vector is cloned into an intermediate vector (which is small and easily manipulated *in vitro*) and recombination within the bacterium is used to clone the sequences. In most systems of this type, the modified T-DNA (containing the gene of interest) is on the same molecule as the transfer apparatus as part of one Ti plasmid. Vectors of this type have become less common as binary vector systems are more easily manipulated.

In binary vectors, the transfer apparatus (i.e. the *vir* genes) and the T-DNA are located on separate plasmids (Figure 4.3). As only the border sequences are needed to define the T-DNA region and the *vir* region is absent, binary vectors are relatively small (and therefore easily manipulated). Box 4.2 describes the development of binary vectors.

Families of binary vectors

It soon became apparent that, although vectors such as pBIN19 (which offer one selectable marker) were sufficient for routine transformation studies, the application of transgenic technology to an increasing number of plant species required a more flexible set of binary vectors.

Thus modern binary vectors tend to come in ‘families’. These offer a choice of selectable marker (both antibiotic and herbicide resistance in many cases) and, in some cases, the opportunity for dual selection by incorporating two selectable markers on the one vector. This dual selection approach allows a much more rigorous selection pressure to be applied to plants, in which it has

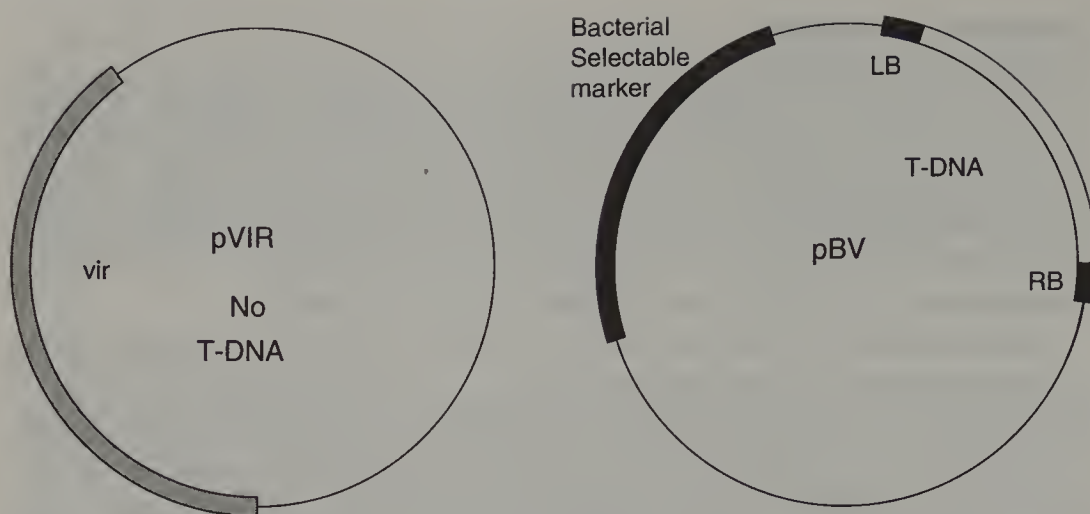


Figure 4.3 The binary vector strategy illustrated by a pair of hypothetical plasmids. One plasmid (pVIR in this case, a derivative of a Ti plasmid) carries the *vir* genes. This plasmid has no T-DNA region, including border sequences, and therefore no T-DNA can be transferred to the plant genome. A smaller binary vector (pBV, which is over 20 times smaller than pVIR) has a T-DNA region defined by the presence of the left and right borders (LB, RB). As there are no *vir* genes on this plasmid (making the binary vector pBV relatively small), their function is supplied *in trans* by pVIR. pVIR would normally be maintained in suitable strains of *Agrobacterium* and is not generally isolated or manipulated. Binary vectors are designed to be easily manipulated using standard *in vitro* techniques and contain selectable markers for selection in bacteria (*E. coli* and *Agrobacterium*) outside the T-DNA region. The T-DNA region will, in all probability, contain a selectable marker for use in plants. pBV is normally maintained in *E. coli*, the gene of interest (the transgene) being cloned into it. Recombinant *E. coli* colonies are identified after bacterial transformation, and the pBV vector is checked to ensure that the transgene is correctly inserted and has not been mutated. pBV can then be transferred, by electroporation, into the *Agrobacterium* strain used for plant transformation (which also harbours pVIR).

traditionally proved difficult to avoid false-positives during the selection process.

Optimisation

Despite the improvements made in vector design and advances in our understanding of both the mechanisms of transgene integration and plant gene expression, plant transformation is still in many ways an imprecise art. Several aspects of vector design are known to influence the efficiency of transgene expression.

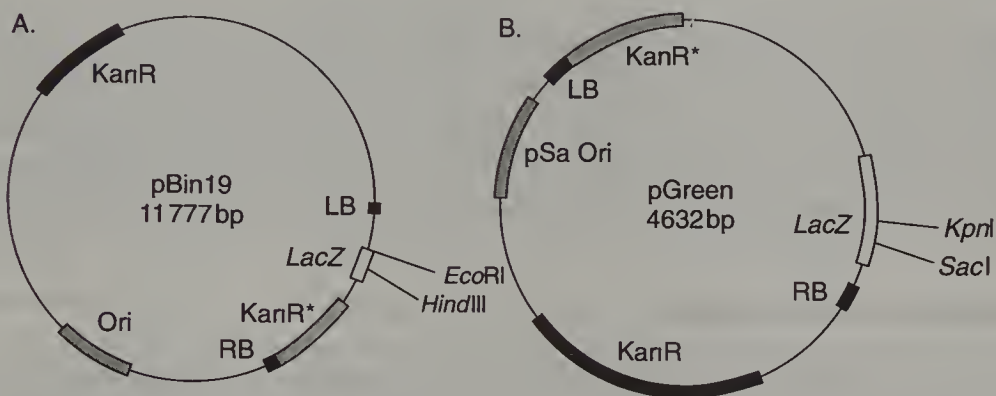
Arrangement of genes in the vector

Analysis of the efficiency of transgene expression allows a few simple ‘rules’ about vector design to be drawn up. If more than one gene is in the vector,

BOX 4.2 The evolution of binary plant transformation vectors

Having looked at the basic requirements for plant transformation vectors we can look at how the design of plant transformation vectors has improved over the years. The binary plant transformation vectors pBIN19 and pGreen can be compared (see Figure).

pBIN19 was one of the first binary plant transformation vectors (developed in the early 1980s) and has been one of the most widely used. pGreen is a new binary vector (one of a family available with different promoters, selectable markers and



Simplified representation of the binary plant transformation vectors (A) pBIN19 and (B) pGreen (in this case pGreen 0029), showing the pertinent features.

pBIN19 is a large plasmid, making manipulations using this plasmid difficult. The selectable marker for selecting transformed plants (KanR*) is located close to the right border (RB), leading to the possibility of false-positives (due to truncation of the T-DNA during transfer, kanamycin-resistant plants that do not express the transgene might be isolated). The multiple cloning site (MCS), situated between the *EcoRI* and *HindIII* sites has 8 restriction enzyme sites. The MCS is in the *LacZ* fragment so blue/white screening for recombinant colonies is possible.

pGreen 0029 is one of a large family of plant transformation vectors. It is much smaller than pBIN19, has a much larger MCS (with over 15 restriction enzyme sites) situated in *LacZ* (between the *KpnI* and *SacI* sites), so blue/white screening is possible. The pSa origin is small and contributes to the small overall size of pGreen 0029.

(Redrawn with permission from figures on the pGreen web-site (http://www.pgreen.ac.uk/a_hom.htm).)

reporter genes). As can be seen, pBIN19 (11 777 bp) is much larger than pGreen (4632 bp) and is therefore that much more difficult to work with. This reduction in size is, at least in part, due to the unusual origin of replication present in pGreen. The pSa origin is in itself small, but in pGreen it has been further divided so that only the replication origin is present on the plasmid. The replicase gene is present on another plasmid and functions in *trans*. The restriction sites in pBIN19 available for use in cloning are limited in comparison with pGreen, although in both cases blue/white screening is possible. The selectable marker, which in both cases is *nptII*, is situated close to the right border in pBIN19, but in pGreen it is situated close to the left border. The advantage of having the selectable marker

BOX 4.2 Continued

next to the left border becomes apparent if one considers that there is a polarity in T-DNA transfer (Chapter 3). Transfer originates from the right border and therefore the selectable marker will be transferred before the gene of interest if the selectable marker is next to the right border. If T-DNA transfer is interrupted this could result in plants expressing the selectable marker (and therefore being isolated as genuine positive transformants) but not containing, or containing a truncated version of, the gene of interest.

different promoters and terminators should be used for each. The use of the same promoter and/or terminator can lead to an increase in gene silencing. Multiple genes on one vector should not be immediately adjacent to each other, and should be in the same orientation. This avoids adjacent inverted repeats that cause plasmid instability in bacteria and increased gene silencing in plants.

Transgene copy number

In any plant transformation protocol, multiple copies of the transgene can be incorporated into the target plant genome. Often high levels of transgene expression are associated with multiple copies of the transgene, but this is not always the case. Multiple T-DNA insertions into the genome may lead to erratic transgene expression.

Transgene position

Transgene integration into the plant genome is basically a random event, and it has been demonstrated that the position of the transgene in the genome can have a marked effect on transgene expression levels. Thus, independent transformants with single copies of the transgene can exhibit large differences in the level of transgene expression. This positional effect on transgene expression remains a problem in plant transformation, despite efforts to find ways of targeting gene integration. Unfortunately, higher plants do not possess an endogenous system enabling homologous recombination (genetic recombination involving the exchange of homologous loci) at a high efficiency (but see Chapter 5) that allows transgenes to be targeted to a particular region of the genome. However, alternatives utilising bacteriophage or yeast site-specific recombinases have been used with some success.

One approach taken to assuage position effects is the inclusion of matrix-attachment regions (MARs). These AT-rich sequences are thought to be involved in maintaining the chromatin in an 'open' structure allowing for gene expression. The inclusion of MARs flanking transgenes has been shown to provide position-independent expression in animal systems, and some, albeit

limited, success has been achieved with their use in plants. MARs may also help to stabilise transgene expression.

Transgene features

It is well documented that heterologous genes, particularly those from non-plant species, tend to express poorly in plants even when driven by a strong promoter. This can be due to a variety of factors, many of which are associated with the 'structure' of the transgene.

Genes from different organisms tend to have different G+C contents, with bacterial genes having a particularly low G+C content. It has been found that a high A+T content in transgenes interferes with mRNA processing in plants, leading to little or no expression of the transgene. Differences in A+T content between the transgene and the isochores (long stretches of DNA with a homogeneous base composition) into which it is integrated may also contribute to the transgene being recognised as 'foreign'. A high A+T content also often results in the presence of sequences (AUUUA) that can destabilise mRNA and may also form potential plant polyadenylation signals. So-called 'cryptic' introns may also be found. These can be efficiently mis-spliced from transgenes (i.e. part of the coding sequence is recognised by the plant as an intron and spliced out) by the plants mRNA processing machinery, effectively resulting in a deletion mutation of the transgene.

Even if the transgene is efficiently transcribed it may not be translated efficiently, probably due to the architecture of the translation initiation region and the presence of codons that are used infrequently by plants. Consensus sequences flanking the ATG initiation codon are known to differ between plants and other species. The inclusion of plant-specific sequences upstream of the translation initiation codon is known to improve translatability. It has also been recently demonstrated that modification (by the insertion, after the translation initiation codon, of codons that are found in a number of highly expressed plant genes) of the region downstream of the translation initiation codon may also improve transgene translatability. The inclusion of the tobacco mosaic virus Ω leader between the promoter and transgene has also been shown to improve translatability.

Mutagenesis can be used to alter the codon usage of a particular sequence; although if the sequence is large and requires a large number of changes, this may prove difficult and/or expensive. Some examples of transgene modification to optimise expression are given in Box 4.3.

Clean gene technology

Since the introduction of transgenic crop plants into agriculture, public concern about the safety of such technology, particularly about the safety of

BOX 4.3 Examples of optimisation of vector components

Modification of the *Bacillus thuringiensis cry1A(b)* gene

The *cry1A(b)* gene encodes a protein that can be used to control insect pests. The wild-type gene was found to be expressed at relatively low levels in transgenic plants. The effect of modifying the sequence of the gene on expression was investigated. Codon usage and overall G+C content was altered and several potential plant polyadenylation signals and ATTTA sequences were removed. These changes were seen to result in large increases (up to 100-fold) in the levels of the *cry1A(b)* gene product.

	Wild-type gene	Modified gene
No. altered bases (%)	0	390 (21%)
No. of modified codons (%)	0	356 (60%)
% G + C content	37	49
No. of potential polyadenylation sites	18	1
No. of ATTTA sequences	13	0

Data taken with permission from Perlak, F. J., *et al.* (1991).

Modification of the *gfp* gene

The *Aequorea victoria* GFP gene is used as a reporter gene in plant transformation experiments. In order to improve its use in plants, and in particular in *Arabidopsis*, it has been extensively modified. A cryptic intron that is efficiently mis-spliced in some plant species was removed by altering codon usage (altering the codon usage has also been demonstrated to increase the expression level). Translatability of the gene was improved by introducing an AACA sequence upstream of the translation initiation codon. In order to prevent build up of green fluorescent protein (GFP) in the nucleoplasm (where it may prove to be mildly toxic) a 5' terminal signal peptide and a 3' endoplasmic reticulum (ER) retention signal (HDEL, His-Asp-Glu-Leu) were introduced to target the GFP to the ER.

(Data taken with permission from Haselhoff, J., *et al.* (1997).)

antibiotic and herbicide resistance genes, has increased to the point where governments have been forced to take action (see Chapter 12). Whatever the scientific merits of the argument, it is a fact that the perceived dangers of transgenic crops are preventing their widespread adoption, particularly in Europe. Even in countries where the growth of such crops is not seen as a major issue (like the USA), the effect is still felt as export markets for foodstuffs derived from genetically engineered crops are affected.

There has therefore been a move to so-called 'clean gene' technologies where the selectable marker is no longer present in the field-grown crop, or, as described earlier, a more acceptable marker gene is used.

In order to develop crops that contain no selectable marker genes, several approaches are available:

1. Theoretically, the use of selectable marker genes could be avoided altogether by simply inserting the transgene of interest. Molecular biology techniques, such as the polymerase chain reaction (PCR), could be used to screen populations of putative transformed plants for those containing the transgene. In practice, this approach would simply be too expensive and labour intensive in most situations.
2. A second approach is to introduce the 'gene of interest' and the selectable marker genes on separate (unlinked) T-DNA molecules. Subsequent segregation analysis of the offspring can be used to identify plants that contain the gene of interest but not the selectable marker. Various strategies to deliver separate T-DNAs have been developed, the most effective appearing to be those in which one binary vector contains several T-DNA regions which are integrated at unlinked sites in the plant genome.
3. An alternative approach is to excise the selectable marker from the plant genome. This can be done by utilising site-specific recombinase systems. Several recombinase systems (phage P1 Cre-*lox*, *Saccharomyces cerevisiae* FLP-*frt* and the R-RS system of *Zygosaccharomyces rouxii*) are available. Vectors in which the selectable marker gene is placed between recognition sites (*lox*, *frt* or RS) for a site-specific recombinase can be used to transform plants. Subsequently, the recombinase (Cre, FLP, R) is expressed (either by transient expression, crossing of plants with recombinase-expressing plants or by using an inducible expression system), resulting in the selectable marker gene being excised from the plant genome.
4. Recently, vectors, in which the selectable marker gene was flanked by sequences that apparently considerably increase the frequency of intrachromosomal recombination leading to excision of the marker gene, have been demonstrated.

Summary

This chapter has looked at the construction of binary vectors for use in *Agrobacterium*-mediated transformation strategies. These binary vectors are, in effect, highly specialised derivatives of the *Agrobacterium tumefaciens* Ti plasmid looked at in Chapter 3. Key to their successful use in plant transformation protocols are the modifications made, resulting in removal of the oncogenes (thereby preventing abnormal plant development) and a general improvement in their ease of use. *Agrobacterium*-mediated transformation with binary vectors has been successfully applied to a wide range of crop plants, but improvements to many aspects of vector design are still being made. Some of these improvements are based on purely scientific grounds, to prevent or reduce gene silencing for instance, whilst others are based on widening the applicability of binary vectors (for example, alternative selectable markers). Other modifications are being made in response, at least in part, to concerns (public and scientific) about the safety of such vectors. These concerns, mainly over selectable marker genes, are particularly evident in

Europe. In the UK, the Royal Society's 1998 report, 'Genetically modified plants for food use', questioned the future use of antibiotic resistance genes and recommended that if they were used they should be removed at an early stage of development of the GM plant. The Advisory Committee on Releases to the Environment (ACRE), in its 'Guidance on Principles of Best Practice in the Design of Genetically Modified Plants', similarly noted the desirability of reducing the amount of 'extraneous' DNA (such as marker genes) in transgenic plants. 'Clean-gene' technologies and alternative selectable markers are therefore becoming increasingly attractive.

Binary vectors and *Agrobacterium*-mediated transformation have contributed to the considerable successes that plant biotechnology has already achieved, and will no doubt continue to contribute well into the future. In the chapters that follow we will look at specific examples of plant biotechnology in practice, including ones that relied on the use of *Agrobacterium*-mediated transformation and binary vectors.

Further reading

Vector construction

- Croy, R. R. D. (1993). Plant selectable genes, reporter genes and promoters. In *Plant molecular biology labfax* (ed. R. R. D. Croy), pp. 149–82. BIOS, Oxford.
- Guerineau, F. and Mullineaux, P. (1993). Plant transformation and expression vectors. In *Plant molecular biology labfax* (ed. R. R. D. Croy), pp. 121–47. BIOS, Oxford.
- Hellens, R., Mullineaux, P. and Klee, H. (2000). A guide to *Agrobacterium* binary Ti vectors. *Trends in Plant Science*, 5, 446–51.
- Hull, G. A. and Devic, M. (1995). The β -glucuronidase (*gus*) reporter gene system. In *Methods in molecular biology* (ed. H. Jones), pp. 125–41. Humana Press, Totowa.
- Jefferson, R. A., Kavanagh, T. A. and Bevan, M. A. (1987). GUS fusions: β -glucuronidase as a versatile gene fusion marker in higher plants. *EMBO Journal*, 6, 3901–7.
- pGreen and pBIN19—web-link 4.1: http://www.pgreen.ac.uk/a_hom.htm

Selectable markers

- Haldrup, A., Noerremark, M. and Okkels, F. T. (2001). Plant selection principle based on xylose isomerase. *In vitro Cellular and Developmental Biology—Plant*, 37, 114–19.
- Haselhoff, J., Siemering, K. R., Prasher, D. C. and Hodge, S. (1997). Removal of a cryptic intron and subcellular localisation of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. *Proceedings of the National Academy of Sciences USA*, 94, 2122–7.
- Joersbo, M. (2001). Advances in the selection of transgenic plants using non-antibiotic marker genes. *Physiologia Plantarum*, 111, 269–71.
- Kaepler, H. F., Carlson, A. R. and Menon, G. K. (2001). Routine utilization of green fluorescent protein as a visual selectable marker for cereal transformation. *In vitro Cellular and Developmental Biology—Plant*, 37, 120–6.
- Penna, S., Sagi, L. and Swennen, R. (2002). Positive selectable marker genes for routine plant transformation. *In vitro Cellular and Developmental Biology—Plant*, 38, 125–8.
- Reed, J., Privalle, L., Powell, M. L., Meghji, M., Dawson, J., Dunder, E., Suttie, J., Wenck, A., Launis, K., Kramer, C., Chang, Y.-F., Hansen, G. and Wright, M. (2001).

Phosphomannose isomerase: an efficient selectable marker for plant transformation. *In vitro Cellular and Developmental Biology—Plant*, 37, 127–32.

Inducible promoters

Kong, H.-G., Fang, Y. and Singh, K. B. (1999). A glucocorticoid-inducible transcription system causes severe growth defects in *Arabidopsis* and induces defence-related genes. *Plant Journal*, 20, 127–33.

Martinez, A., Sparks, C., Hart, C. A., Thompson, J. and Jepson, I. (1999). Ecdysone agonist inducible transcription in transgenic tobacco plants. *Plant Journal*, 19, 97–106.

Zuo, J. and Chua, N.-H. (2000). Chemical-inducible systems for regulated expression of plant genes. *Current Opinion in Biotechnology*, 11, 146–51.

Transgene integration and expression

Fang, R.-X., Nagy, F., Sivasubramaniam, S. and Chua, N.-H. (1989). Multiple *cis* regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. *Plant Cell*, 1, 141–50.

Gallie, D. R. (1998). Controlling gene expression in transgenics. *Current Opinion in Plant Biology*, 1, 166–72.

Kumar, S. and Fladung, M. (2001). Controlling transgene integration in plants. *Trends in Plant Science*, 6, 155–9.

Miki, B. (2002). Transgene expression and control. *In vitro Cellular and Developmental Biology—Plant*, 38, 139–45.

Perlak, F. J., Fuchs, R. L., Dean, D. A., McPherson, S. L. and Fischhoff, D. A. (1991). Modification of the coding sequence enhances plant expression of insect control protein genes. *Proceedings of the National Academy of Sciences USA*, 88, 3324–8.

Clean-gene technology

Ebinuma, H. and Komamoto, A. (2001). MAT (multi-auto-transformation) vector system. The oncogenes of *Agrobacterium* as positive markers for regeneration and selection of marker free transgenic plants. *In vitro Cellular and Developmental Biology—Plant*, 37, 103–13.

McCormac, A. C., Elliott, M. C. and Chen, D. F. (1999). pBECKS2000: a novel plasmid series for the facile creation of complex binary vectors which incorporates ‘clean-gene’ facilities. *Molecular and General Genetics*, 261, 226–35.

McCormac, A. C., Fowler, M. R., Chen, D. F. and Elliott, M. C. (2001). Efficient co-transformation of *Nicotiana tabacum* by two independent T-DNAs, the effect of T-DNA size and implications for genetic separation. *Transgenic Research*, 10, 143–55.

Puchta, H. (2000). Removing selectable marker genes: taking the short cut. *Trends in Plant Science*, 5, 273–4.

Best practice in vector design (ACRE report)—web-link 4.2:

<http://www.defra.gov.uk/environment/acre/bestprac/consult/guidance/bp/index.htm>

The Royal Society—web-link 4.3: <http://www.royalsoc.ac.uk/>

5

The genetic manipulation of herbicide resistance

Introduction

The preceding four chapters have given an overview of the tools of plant biotechnology that can be used for genetic manipulation. The remaining section of the book will look in detail at the major targets for the genetic manipulation of crops and how they are being approached or have been achieved. Thus, each chapter will describe the scientific strategies that could be used to attain a particular goal for crop improvement—why certain genes might be useful, how they might be transformed into plants and how their expression might be regulated. In addition, some of the wider issues surrounding GM crops will be explored—why have certain targets received so much commercial attention, which GM strategies have actually been successful in the field and are the concerns about the safety of GM crops justified?

The first chapter in this sequence of different GM applications deals with the genetic manipulation of herbicide-resistant plants. Herbicide resistance was one of the first GM traits to be tested in the field, and subsequently for commercial production. Table 5.1 shows that the greatest number of trial sites set up during the period 1987–1997 were to test crops carrying this trait, and that herbicide-resistant crops form more than three-quarters of the total released GM crops in terms of area of commercial planting. (Indeed, note that of all the traits described in this book, only herbicide and insect resistance (see Chapter 6) are widely grown, as yet.) One of the aims of this chapter is to explain why there has been such rapid development of this particular trait, and why crops genetically modified to be herbicide resistant remain the most widely grown of all GM crops.

There are scientific and commercial reasons why herbicide-resistant crops have been developed so rapidly compared to other GM traits. The scientific reasons include:

1. Considerable information was already available about the modes of action of certain herbicides, and about the affected biochemical pathways.

2. Biological resources were available to gain an understanding of the mechanisms of resistance, and provide genes for genetic manipulation, including:
- resistant bacteria via the laboratory or from the wild;
 - tolerant plants selected in tissue culture;
 - field-selected resistant crops and weeds.
3. Single-gene mechanisms to obtain resistance were relatively simple to devise. (This is not the case with some of the traits discussed in subsequent chapters.)

The major commercial impetus for their development flowed from the advantage to the agrochemical/seed industries of producing crops resistant to specific herbicides, particularly those manufactured/owned by the same company. The engineering of a crop to be resistant to a particular herbicide allows it to be treated at the optimal times for the reduction of weeds, hence greatly extending the applications and effectiveness of that herbicide. The role of the agrochemical industry in driving forward this technology (and GM crop science in general) will be returned to in this and subsequent chapters.

Table 5.1 Most frequent traits in transgenic crop field trials and commercial plantings
(Data from James, 1997, 2001: web-link 1)

Trait	Number of trial sites (%) in the USA (1987–1997)	Area of commercial planting (2000)
Herbicide tolerance	30	32.7 M hectares (74%)
Insect resistance	24	8.3 M hectares (19%)
Herbicide + insect resistance		3.1 M hectares (7%)
Product quality	21	—
Viral resistance	10	—
Fungal resistance	4	—
Agronomic properties	4	—
Others	7	—

The use of herbicides in modern agriculture

Weeds have a significant effect on the yield and quality of crops, as a result of competition for light and nutrients, contamination of the harvested crop and because weed populations harbour pests and diseases. Thus weeds are one of the three classes of biotic stress that have a major impact on the proportion

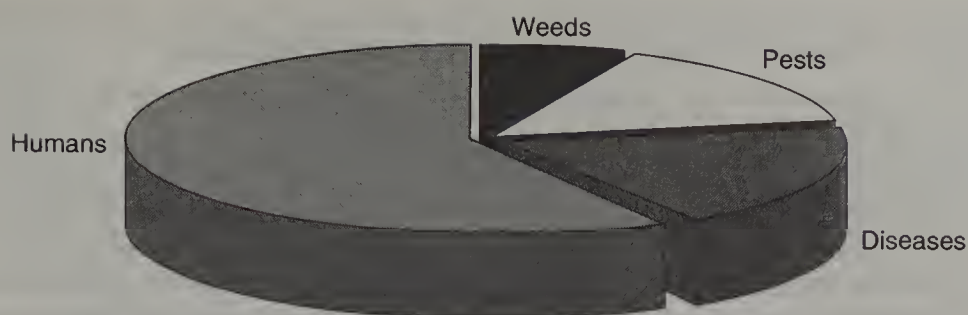


Figure 5.1 The relative proportions of total world crop production lost to weeds, pests and diseases. The figure shows that of the total potential yield, only 63% is available for human consumption. The remainder is lost in more or less equal proportions to weeds, pests and diseases.

of world crop yield available for human consumption. Figure 5.1 indicates that the reduction in world crops caused by weeds, pests and diseases amounts to over one-third of the potential yield, with a roughly equal contribution made by each. Modern agriculture has developed a range of effective herbicides (weedkillers) to tackle the effects of weeds on crop yield. Some of the most useful of these are broad-spectrum herbicides because they are active against a wide range of weeds. However, these can only be used at times when the crop is not itself vulnerable to herbicide action. The development of herbicide-resistant crops therefore offers the opportunity to spray crops at the most effective time to kill weed species without damaging the crop plants. However, it should be noted that some crop plants are naturally resistant to certain herbicides, and that resistant strains may appear through the normal processes of mutation and natural selection. Thus, the concept of herbicide-resistant crops is neither novel nor unique to GM technology.

What types of compounds are herbicides?

By definition, herbicides are much more toxic to plants than to animals. Therefore, it is not surprising that they generally affect ‘plant-specific’ biological processes. Consider for a moment the type of biochemical pathways that are likely to occur in plants but not in animals. Although one’s first thought is likely to be of photosynthesis, many other processes are potential targets. Remember that plants are autotrophic, i.e. they can synthesise all their macromolecular components *de novo*. For example, the compounds that are essential in the human diet (vitamins, essential amino acids, etc.) are synthesised by plants, and these biosynthetic pathways are therefore possible targets for herbicides. Many of these pathways are located in the chloroplast (Box 5.1), and this has implications for the targeting of many of the transgenic proteins designed to enhance herbicide resistance.

The herbicidal activity of many herbicides has been found to result from the specific inhibition of a single enzyme/protein. Table 5.2 shows a classification

BOX 5.1 Plastids as major biosynthetic factories of plant cells

The plastids are major contributors to the biosynthetic activity in plant cells. The chloroplast is the sole site of the photoreduction of carbon dioxide to phosphoglyceric acid, the photo-oxidation of water and photophosphorylation. In addition, all an animal's essential dietary amino acids are synthesised in chloroplasts. The vast majority of lipid biosynthesis occurs in this organelle, as does the synthesis of purines and pyrimidines. The plastids also play a major role in the assimilation of nitrogen, reducing nitrite (produced by reduction of nitrate in the cytosol) to ammonia, and in the assimilation of sulphur, reducing sulphate to sulphide. The chlorophyll pigments of the chloroplast are also produced *in situ*, as are the carotenoid pigments of chromoplasts. Starch biosynthesis is also an important pathway in chloroplasts and is the predominant function of amyloplasts.

The chloroplast is also responsible for a proportion of plant cell protein synthesis. Note that the chloroplast genome encodes all the components required for the plastid translational machinery, as well as for a number of the chloroplast proteins (Chapter 1).

of the major chemical groups of herbicides according to their modes of action. It can be seen that herbicides belong to a wide range of different chemical families, with about 15 broad classes of mode of activity. It is also worth noting that most herbicides only have one mode of action, and that certain enzymes seem to be relatively vulnerable to herbicide activity. For example, five chemical families of herbicide target the enzyme acetolactate synthase (ALS). ALS catalyses the first reaction in the branched-chain amino acid biosynthetic pathways (see Figure 5.5), and Table 5.3 shows two common classes of herbicide with widely differing chemical structures (sulphonylureas and imidazolinones) that target this enzyme. As will be seen, knowledge of the herbicide's site of action is essential for some strategies aimed at engineering resistance, but less important for others.

The wide range of chemical families shown in Table 5.2 means that herbicides differ greatly in other properties besides their mode of action. For example, they can also be classified according to their:

- site of uptake into the plant (root vs. shoot);
- degree of translocation within the plant (systemic vs. contact);
- time of application (preplanting, pre-emergence, postemergence or preharvesting).

The widely varied chemical properties of these compounds also means they will differ greatly with regard to toxicity, environmental persistence and biodegradability. The environmental impact of herbicide-resistant crops will be considered at the end of this chapter. However, this is a useful place to make the point that an assessment of the risks and benefits of growing a particular

herbicide-tolerant crop must take account of the properties of that particular herbicide.

Table 5.3 shows the chemical structure of a few common herbicides that affect well-characterised biochemical targets. This again emphasises the point that herbicides are a heterogeneous group of compounds with differing modes of action. In consequence, most transgenic strategies for resistance to a particular herbicide have to be specifically designed for that class of herbicide. Rather than describe in detail all the herbicides against which resistance has been engineered, this chapter will adopt a case-study approach in order to highlight general approaches and strategies.

Strategies for engineering herbicide resistance

The engineering of herbicide resistance demonstrates the way in which quite different strategies can be used to achieve the same objective. Mullineaux (1992) identifies four distinct strategies for engineering herbicide resistance. These are:

1. *Overexpression of the target protein.* This strategy effectively involves titrating the herbicide out by overproduction of the target protein. For example, if the herbicide is a specific inhibitor of one particular enzyme, production of sufficient excess enzyme will partially overcome the inhibition. Overexpression can be achieved by the integration of multiple copies of the gene and/or the use of a strong promoter plus translational enhancer to drive expression of the gene.
2. *Mutation of the target protein.* The logic behind this approach is to find a modified target protein that substitutes functionally for the native protein and which is resistant to inhibition by the herbicide, and to incorporate the resistant target protein gene into the plant genome. Several sources of resistant proteins can be exploited. Note that both the overexpression and mutated target protein strategies require knowledge of the mode of action of the herbicide.
3. *Detoxification of the herbicide, using a single gene from a foreign source.* Detoxification is a means of converting the herbicide to a less toxic form and/or removing it from the system. This strategy can be contrasted with the previous two because it does not require a detailed knowledge of the site of action. Table 5.4 shows several examples of specific detoxification reactions for common herbicides.
4. *Enhanced plant detoxification.* The aim here is to improve the natural plant defences against toxic compounds. This requires detailed information about endogenous plant detoxification pathways and the mechanisms by which compounds are recognised and targeted for detoxification by the plant.

Table 5.2 Classification of herbicides according to their mode of action (Data from the Herbicide Resistance Action Committee (HRAC) and the Weed Science Society of America)

HRAC group	Mode of action	Chemical family	Example
A	Inhibition of acetyl-CoA carboxylase (ACCase)	Aryloxyphenoxy-propionates ‘FOPs’ Cyclohexanediones ‘DIMs’	
B	Inhibition of acetolactate synthase (ALS) (acetohydroxyacid synthase AHAS)	Sulfonylureas Imidazolinones Triazolopyrimidines Pyrimidinyl(thio)benzoate Sulfonylaminocarbonyl-triazolinones	Chlorsulfuron Imazapyr
C1	Inhibition of photosynthesis at photosystem II	Triazines Triazinones Triazolinone Uracils Pyridazinones Phenylcarbamates	Atrazine
C2	Inhibition of photosynthesis at photosystem II	Ureas Amides	
C3	Inhibition of photosynthesis at photosystem II	Nitriles Benzothiadiazinone Phenylpyridazines	Bromoxynil
D	Photosystem-I-electron diversion	Bipyridyliums	Paraquat
E	Inhibition of protoporphyrinogen oxidase (PPO)	Diphenylethers Phenylpyrazoles N-phenylphthalimides Thiadiazoles Oxadiazoles Triazolinones Oxazolidinediones Pyrimidindiones Others	
F	Bleaching: Inhibition of carotenoid biosynthesis at the phytoene desaturase step (PDS)	Pyridazinones Pyridinecarboxamides Others	
F2	Bleaching: Inhibition of 4-hydroxyphenyl-pyruvate-dioxygenase (4-HPPD)	Triketones Isoxazoles Pyrazoles Others	

Table 5.2 *Continued*

HRAC group	Mode of action	Chemical family	Example
F3	Bleaching: Inhibition of carotenoid biosynthesis (unknown target)	Triazoles Isoxazolidinones Ureas Diphenylethers	
G	Inhibition of EPSP synthase	Glycines	Glyphosate
H	Inhibition of glutamine synthetase	Phosphinic acids	Glufosinate-ammonium Bialaphos
I	Inhibition of DHP (dihydropteroate) synthase	Carbamates	Asulam
K1	Microtubule assembly inhibition	Dinitroanilines Phosphoramidates Pyridines Benzamides Benzenedicarboxylic acids	Oryzalin
K2	Inhibition of mitosis/microtubule organisation	Carbamates	
K3	Inhibition of cell division	Chloroacetamides Acetamides Oxyacetamides Tetrazolinones Others	
L	Inhibition of cell wall (cellulose) synthesis	Nitriles Benzamides Triazolocarboxamides	
M	Uncoupling (membrane disruption)	Dinitrophenols	
N	Inhibition of lipid synthesis—not ACCase inhibition	Thiocarbamates Phosphorodithioates Benzofuranes Chlorocarbonic acids	TCA
O	Action like indole-acetic acid (synthetic auxins)	Phenoxy-carboxylic acids Benzoic acids Pyridine carboxylic acids Quinoline carboxylic acids Others	2,4-D Dicamba Picloram
P	Inhibition of auxin transport	Phthalamates Semicarbazones	
Z	Unknown	Arylamino-propionic acids Pyrazolium Organoarsenicals	

Table 5.3 The structures of common herbicides affecting biochemical targets

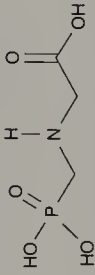
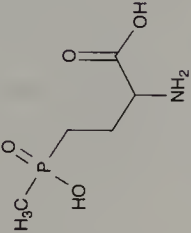
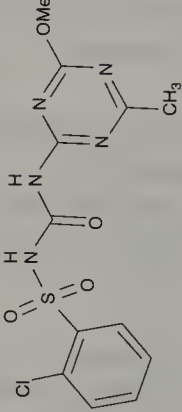
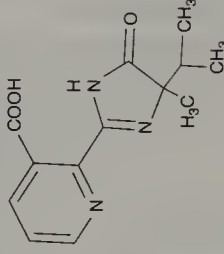
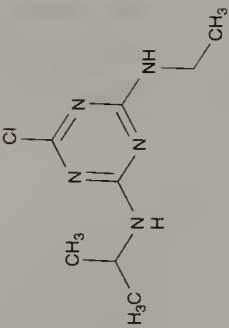
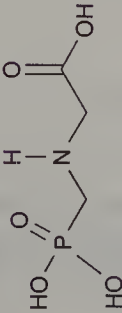
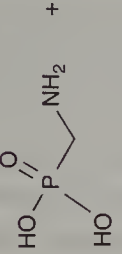
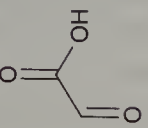
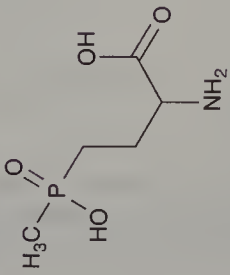
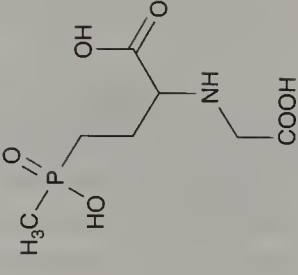
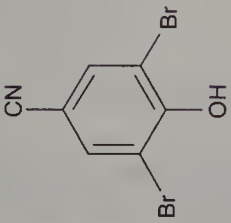
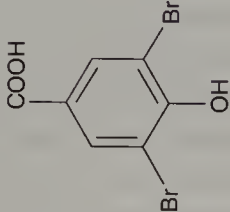
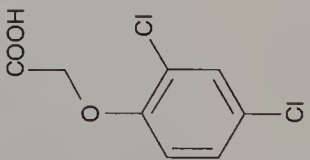
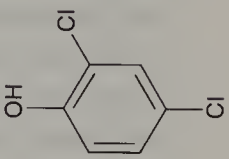
Class of herbicide	Compound/herbicide	Chemical formula	Inhibited pathway	Target protein
Glycine	Glyphosate (Roundup)		Aromatic amino acids	5-Enolpyruvyl shikimate 3-phosphate synthase
Phosphinic acid	Phosphinothricin (Basta)		Nitrogen assimilation	Glutamine synthase
Sulphonylurea	Chlorsulphuron (Glean)		Branched-chain amino acids	Acetolactate synthase
Imidazolinone	Imazathapyr		Branched-chain amino acids	Acetolactate synthase
S-triazine	Atrazine		Photosynthesis	33-kDa protein

Table 5.4 Herbicide detoxification reactions

Compound/herbicide	Enzyme	Metabolite	Source organism
	Oxidoreductase	 	<i>Ochrabactrum anthropi</i>
	Acetyltransferase		<i>Streptomyces hygroscopicus</i>
	Nitrilase		<i>Klebsiella ozaenae</i>
	Monoxygenase		<i>Alcaligenes eutrophus</i>

CASE STUDY 1 **Glyphosate resistance**

Glyphosate is a broad-spectrum herbicide that is reputedly effective against 76 of the world's worst 78 weeds, and is marketed as 'Roundup' by the American chemical company Monsanto. It is a simple glycine derivative (see Table 5.2) that acts as a competitive inhibitor of the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS). Figure 5.2 shows the similarity in structure between glyphosate and one of the substrates of the enzyme, phosphoenolpyruvate (PEP). Glyphosate binds more tightly to the EPSPS-shikimate-3-phosphate complex than does PEP—its dissociation rate from the complex is 2300 times slower than PEP. Consequently, EPSPS is effectively inactivated once glyphosate binds to the enzyme-substrate complex. EPSPS is a key enzyme in the biosynthetic pathways of the aromatic amino acids phenylalanine, tyrosine and tryptophan (Box 5.2). Thus, the herbicidal activity of glyphosate results from its inhibition of the biosynthesis of aromatic amino acids and other products of the shikimate pathway.

From this knowledge of the mode of action of glyphosate, is it possible to predict what effects the herbicide is likely to have on the growth of a treated plant? The first deduction is that protein synthesis will be blocked due to the insufficient supply of aromatic amino acids. The most immediate effects of this inhibition would be expected in regions of the plant involved in rapid growth and division, including the meristems. Certain specialised organs, such as the developing endosperm, also vigorously accumulate proteins. In addition, other pathways will be affected by the depletion of aromatic amino acids. The shikimate pathway

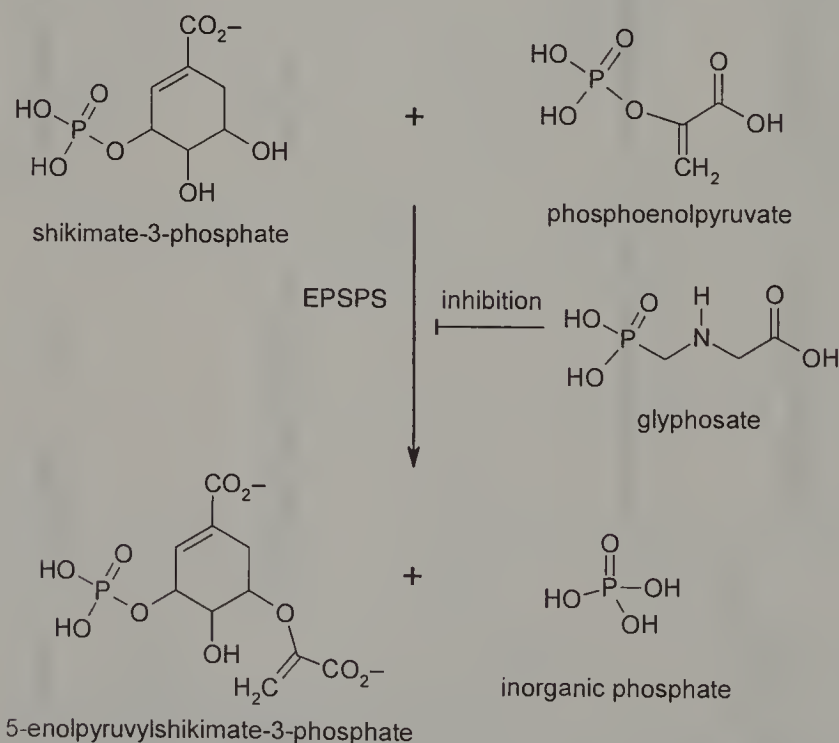
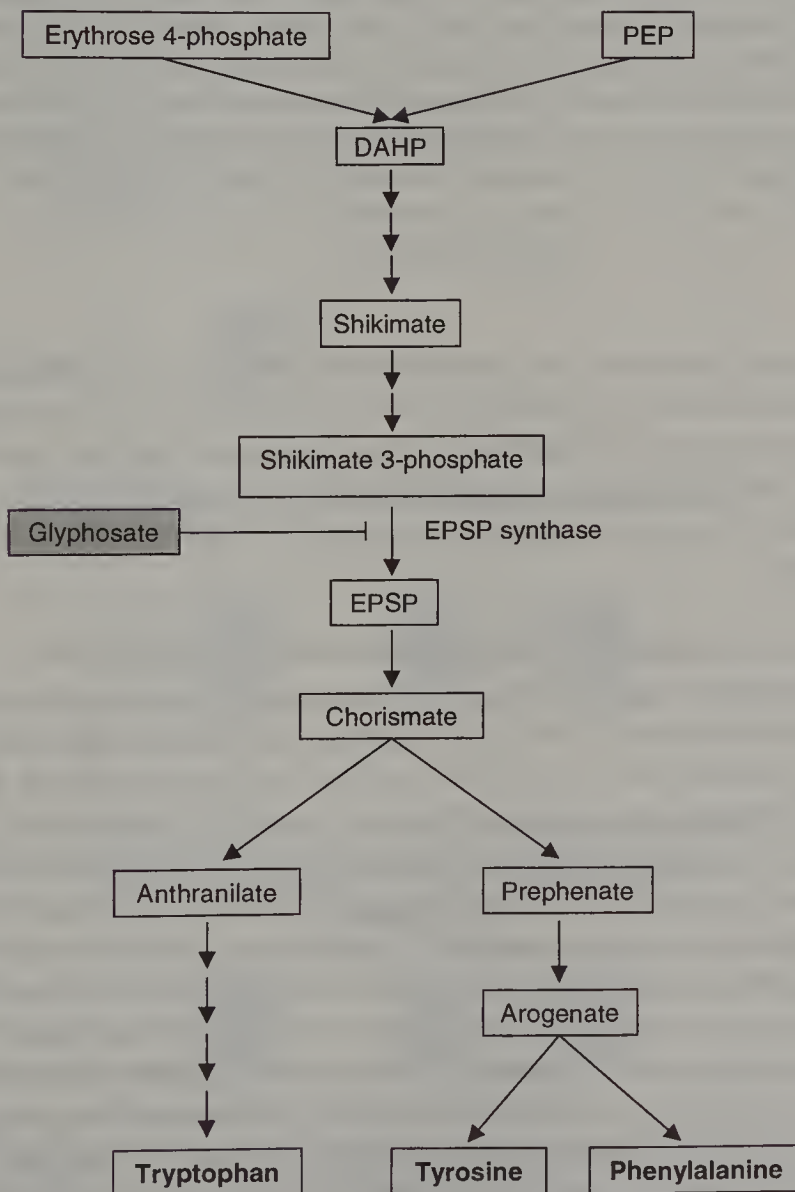


Figure 5.2 Glyphosate and the reaction catalysed by EPSPS. The biochemical reaction catalysed by EPSPS involves the addition of phosphoenolpyruvate to shikimate-3-phosphate. The structural similarity between glyphosate and PEP explains the competitive inhibition of EPSPS by the herbicide.

BOX 5.2 Aromatic amino acid biosynthesis pathways

The biosynthesis of the aromatic amino acids and related compounds shares a common biochemical pathway up to the formation of chorismate. The pathway to chorismate starts with the condensation of phosphoenolpyruvate (PEP) (from glycolysis) and erythrose 4-phosphate (from the pentose phosphate pathway) to form 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP). The next step involves a complex redox/cyclisation reaction producing 3-dehydroquinate. The next two reactions (to form 3-dehydro-shikimate and then shikimate) are catalysed by a bifunctional enzyme in plants. Shikimate kinase then



Biosynthetic pathways of the aromatic amino acids.

BOX 5.2 Continued

phosphorylates shikimate to produce shikimate 3-phosphate, one of the substrates of 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS). EPSPS adds the enolpyruvyl side chain to shikimate 3-phosphate to form EPSP. Finally, chorismate is formed by the elimination of phosphate from EPSP.

Chorismate is the precursor of the phenolic and indole rings of the aromatic amino acids, and of other aromatic compounds. Chorismate mutase is the committing enzyme for phenylalanine and tyrosine synthesis, forming prephenate. Prephenate aminotransferase utilises glutamate or aspartate as amino donors to form aroenate, the immediate precursor of phenylalanine (via aroenate dehydratase) and tyrosine (via aroenate dehydrogenase). These amino acids are themselves the precursors of a wide range of secondary products, including lignins, flavonoids, hydroxycinnamic acids and alkaloids.

Alternatively, chorismate may be converted to anthranilate by anthranilate synthase, an enzyme complex with two catalytic subunits—the α -subunit catalysing the amination of chorismate and the removal of the enolpyruvyl side chain, whilst the β -subunit has glutamine aminotransferase activity. Four subsequent steps are involved in the biosynthesis of tryptophan. Tryptophan itself is the precursor of several important secondary products, including IAA, indole alkaloids, indole glucosinolates, phytoalexins and acridone alkaloids.

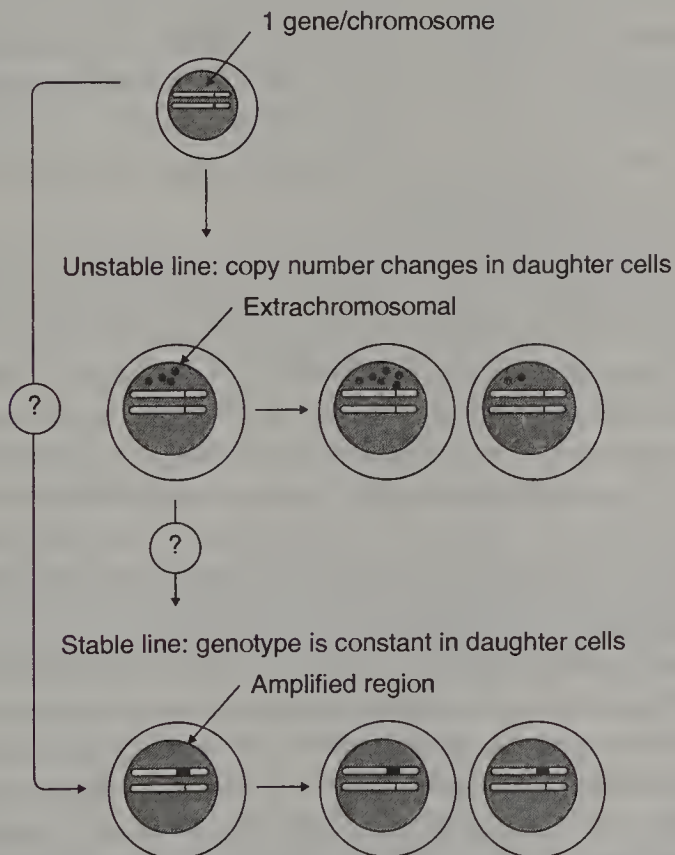
supplies aromatic precursors for a range of phenolic compounds, including lignins, alkaloids and flavonoids (Box 5.2). Indeed, 20% of the carbon fixed by plants flows through this pathway, primarily for lignin biosynthesis. Indole compounds other than tryptophan are produced by the same pathway, so auxin (indole-acetic acid, IAA) biosynthesis will also be affected by this herbicide. It is therefore not surprising that glyphosate has such a profound effect on plants.

Strategy 1 for glyphosate resistance: overexpression of a plant EPSPS gene

Of the four strategies described above, three have actually been tested in the laboratory, two of which form the basis of the current commercial plantings of glyphosate-resistant crops. It is therefore instructive to compare the three strategies. One of the earliest approaches to engineering glyphosate resistance involved the overexpression of a plant EPSPS gene. This was facilitated by the isolation of petunia cDNA from glyphosate-resistant tissue cultures. The stepwise selection of petunia cells capable of growing in the presence of increasing amounts of glyphosate led to the isolation of cultures in which the levels of EPSPS enzyme were much higher than normal. This was found not to be due to increased expression of the EPSPS gene, but was rather the result of gene amplification, such that there were multiple (up to 20) copies of the EPSPS gene in an otherwise normal petunia genome (see Box 5.3). The EPSPS enzyme was not itself mutated—i.e. the resistance was simply due to the increased amount of enzyme. However, the high levels of EPSPS mRNA made it simpler to isolate the cDNA for this gene and use it for re-introduction into plants.

BOX 5.3 Gene amplification

In prokaryotes, the isolation of a strain resistant to a particular toxic compound would usually result from the selective advantage derived from mutations in the gene coding for the affected protein itself, or in related genes involved in transport, detoxification or gene regulation. In contrast, the selection of eukaryotic cells in culture resistant to a particular toxic compound often results from the amplification of the gene encoding the target enzyme, rather than from mutations in that gene. This process is particularly observed during a step-wise selection procedure, in which the level of toxin is increased gradually in each round of selection. One of the best characterised examples is that of methotrexate resistance in cultured mammalian cells. The anticancer drug methotrexate is an inhibitor of the enzyme dihydrofolate reductase (DHFR), which supplies single carbon units for, amongst other things, thymidine synthesis. It is therefore particularly important in cycling cells for the replication of DNA. A step-wise selection of methotrexate-resistant cells led to the predominance of cells in which the DHFR gene had been specifically amplified in tandem 40–400-fold. This indicates that gene duplication occurs more frequently than mutation in eukaryotes. The step-wise selection applies a quantitative selection pressure for repeated gene duplications, resulting in a process of ‘accelerated evolution’.



The *dhfr* gene can be amplified to give unstable copies that are extrachromosomal (double minutes) or stable (chromosomal). Extrachromosomal copies arise at early times. (with permission from Lewin Genes VII 2000 Oxford University Press)

The observation that the endogenous mechanism of resistance that had been selected for in the petunia cells was one of excess normal EPSPS (as a result of gene amplification) indicates that Strategy 1—overexpression of the target protein—should be feasible. The effect of overexpressing the EPSPS gene in the transgenic petunia was tested in one of the earliest experiments on the engineering of herbicide resistance. The EPSPS cDNA was fused to the CaMV 35S promoter and a *nos* terminator sequence in the vector pMON 546 (Figure 5.3) and transformed into petunia using *Agrobacterium*. The use of the plant gene enabled the researchers to avoid one of the major obstacles to transgene expression—protein targeting. As noted above, many of the potential target pathways of herbicides are located in the plastids. The targeting of proteins to the plastid is discussed in Box 5.4. Note that no further manipulation of the cDNA was required in order to target this protein to the plastid site of activity, since the plant EPSPS cDNA sequence contains its own transit peptide.

The result of this experiment was a 40-fold increase in EPSPS activity in the transgenic plants and a tolerance to glyphosate sprayed in the field at a dose 2–4 times higher than that required to kill wild-type plants.

Strategy 2 for glyphosate resistance: mutant EPSPS genes

Mutated EPSPS genes have been isolated from a number of glyphosate-resistant bacteria. It is instructive to compare two of the early experiments using glyphosate-resistant genes from bacteria. In one, a mutated *aroA* gene from *Salmonella typhimurium* was inserted between the promoter and terminator sequences of the *ocs* gene of the *Agrobacterium tumefaciens* Ti plasmid. Only a moderate increase in herbicide tolerance was obtained. With reference to Box 5.4, you should be able to explain why resistance was limited—the prokaryote gene did not have a plastid transit peptide sequence, so the resistant enzyme was not transported into the chloroplast.

The requirement for a functional plastid transit peptide was demonstrated by the construction of a hybrid EPSPS gene by fusion of the C-terminal end of a mutated *aroA* gene from *E. coli* to the N-terminal end of the petunia EPSPS cDNA sequence containing the transit peptide sequence (Figure 5.3). Expression of the chimeric enzyme increased glyphosate tolerance in transgenic tobacco from 0.01 mmol l^{-1} to 1.2 mmol l^{-1} glyphosate. This experiment was an important validation of the strategy, and demonstrated the feasibility of expressing prokaryotic genes incorporated into the plant nuclear genome, given appropriate promoter and termination signals. However, note that the expression of prokaryotic genes is often not optimal in transgenic plants and extensive modifications may be required to obtain high levels of expression (see Chapter 4).

These early experiments provided useful evidence for the feasibility of Strategy 2 (the resistant target protein strategy), but also revealed problems associated with the strategy. One is that a mutant enzyme with reduced affinity for a competitive inhibitor may also have a lower affinity (increased K_M) for the substrate. This proved to be the case with the *E. coli* and *S. typhimurium* genes. For this reason, glyphosate-resistant genes from other sources have been tested for their effectiveness in different plants. A gene from the herbicide-resistant *A. tumefaciens* strain CP4 encodes an EPSPS that is resistant to glyphosate, but retains a low K_M for phosphoenolpyruvate. This gene, in conjunction with an enhanced CaMV 35S promoter (see Chapter 4, Figure 4.2) and a chloroplast transit peptide sequence from *Arabidopsis* or petunia (Figure 5.3), is incorporated into the current range of Monsanto's major dicotyledo-

nous Roundup Ready crops (soybean, cotton and oilseed rape). On the other hand, Roundup Ready maize contains a construct optimised for monocotyledonous crops, with a resistant EPSPS gene from maize (isolated after mutagenesis and selection in tissue culture), fused to a rice promoter and maize chloroplast transit peptide sequence.

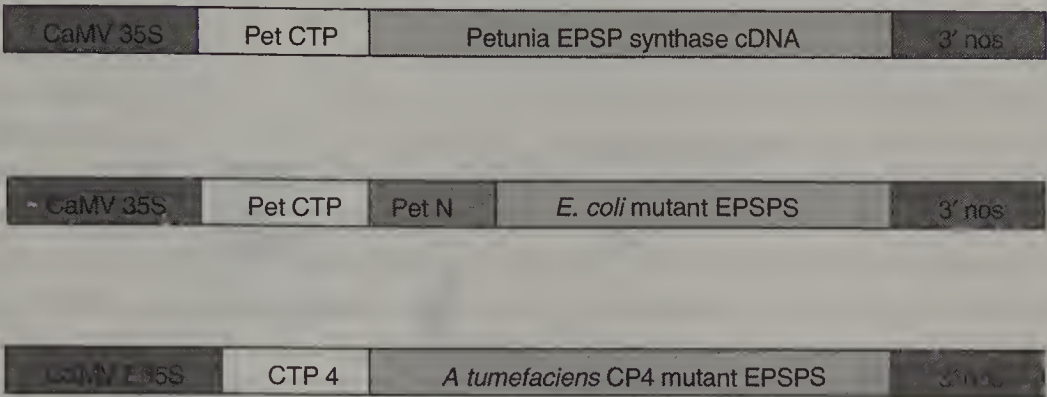


Figure 5.3 Constructs for engineering glyphosate resistance. Maps of some of the gene constructs that have been used to engineer glyphosate resistance are shown. The upper two constructs are early experimental designs that were used to demonstrate the feasibility of engineering herbicide resistance. The first construct directed overexpression of a wild-type petunia EPSPS gene, whilst the second shows a resistant *E. coli* gene fused to the N-terminal region of the petunia gene, including the chloroplast transit peptide. The third construct is one currently used to generate Roundup Ready crops such as soybean and cotton using a resistant EPSPS gene from *A. tumefaciens* with an enhanced CaMV 35S promoter (E355).

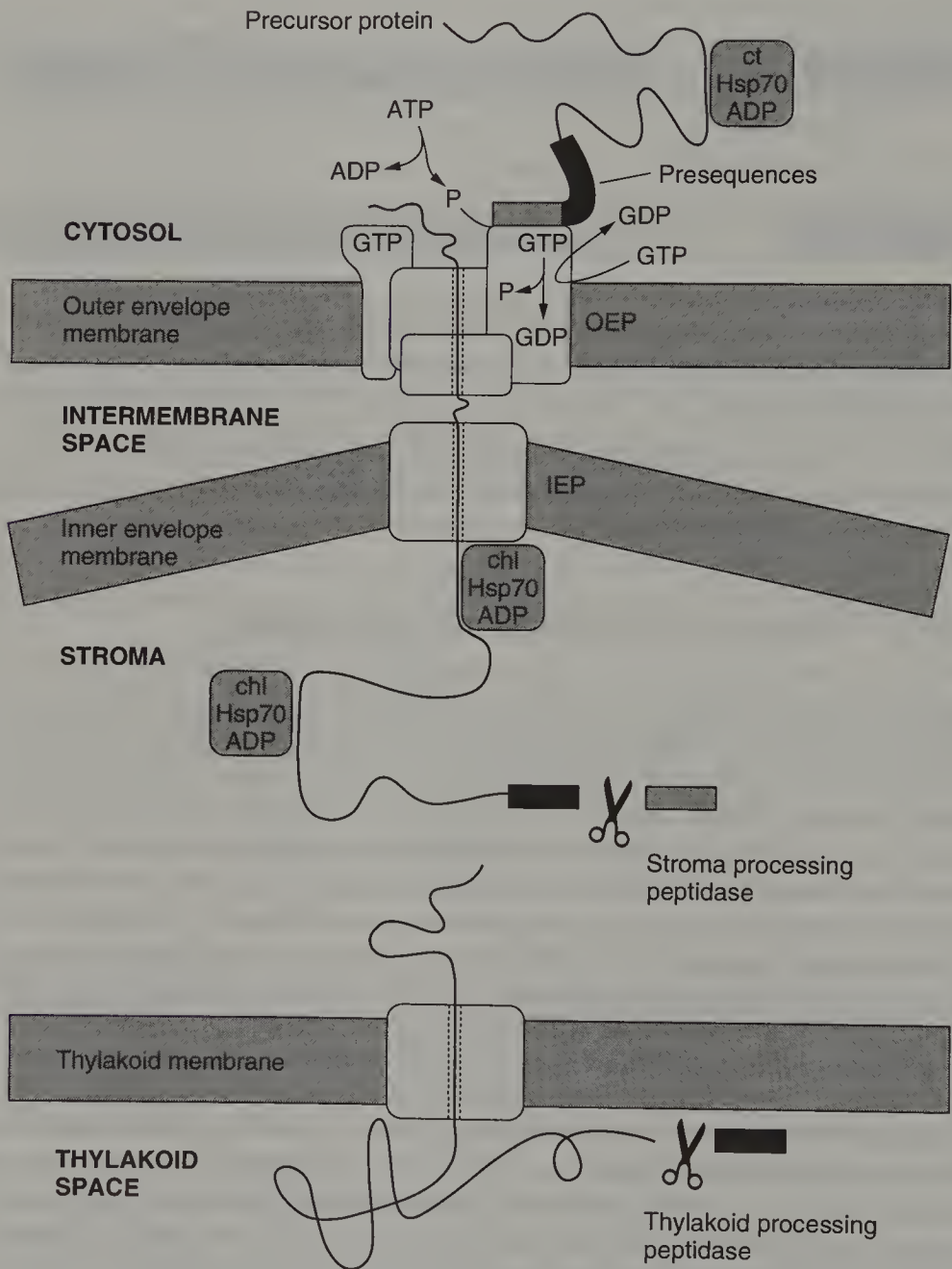
BOX 5.4 Chloroplast transit peptides and protein targeting

Some chloroplast proteins are encoded by the plastid genome and translated on 70S ribosomes in the organelle. However, most chloroplast proteins are encoded in the nuclear genome and translated on 80S ribosomes in the cytoplasm. There must therefore be mechanisms for the recognition and transport of proteins destined for the chloroplast. Note that the chloroplast envelope comprises an outer and inner membrane, and that the thylakoids comprise an additional internal membrane system. Thus there are three distinct compartments in the chloroplast: the intermembrane space of the envelope; the stroma; and the thylakoid lumen. Depending upon the precise destination, a chloroplast protein might be transported across three membrane systems. The means by which all plastid proteins are transported into the chloroplast is by recognition of a sequence of about 40–50 amino acids at the N-terminal end of the protein (the transit peptide). This peptide directs translocation into the stroma, where a specific peptidase removes the transit peptide. Note that the process is ‘post-translational’, i.e. the protein is transported after synthesis, unlike the co-translational transfer of membrane-bound and secreted proteins into the ER during synthesis on rough ER-bound ribosomes. The transport process into the stroma appears to involve a complex protein import apparatus that spans the inner and outer membranes.

Targeting into the thylakoids requires a bipartite transit peptide. Removal of the stromal targeting sequence exposes a second transit peptide that acts as a luminal targeting peptide. This directs the protein across the thylakoid membrane into the thylakoid lumen, where it is also removed by a specific protease.

BOX 5.4 *Continued*

Transit peptides will direct the transfer of a chimeric non-plastid protein into the chloroplast. Hybrid transgenes containing an N-terminal transit peptide can therefore be constructed so as to target the protein to a specific chloroplast compartment.



Protein transport into the chloroplast. (Taken with permission from Heldt H-W. Plant Biochemistry and Molecular Biology 1997 Oxford University Press)

Strategy 3 for glyphosate resistance: detoxification by heterologous genes

An alternative strategy has been developed for engineering glyphosate tolerance based upon a specific detoxification mechanism. In soil microorganisms, glyphosate can be degraded by cleavage of the C–N bond, catalysed by an oxidoreductase, to form aminomethylphosphonic acid (AMPA) and glyoxylate (Table 5.4). A gene encoding the enzyme glyphosate oxidase (GOX) has been isolated from a soil organism, *Ochrobactrum anthropi* strain LBAA, and modified by addition of a transit peptide. Transgenic crops such as oilseed rape transformed with this gene show very good glyphosate resistance in the field. However, this strategy is not generally used in isolation. Monsanto now employ a dual strategy for oilseed rape (canola), in which both the resistant *Agrobacterium* CP4 EPSPS gene and the GOX gene are expressed. In addition to enhanced glyphosate resistance, this approach avoids the accumulation of the herbicide in the resistant plant, because the glyphosate is broken down into relatively harmless products (glyoxylate is a normal plant metabolite, and AMPA can be converted to glycine). This highlights one significant difference between Strategy 2 and Strategy 3. Strategy 2 enables the plant to function in the presence of the herbicide and therefore the herbicide may accumulate to higher levels than those normally found in that crop. In contrast, the detoxification strategy should result either in the destruction of the herbicide, or in the accumulation of a conjugate less harmful than the original compound.

Pleiotropic effects of transgenes

It should be noted at this stage that the insertion of a transgene into a plant may result in unforeseen and perhaps undesirable effects. Roundup Ready crops have not been without problems of this type. One phenomenon encountered by Roundup Ready soybeans during hot weather has been splitting of the stems. It has been suggested that this occurs due to the 20% higher lignin content of these plants. A look back to the mode of action of glyphosate should indicate why the introduction of an additional EPSPS gene could result in increased lignin biosynthesis (see Box 5.2). It should be remembered that the plant EPSPS enzyme will still be functional under normal growth conditions in the absence of the herbicide, and that the expression of additional enzyme from the transgene may affect the balance of the relevant metabolic pathways. This lesson will be returned to in subsequent chapters, particularly Chapters 10 and 11 dealing with the manipulation of plant metabolic pathways.

CASE STUDY 2 **Phosphinothricin**

Glyphosate resistance is one of the most widespread commercial GM traits. The closest rival to glyphosate in terms of the number and acreage of resistant crops is the herbicide phosphinothricin (PPT) or glufosinate. Whilst both are broad-spectrum herbicides, glyphosate is particularly effective against grasses, whilst phosphinothricin is more effective against broad-leaved weeds and least effective against perennials and volunteer

cereals. (In this chapter, phosphinothricin will be used in preference to glufosinate to avoid confusion between glufosinate and glyphosate. Note also that glufosinate is usually applied as the ammonium salt, and is commonly called glufosinate ammonium.) Phosphinothricin is unusual amongst herbicides in being derived from a natural product. Bialaphos is a tripeptide of the form PPT–Ala–Ala (see Table 5.3), produced by certain *Streptomyces* species. It can be applied directly as a herbicide and has been marketed as such under various trade names, for example ‘Herbiace’ (Meiji Seika). Bialaphos is converted to the active form L-phosphinothricin by proteolytic removal of the alanine residues. Phosphinothricin was marketed by the German company Hoechst, under the trade name ‘Basta’. One point to be aware of in tracking the progress of a particular transgenic crop is the dynamic nature of the agrochemical and biotechnology sector—there have been a number of take-overs and mergers resulting in company name changes. Thus, Hoechst has since undergone a series of mergers such that the Basta brand name has been owned successively by AgrEvo and now Aventis. A different formulation of phosphinothricin is also marketed by Aventis under the brand name Liberty, and complements the LibertyLink lines of transgenic phosphinothricin-resistant crops produced by the same company, which are described below.

The herbicidal action of phosphinothricin is a result of its competitive inhibition of glutamine synthetase (GS). Figure 5.4 shows the similarity in size and charge between phosphinothricin and glutamate—one of the substrates of GS. The immediate effect of inhibiting GS is the accumulation of ammonia to toxic levels, which rapidly kills the plant cells. The disruption of glutamine synthesis also inhibits photosynthesis, and it is the combined effects of ammonium toxicity and inhibition of photosynthesis that account for the herbicidal activity of phosphinothricin. Uptake of phosphinothricin is through the leaf, the speed of which is dependent on many factors, including plant species, stage of growth, air humidity, temperature and rate of application. Translocation within the plant is limited (unlike glyphosate), and varies according to species. Some limited systemic activity may occur as a result of movement around the leaf, from leaf to leaf and from leaf to roots. This may be sufficient to suppress the regrowth of perennial weeds that are not killed outright by contact activity. However, it will often not provide the ‘roots and all’ kill seen by glyphosate for many perennial grass weeds.

Strategy for Basta resistance

The natural occurrence of bialaphos provides a lead to follow in devising a strategy for engineering resistance. Toxic compounds such as PPT, with a simple structural homology to a common substrate such as glutamate, are likely to be toxic to the host organism. The fact that the compound is synthesised as an inactive precursor is indicative of this fact. It is therefore not surprising that *Streptomyces* spp. also contain a detoxification gene that protects the organism from the toxic effects of PPT. The *bar* gene of *Streptomyces hygroscopicus* and the closely related *pat* gene of *Streptomyces viridochromogenes* code for the enzyme phosphinothricin acetyltransferase (PAT). The addition of an acetyl group to the amino group of phosphinothricin inactivates the compound (Figure 5.4). Thus, transferring this gene to a plant should, in theory, provide resistance against phosphinothricin. This approach to the engineering of phosphinothricin resistance was developed by Plant Genetic Systems under contract from Hoechst. (PGS was subsequently acquired by AgrEvo). The *bar* gene has now been integrated into many different plants, usually under the control of

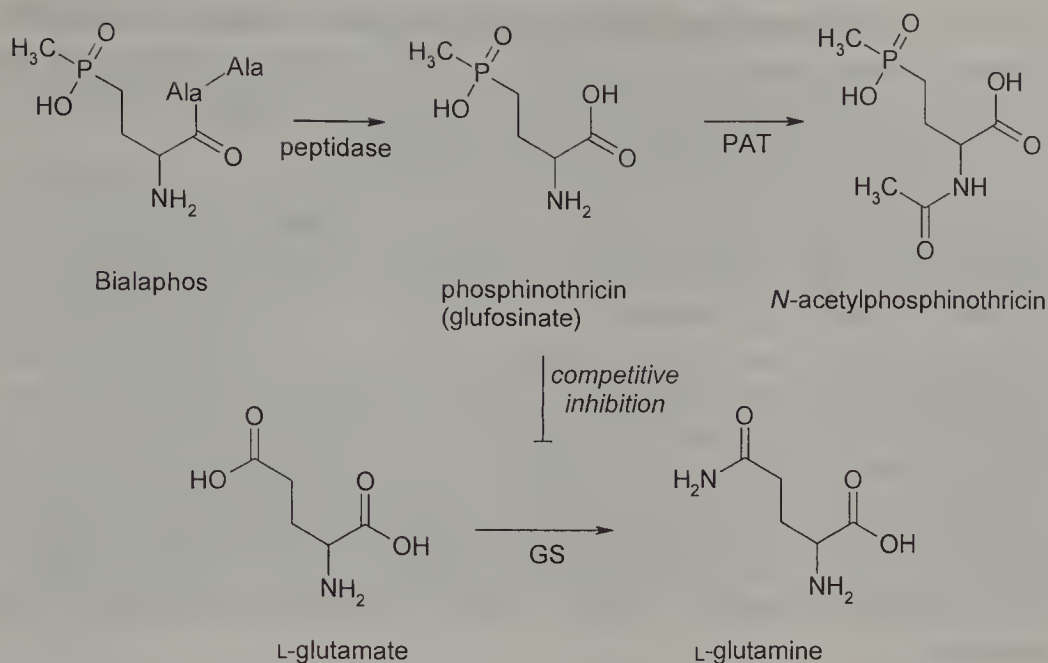


Figure 5.4 The formation, mode of action and detoxification of phosphinothricin. The conversion of bialaphos to phosphinothricin involves removal of the two alanine residues by a peptidase. The compound acts as a competitive inhibitor of glutamine synthase (GS), and the figure highlights the structural similarity between phosphinothricin and the substrate L-glutamate. The detoxification reaction catalysed by phosphinothricin acetyltransferase (PAT) is also shown.

the 35S promoter. The current major LibertyLink crop lines supplied by Aventis are oilseed rape, maize and chicory.

The *bar* gene has also proved to be useful as a selectable marker for the transformation and regeneration of transgenic plants. It provides an alternative to selection with antibiotics such as kanamycin, to which different species have a highly varied response (see Chapters 3 and 4). Transgenic plants can be selected directly on PPT medium, but caution is required. Inhibition of GS by PPT causes NH_3 accumulation in non-transgenic material and hence death of the plant tissue, but accumulation of NH_3 in the non-transformed tissue can also cause problems of toxicity to neighbouring transformed cells.

Prospects for plant detoxification systems

One other strategy that has yet to be fully exploited is the possibility of enhancing endogenous plant detoxification mechanisms. Many xenobiotic (foreign) compounds are detoxified in plants but the pathways may involve more than one step, such as hydroxylation, conjugation and transport stages, so it may prove difficult to identify single-gene mechanisms to engineer resistance. The hydroxylation of compounds involves enzymes such as the cytochrome P450 monooxygenases, which form a large gene family. For example, the analysis of weeds resistant to the herbicide bromoxynil (Table 5.4) revealed that the bromoxynil was being detoxified by an endogenous cytochrome P450 monooxygenase. This offers the opportunity to use

endogenous plant genes to enhance resistance against a range of herbicides. However, more research is required to identify which members of the cytochrome P450 gene family are specific for particular classes of xenobiotics, and which have roles in normal metabolic pathways.

Plant detoxification pathways often involve conjugation to glutathione by glutathione *S*-transferase (GST) activity, and specific transport of the conjugate into the vacuole. GSTs also comprise a large gene family, some members of which are known to be involved in endogenous metabolic reactions. In some cases, the hydroxylation and conjugation pathways operate in concert. Hence, the resistance of maize to atrazine is ascribed to a two-step pathway involving both 2-hydroxylation and conjugation to glutathione (Figure 5.5). Thus, there is the potential here to enhance endogenous systems, or transfer systems between plant species, once more information about the functions of these large gene families is available. The exploitation of functional genomics techniques (Chapter 1) will accelerate the acquisition of this knowledge.

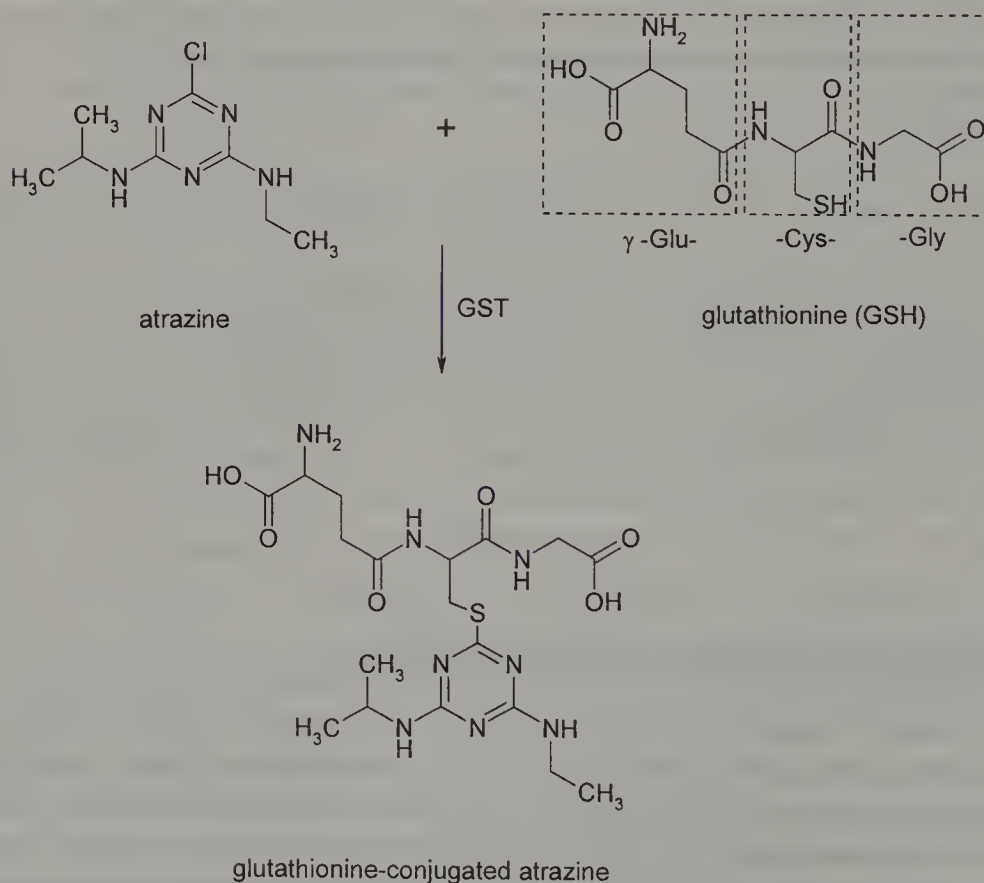


Figure 5.5 The conjugation of atrazine to glutathione. The detoxification of atrazine is a two-stage process involving a 2-hydroxylation step (removing the chlorine residue) prior to the addition of glutathione. Glutathione is a tripeptide in which the key residue is the middle cysteine. The SH group is involved in a number of redox and conjugation reactions (see also Chapter 9).

Commercialisation of herbicide-resistant plants to date

This chapter has described the two most widespread examples of genetically manipulated herbicide resistance. Table 5.5 indicates those crops resistant to a number of other herbicides that have been developed to the field-trial stage. It is important to note that some of these have not been widely planted commercially. The atrazine-resistant crops, for example, have not been developed further, given the environmental concerns about the use of this class of persistent herbicides.

The table gives the strategy employed in each case. All these examples use either the mutant target enzyme approach, or the prokaryotic detoxification gene approach, which have been described in detail when considering the glyphosate and phosphinothricin case studies. However, one more example is worth exploring further, because it demonstrates the use of homologous recombination to modify an endogenous plant gene to engineer herbicide resistance.

CASE STUDY 3 **Engineering imidazolinone resistance by targeted modification of endogenous plant genes**

As stated near the beginning of this chapter, two important classes of herbicide (sulphonylureas and imidazolinones) have the same mode of action, that is they inhibit the enzyme acetolactate synthase (ALS) (or more correctly, acetohydroxy-acid synthase or AHAS) (see Table 5.3 and Figure 5.6). This blocks the synthesis of the aliphatic amino acids isoleucine and/or leucine and valine. Herbicide-resistant forms of this enzyme have been isolated from a range of species, from bacteria and yeast to plants, and some of these have been used to engineer herbicide resistance by standard 'Strategy 2' approaches. The sequence analysis of various resistant ALS genes also provided information about the precise changes in amino acid sequence responsible for the resistant character. Researchers at Pioneer Hi-bred International used this type of information to predict that a single base change in the endogenous maize ALS gene family would be sufficient to change Ser 621 (encoded by AGT) to Asn 621 (AAT) and that this should confer resistance to the herbicide Lightning (a mixture of imazathapyr [Table 5.3] and imazapyr). This was actually achieved by bombarding maize cells with an oligonucleotide made up of a combination of DNA and RNA bases, with a 32-base section having exact homology to the target sequence of the endogenous maize genes, apart from the single base mismatch to give the desired mutation. Cells that were able to grow and develop callus on medium containing imazathapyr were selected, and plants were regenerated at an estimated transformation efficiency of 10^{-4} .

Some nine independently transformed plants were tested for Lightning tolerance—three were resistant to fourfold the normal field dose, whilst four others were only slightly injured by the fourfold dose and were able to tolerate the normal field dose. The remaining two plants were as susceptible to the herbicide as control plants. There was a direct correlation between Lightning resistance and the sequence of ALS genes in the transformed plants. The highly resistant plants contained the predicted AGT to AAT mutation, whilst the

Table 5.5 Herbicide-resistant crops now in the field

Class of herbicide	Compound/herbicide	Transgene/mechanism	Companies	Crops
Glycine	Glyphosate (Roundup)	<i>Agrobacterium</i> CP4-resistant gene	Monsanto	Soybean, rape, tomato
	Glyphosate (Roundup)	Maize resistant gene	Monsanto	Maize
	Glyphosate (Roundup)	Oxidoreductase detoxification	Monsanto	Maize, rape, soybean
Phosphinic acid	Phosphinothricin (Basta), (Liberty)	<i>bar</i> gene—phosphinothricin acetyltransferase detoxification	Hoechst/AgrEvo/Aventis	Maize, rice, wheat, cotton, rape, potato, tomato, sugar beet
			Novartis/Syngenta	
Sulphonylurea	Chlorsulphuron (Glean)	Mutant plant acetolactate synthase	DuPont—Pioneer Hi-Bred	Rape, rice, flax, tomato, sugar beet, maize
Imidazolinone	(Arsenal)	Mutant plant acetolactate synthase	American Cyanamid	
S-triazines	Atrazine (Lasso)	Mutant plant chloroplast <i>psbA</i> gene	DuPont, Ciba-Geigy/Novartis	Soybean
Nitriles	Bromoxynil (Buctril)	Nitrilase detoxification	Calgene	Cotton, rape, potato, tomato
Phenoxy-carboxylic acids	2,4-D	Monooxygenase detoxification	Schering/AgrEvo	Maize, cotton

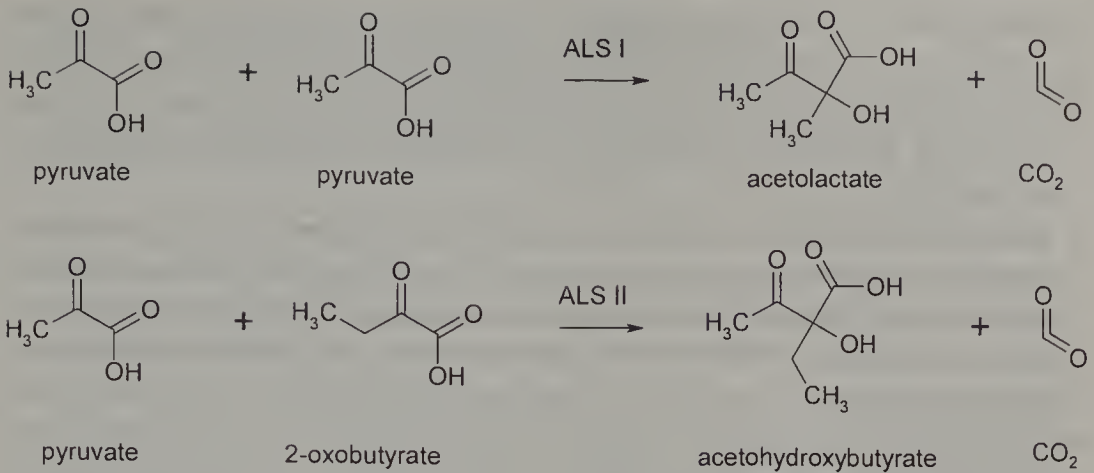


Figure 5.6 The reactions catalysed by acetolactate synthase forms I and II. The formation of acetolactate by ALS I is the first step in the biosynthesis of leucine and valine, whilst the similar reaction catalysed by ALS II is a key step in isoleucine synthesis.

moderately resistant plants showed mutations within a few bases of the target site. In contrast, the sequence of the ALS genes was not altered in the susceptible plants.

The conclusion drawn from this work is that although homologous recombination is a rare event in plants, it can be used to change endogenous gene sequences under certain circumstances. Does this therefore mean that many more traits will be engineered by this type of procedure in the future? There are certainly many advantages to the manipulation of an endogenous gene *in situ*, since it avoids many of the problems associated with transgene cloning, modification, expression and targeting, as well as the requirement for a selectable marker gene, unpredictable effects following random insertion into the genome, etc. However, this procedure will only be suitable for certain types of genetic modification. First, the precise base change needed to produce the desired mutation must be known. Second, the targeted gene mutation must confer a trait that is selectable at this very low frequency of transformation. This is ideal for herbicide resistance, but generally not for the types of trait such as pest and disease resistance that are discussed in the next three chapters.

The environmental impact of herbicide-resistant crops

The predominance of herbicide-resistant crops in terms of their development and commercial growing was highlighted at the beginning of this chapter. Some of the reasons for this trait maintaining its position in the GM crops ‘league tables’ have been discussed. Scientifically, herbicide resistance has a number of advantages that facilitated its rapid development. A number of different single-gene strategies using accessible genes are possible, and the trait itself can be used as a selectable marker. These factors, combined with the research and development impetus provided by the agrochemical industry,

inevitably led to their prime position. However, it should be noted that herbicide-resistant crops do not top the league table just because they had such a favourable 'head start'. They have maintained their leading position because of the rapid adoption of these crops in countries where the technology has been accepted.

In the USA, where the greatest area of GM crops are grown, the proportion of herbicide-resistant soybean rose from 17% of the soybean acreage in 1997 to 68% in 2001. Similarly, herbicide-resistant cotton expanded from 10% of the cotton acreage in 1997 to 56% in 2001. This rapid adoption of herbicide-resistant crops begs a number of questions. The first one is whether farmers are adopting these crops so rapidly because there is a clear commercial advantage for them, and if so, what is it? The second one relates to the environmental impact of this rapid switch in growing patterns, and hence herbicide usage patterns. Box 5.5 presents one analysis of the economic benefits of growing glyphosate-resistant soybean, and its impact on herbicide use. This indicates that although the use of glyphosate increased compared to that for non-transgenic soybean, the total use of herbicides decreased. Indeed, the major beneficiary of this technology was the farmer rather than the producer, mainly by way of a marked overall reduction in herbicide costs. Thus, the assumption that herbicide-resistant crops will inevitably lead to a greater use of herbicides needs to be qualified. The use of the target herbicide may well increase, but this may displace or reduce the requirement for other herbicides, such that there is a net reduction in their use. Since the environmental impact of herbicides such as glyphosate and phosphinothricin is thought to be lower than many of the compounds they are displacing, the producers of herbicide-resistant crops can claim that the adoption of their crops will have a positive effect on the environment. However, there are other environmental concerns about the rapid adoption of herbicide-resistant crops. One area of concern is the reduction of biodiversity as a result of the more efficient removal of weed species from arable land. These concerns relate more widely to current intensive agricultural practices, and lie outside the direct scope of this book. Another area of concern relates to the possibility of encouraging the development of herbicide-resistant weeds, and is discussed below.

The development of 'super weeds'

One of the major concerns about the use of herbicide-resistant crops is that their introduction will lead to the appearance of 'super weeds', which could evade control by the commonly used herbicides. It is important to put these fears into context. We have previously discussed the fact that some crops are already resistant to certain herbicides, and that herbicide-resistant weeds may appear naturally wherever there is a selective pressure created by the repeated use of the same herbicide. Thus, the problem of herbicide-resistant weeds is not a new phenomenon unique to GM technology and, bearing in mind the

BOX 5.5 Roundup Ready soybean—glyphosate usage

It has been calculated that from 1997 to 1998 there was an 81% increase in glyphosate use (up by 5540 tonnes) in line with an increase in the proportion of GM soybean from 13% to 36% of the crop. However, this was more than compensated for by a reduction in the use of other herbicides on soybean of 8990 tonnes, resulting in a net reduction in total herbicide use of 3360 tonnes (–9.7%). Since glyphosate has better properties than some of the other herbicides in terms of toxicity, environmental persistence and biodegradability, this is claimed to be of net benefit to the environment. This also has an impact on the production costs of the crop. For example, typical 1999 production costs (in \$US) of Roundup Ready and non-GM soybean have been calculated to be ~\$90 and ~\$100 per acre, respectively, showing a net saving of \$10 per acre. The large saving in herbicide costs more than compensates for the technology fee charged for the Roundup Ready crop.

It is clear that there is a significant reduction in production costs associated with the use of Roundup Ready soybeans, derived principally from the reduced use of herbicides. In addition, it has been estimated that most of the economic benefits accrue to the farmer, as demonstrated in the table below.

Beneficiaries	Estimated benefits (\$USm)	Distribution of benefits (%)
Seed companies	32	3
US consumer	42	4
Technology inventor	74	7
US farmer	769	76
Total benefits	1061	100

Data from Falck-Zepada, Traxler and Nelson (2000) USDA Economic Research Service. Weblink 5.9

large armoury of herbicides listed in Table 5.2, is not one that need pose a major threat. Nevertheless, it is sensible to ask whether the widespread use of herbicide-resistant crops will make a difference to the number, rate of appearance and type of resistant weeds because: (1) they encourage the repeated use of the same herbicide, and (2) the herbicide resistance gene is transferred to a weed population by one of the processes of ‘gene flow’ described below.

Herbicide-resistant weeds could theoretically arise by three types of mechanism:

1. The herbicide-resistant crop itself appears as a ‘volunteer’ weed in fields where rotational crops are grown. This volunteer population could reproduce outside of cultivation and form a self-sustaining weed population.
2. Pollen from the herbicide-resistant crop fertilises weedy relatives of the crop plant, producing herbicide-resistant hybrids. Subsequent

backcrossing of these hybrids with the weed species could lead to introgression of the herbicide-resistant trait into the weed population.

3. There is horizontal gene transfer (as opposed to the vertical transmission in points 1 and 2) by other mechanisms (e.g. viruses) that spreads the herbicide-resistant trait into a much wider range of plant species.

Just as the environmental impact of each herbicide should be considered separately, it is also important to treat different crops on a case-by-case basis. Thus, certain crops are normally more prone to form volunteer weeds than others, and the presence of a herbicide-resistant strain could be a problem if that herbicide is normally used to clear the volunteers.

However, the large arsenal of herbicides shown in Table 5.2 should ensure that volunteer weeds can be cleared from cultivated land for rotational crops. It must be borne in mind that “weediness” is a complex mix of characteristics that do not normally include herbicide resistance. When considering gene flow from a GM crop to other crops and weed species, the nature of each individual crop must also be considered. For example, the spread of pollen from out-crossing crops is more of a problem than with crops that are self-crossing. Another factor to consider is the proximity of closely related weed species. Thus, maize, soybean and cotton have no compatible wild-type relatives in the countries where GM varieties are widely grown (USA and Canada). Wheat, oilseed rape and sugar beet are grown in close proximity to related species, but only the latter two are out-crossing. For this reason, oilseed rape and sugar beet have been used as “worst case analysis” crops for the study of gene flow into weeds. The studies that have been carried out to date tend to indicate that there is detectable gene flow from crops to weed species, but that transgene introgression under field conditions is probably rare. In order to minimise this potential gene flow, it is necessary to predict the distance that pollen can travel from each GM crop, and to establish appropriate buffer distances to prevent the fertilisation of weedy relatives. However, many of the studies to determine buffer distances have been on a relatively small scale. A recent large-scale study using herbicide-resistant oilseed rape (produced by non-transgenic means) does indicate that oilseed rape pollen can travel much further than had been suspected, but that nevertheless, the amount of gene flow to non-resistant plants was minimal.

There is as yet little evidence that horizontal gene flow between plants poses a particular problem for the transfer of transgenes into unrelated species. What is clear about the entire issue of environmental impact of herbicide resistant crops is that much more research is required to establish the full extent of all types of gene flow between plants in the environment. Thus, a recent review of the environmental risks and benefits of GM crops concludes that:

1. The risks and benefits of GM crops are not entirely certain or universal.
2. The ability to predict ecological impacts of any introduced species (GM or not) is difficult, and available data have limitations.

3. Some benefits and risks may exist that have not yet been identified or addressed in published literature.
4. The quantity and quality of different GM crops that may eventually be developed merit special consideration for risk assessment.
5. Better evaluation of potential benefits will help risk managers know how these balance with potential risks.
6. Measures developed to prevent gene transfer to wild plants can reduce the potential environmental impacts and prolong potential benefits.

Summary

This chapter has given an overview of the genetic manipulation of herbicide-resistant plants. The wide range of different chemical families of herbicides and their modes of action have been described. Four basic strategies for the genetic engineering of herbicide resistance have been discussed with reference to a number of case studies. It has been concluded that Strategy 2 (resistant target protein) and Strategy 3 (specific detoxification) are currently the most effective approaches. However, future developments may well include the more precise enhancement of endogenous detoxification mechanisms and the targeted mutation of endogenous resistance genes. Concerns about the environmental impact of herbicide-resistant crops have been touched upon, with a view to returning to these wider issues in Chapter 12.

The next four chapters deal with the genetic manipulation of resistance to other stresses—pests, diseases and abiotic stress. Some of the concepts introduced in this chapter will be recurring themes in the subsequent chapters. However, we will also introduce new ideas and techniques in a logical sequence as we progress through the next chapters.

Further reading

Plant Biochemistry

- Buchanan, B. B., Gruissem, W. and Jones, R. L. *Biochemistry and Molecular Biology of Plants* (2000) American Society of Plant Physiologists. Rockville, Maryland, USA
- Heldt, H.-W. *Plant Biochemistry and Molecular Biology*. (1997) Oxford University Press, Oxford UK

Herbicides

- Herbicide Resistance Action Committee and the Weed Science Society of America—web-link 5.1: <http://www.plantprotection.org/HRAC/MOA.html>
- Pesticide Action Network Pesticide Database-web-link 5.2: <http://www.pesticideinfo.org>

GM crops-general

- AgBioTechNet-web-link 5.3: <http://www.agbiotech.net/>
- GEO-PIE Genetically Engineered Organisms—Public Issues Education Project. Cornell University-web-link 5.4: <http://www.geo-pie.edu/traits/herbres.html>

James C. International Service for the Acquisition of Agri-biotech Applications-web-link 5.5: <http://www.isaaa.org/>

Northern Light special edition on GM foods-web-link 5.6: <http://www.special.northernlight.com/gmfoods/>

Herbicide resistant crops

AgBiotechInfonet—web-link 5.7:

<http://www.biotech-info.net/herbicide-tolerance.htm/>

Felsot A. Articles in issues 173, 175, 176 and 178. Agrichemical and Environmental News 2000–2001—web-link 5.8: <http://www.tricity.wsu.edu/aenews>

Cole, D. J. (1994). Molecular mechanisms to confer herbicide resistance. In *Molecular biology in crop protection* (ed. G. Marshall and D. Walters), pp. 146–76. Chapman and Hall, London UK

Mullineaux, P. M. (1992). Genetically engineered plants for herbicide resistance. In *Plant genetic manipulation for crop protection* (ed. A. M. R. Gatehouse, V. A. Hilder and D. Boulter), pp. 75–107. C. A. B. International, Wallingford UK

Homologous recombination

Zhu, T., Peterson, D. J., Tagliani, L., St Clair, G., Baszczynski, C. L. and Bowen, B. (1999). Targeted manipulation of maize genes *in vivo* using chimeric RNA/DNA oligonucleotides. *Proceedings of the National Academy of Sciences USA*, 96, 8786–8773.

Zhu, T., Mettenburg, K., Peterson, D. J., Tagliani, L. and Baszczynski, C. L. (2000). Engineering herbicide-resistant maize using chimeric RNA/DNA oligonucleotides. *Nature Biotechnology* 18, 555–8.

Environmental impact

Rieger, M. A., Lamond, M., Preston, C., Powles, S. B. and Roush, R. T. (2002). Pollen-mediated movement of herbicide resistance between commercial canola fields. *Science*, 296, 2386–8.

Senior, I. J. and Dale, P. J. (2002). Herbicide-tolerant crops in agriculture: oilseed rape as a case study. *Plant Breeding*, 121, 97–107.

Wolfenbarger, L. L. and Phifer, P. R. (2000). The ecological risks and benefits of genetically engineered plants. *Science*, 290, 2088–93.

Herbicide use and commercial benefits

USDA Economic Research Service—web-link 5.9: <http://www.ers.usda.gov/>*

Agrochemical Industry

Aventis—web-link 5.10: <http://www.aventis.com>

Biotechnology Industry Organisation—web-link 5.11: <http://www.bio.org/foodag/>

Council for Biotechnology Information—web-link 5.12: <http://www.whybiotech.com>

DuPont—web-link 5.13: <http://www.dupont.com/biotech/>

Montsanto company—web-link 5.14: <http://www.monsanto.com/com>

Monsanto Biotechnology Knowledge Centre—web-link 5.15: <http://www.biotechknowledge.monsanto.com>

Pioneer Hi-Bred—web-link 5.16: <http://www.pioneer.com/biotech/default.htm>

6

The genetic manipulation of pest resistance

Introduction

In Chapter 5, the scale of the problem of world crops lost to weeds, pests and diseases was discussed. It has been estimated that 13% of the potential world crop yield is lost to pests (see Figure 5.1). Plant pests range from nematodes to birds and mammals, but insect pests cause a major proportion of the total pest damage to crops. This chapter will therefore focus upon biotechnological solutions to the problem of insect pest damage to crops.

Far from reducing pest damage, modern agricultural practices have, if anything, exacerbated the problem. One of the contributory factors to this exacerbation has been the breeding out of endogenous pesticidal traits. Some defensive characters may have been lost accidentally during the selection process for other properties, but pesticidal traits may also have been removed 'deliberately' because they affect other characteristics like yield and quality. For whatever reason, the fact remains that many current elite cultivars have less natural resistance to pests than did their predecessors. This problem is compounded by the practice of growing monocultures, since the growing of a single crop over a wide area in repeated years encourages the build-up of pests. These negative effects of modern agriculture on pest damage have encouraged an increasing reliance on chemical pesticides, but this has tended to drive a cycle in which each new pesticide just keeps pace with the appearance of resistance to the previous pesticide in the insect population. Against this background, the possibility of developing insect-resistant crops that could help to reduce insect damage, whilst reducing the reliance on chemical insecticides, has a number of attractions. This chapter will consider the scientific strategies for achieving pest resistance before returning to discuss the potential benefits and drawbacks of GM approaches.

The nature and scale of insect pest damage to crops

Before examining GM strategies for developing insect resistance, it is useful to consider some of the characteristics of the insects causing the damage. The

Table 6.1 Common insect pests of major crops

Insect species	Common name of pest	Order	Crops affected
<i>Ostrinia nubilalis</i>	European corn borer	Lepidoptera	Maize
<i>Heliothis virescens</i>	Tobacco budworm	Lepidoptera	Tobacco, cotton
<i>Heliothis armigera</i>	Old world bollworm tomato fruitworm	Lepidoptera	Cotton, tomato
<i>Helicoverpa zea</i>	Cotton bollworm	Lepidoptera	Cotton
<i>Manduca sexta</i>	Tobacco hornworm	Lepidoptera	Tobacco, tomato, potato
<i>Spodoptera littoralis</i>	Cotton leafworm	Lepidoptera	Maize, rice, cotton, tobacco
<i>Leptinotarsa decemlineata</i>	Colorado beetle	Coleoptera	Potato
<i>Callosobruchus maculatus</i>	Cowpea seed beetle	Coleoptera	Cowpea, soybean
<i>Tribolium confusum</i>	Confused flour beetle	Coleoptera	Cereal flours
<i>Locusta migratoria</i>	Locust	Orthoptera	Grasses
<i>Nilaparvata lugens</i>	Brown plant hopper	Homoptera	Rice

first point to make is that, whilst some adult insects feed off plants and can damage crops (think of a plague of locusts), most of the problems are caused by insect larvae. The major classes of insect that cause crop damage are the orders Lepidoptera (butterflies and moths), Diptera (flies and mosquitoes), Orthoptera (grasshoppers, crickets), Homoptera (aphids) and Coleoptera (beetles). Table 6.1 lists a few of the common pests that cause extensive damage to some of the major crops of the world.

GM strategies for insect resistance: The *Bacillus thuringiensis* approach

In this chapter, two approaches will be compared: first, the use of bacterial insecticidal genes to provide protection from pest damage; and, second, the potential for using endogenous plant protection mechanisms (a ‘Copy Nature’ approach). By far the most widespread example of the first approach is the use of the *cry* endotoxin genes from *Bacillus thuringiensis*.

B. thuringiensis was discovered by Ishiwaki in 1901 in diseased silkworms, and was subsequently classified and named after its isolation from the gut of diseased flour moth larvae in Thuringberg, by Ernst Berliner. The adverse effect of the bacteria on the insect larvae was subsequently identified as arising

Table 6.2 Classification of the insecticidal crystal protein genes of *Bacillus thuringiensis*

<i>cry</i> gene families	Protein size (kDa)	<i>B. thuringiensis</i> subspecies/strain of holotype	Susceptible insect class
<i>cry1Aa</i> (1–14)	133	kurstaki	Lepidoptera
<i>cry1Ab</i> (1–16)	130	berliner	Lepidoptera
<i>cry1Ac</i> (1–15)	133	kurstaki	Lepidoptera
<i>cry1Ad-g</i>	133	aizawai	Lepidoptera
<i>cry1Ba</i> (1–4)	140	kurstaki	Lepidoptera
<i>cry1Bb-g</i>	1340	EG5847	Lepidoptera
<i>cry1Ca</i> (1–8)	134	entomocidus	Lepidoptera
<i>cry1Cb</i> (1–2)	133	galleriae	Lepidoptera
<i>cry1Da</i> (1–2)	132	aizawai	Lepidoptera
<i>cry1Db</i> (1–2)	131	BTS00349A	
<i>cry1Ea</i> (1–6)	133	kenyae	Lepidoptera
<i>cry1Eb1</i>	134	aizawai	Lepidoptera
<i>cry1Fa</i> (1–2)	134	aizawai	Lepidoptera
<i>cry1Fb</i> (1–5)	132	morrisoni	
<i>cry1Ga</i> (1–2)	132	BTS00349A	
<i>cry1Gb</i> (1–2)	133	wuhanensis	Lepidoptera
<i>cry1Ha-b</i>	133	BTS02069AA	
<i>cry1Ia</i> (1–9)	81	kurstaki	Lepidoptera
<i>cry1Ib-e</i>	81	entomocidus	Lepidoptera & Coleoptera
<i>cry1Ja-d</i>	133	EG5847	Lepidoptera
<i>cry1Ka1</i>	137	morrisoni	Lepidoptera
<i>cry2Aa</i> (1–10)	71	kurstaki	Lepidoptera & Diptera
<i>cry2Ab</i> (1–5)	71	kurstaki	Lepidoptera
<i>cry2Ac</i> (1–2)	70	shanghai	Lepidoptera
<i>cry3Aa</i> (1–7)	73	tenebrionis	Coleoptera
<i>cry3Ba</i> (1–2)	75	tolworthi	Coleoptera
<i>cry3Bb</i> (1–3)	74	EG4961	Coleoptera
<i>cry3Ca1</i>	73	kurstaki	Coleoptera
<i>cry4Aa</i> (1–3)	135	israelensis	Diptera
<i>cry4Ba</i> (1–5)	128	israelensis	Diptera
<i>cry5Aa1</i>	152	darmstadiensis	Nematodes
<i>cry5Ab1</i>	142	darmstadiensis	Nematodes
<i>cry5Ac1</i>	135	PS86Q3	Hymenoptera
<i>cry5Ba1</i>	140	PS86Q3	Hymenoptera
<i>cry6Aa</i> (1–2)		PS52A1	Nematodes
<i>cry6Ba1</i>		PS69D1	Nematodes
<i>cry7Aa1</i>	129	galleriae	Coleoptera
<i>cry7Ab</i> (1–2)	130	dakota	Coleoptera
<i>cry8A-D</i>	131	kumamotoensis	Coleoptera
<i>cry9Aa</i> (1–2)	130	galleriae	Lepidoptera
<i>cry9Ba1</i>		galleriae	Lepidoptera
<i>cry9Ca1</i>	130	tolworthi	Lepidoptera
<i>cry9Da</i> (1–2)	132	japonensis	
<i>cry10Aa1</i>	78	israelensis	Diptera
<i>cry11Aa</i> (1–2)	72	israelensis	Diptera
<i>cry11Ba-b</i>	81	jegathesan	Diptera
<i>cry12 – 40</i>	various	various	various

from a number of toxins produced by the bacteria. The bacterium produces an insecticidal crystal protein (ICP) which forms inclusion bodies of regular bipyramidal or cuboidal crystals during sporulation. ICPs are one of several classes of endotoxins produced by the sporulating bacteria, hence they were originally classified as δ -endotoxins, to distinguish them from other classes of α -, β - and γ -endotoxins.

The structure of the *cry* genes and their δ -endotoxin products have been well characterised. The *cry* genes are carried on plasmids and belong to a superfamily of related genes. Table 6.2 shows the classification of the *cry* gene superfamily according to size and sequence similarities. This classification was introduced in 1998 and supersedes previous nomenclature that may still appear in the literature. The sequence comparison indicates a large number of distinct families (*cry1*–*cry40* at the time of publication) and within each family there may be three further levels or ranks of subfamily. Hence, *cry1Aa1* is very closely related to the other *cry1Aa* genes, which are all closely related to the other *cry1A* genes, which all form part of the *cry1* family. Table 6.3 shows that the strains of *B. thuringiensis* produce a wide range of different crystal proteins. Thus, it is not just a case of each strain containing one specific *cry* gene encoding one particular crystal.

Apart from sequence similarity, it is apparent that there is a large difference in size between different Cry proteins, though they tend to cluster as either large (~130 kDa) or small (~70 kDa) proteins. Although there is a considerable difference in size between the different subfamilies, they share a common active core comprising three domains. Figure 6.1 shows a simple alignment of different classes of Cry protein, showing the common core domains. It can be seen that the N-terminal end of each gene has a similar organisation, despite the great difference in overall length. In fact, the larger proteins such as the Cry1 group are inactive, and are activated by proteolytic removal of the C-terminal sequence.

Figure 6.2 shows the structure of the active core of the Cry1A protein, exemplifying the three discrete domains, each with a quite different folding structure. Domain I at the N-terminal end comprises a series of α -helices arranged in a cylindrical formation. The open cylindrical structure of this domain is thought to be responsible for creating a pore through the membrane of the insect midgut. Domain II comprises a triple β -sheet and is involved in receptor recognition. Domain III is a β -sandwich, with a number of putative roles including protection from degradation, toxin/bilayer interactions and receptor binding.

The mode of action of δ -endotoxins involves a specific interaction between the protein and the insect larva midgut. After ingestion by an insect larva, the protein crystals are solubilised in its midgut. The larger proteins such as the 130-kDa Cry1 group are proteolytically cleaved at this stage to release the active 55–70-kDa active fragment of the protein. This interacts with high-affinity receptors in the midgut brush-border membrane. The result of this

Table 6.3 The range of insecticidal crystal proteins in individual *Bacillus thuringiensis* strains

<i>B.t.</i> subpecies and strains	Crystal protein
<i>aizawai</i>	Cry1Aa, Cry1Ab, Cry1Ad, Cry1Ca, Cry1Da, Cry1Eb, Cry1Fa, Cry9Ea, Cry39Aa, Cry40Aa
<i>entomocidus</i>	Cry1Aa, Cry1Ba, Cry1Ca, Cry1Ib
<i>galleriae</i>	Cry1Ab, Cry1Ac, Cry1Da, Cry1Cb, Cry7Aa, Cry8Da, Cry9Aa, Cry9Ba
<i>israelensis</i>	Cry10Aa, Cry11Aa
<i>japonensis</i>	Cry8Ca, Cry9Da
<i>jegathesan</i>	Cry11Ba, Cry19Aa, Cry24Aa, Cry25Aa
<i>kenyae</i>	Cry2Aa, Cry1Ea, Cry1Ac
<i>kumamotoensis</i>	Cry7Ab, Cry8Aa, Cry8Ba
<i>kurstaki</i> HD-1	Cry1Aa, Cry1Ab, Cry1Ac, Cry1Ia, Cry2Aa, Cry2Ab
<i>kurstaki</i> HD-73	Cry1Ac
<i>kurstaki</i> NRD-12	Cry1Aa, Cry1Ab, Cry1Ac
<i>morrisoni</i>	Cry1Bc, Cry1Fb, Cry1Hb, Cry1Ka, Cry3Aa
<i>tenebrionis</i>	Cry3Aa
<i>tolworthi</i>	Cry3Ba, Cry9Ca
<i>wuhanensis</i>	Cry1Bd, Cry1Ga, Cry1Gb

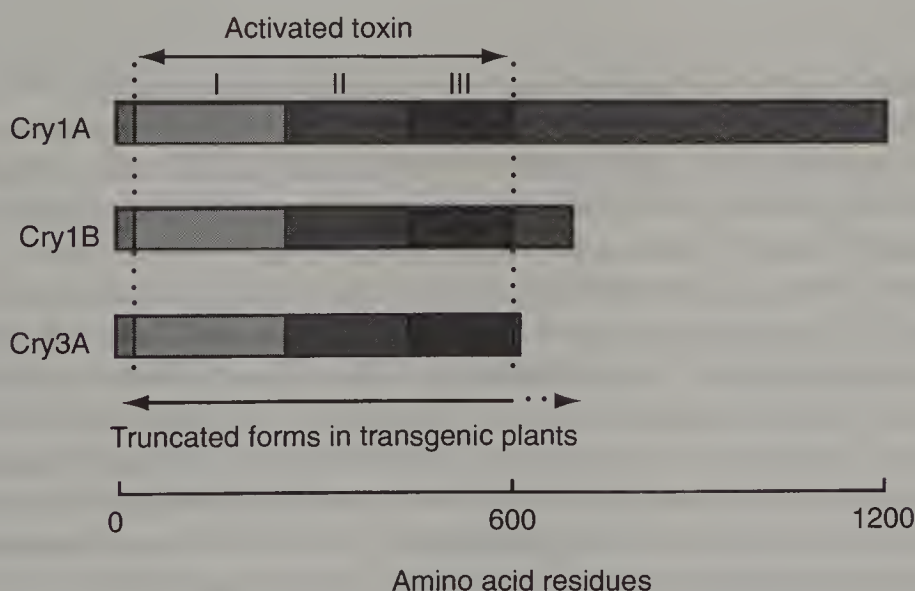


Figure 6.1 Comparison of the structures of different classes of Cry protein. Three different classes of Cry protein are compared to demonstrate the differences in overall size, but the alignment of the common core region that comprises the activated toxin. The organisation of the core into domains I, II and III (see Figure 6.2) is also indicated. The N- and C-terminal extensions are trimmed by insect gut proteases to release the active toxin. In some transgenic plants, a truncated, active form of the protein is produced directly. (Redrawn with permission from de Maagd, *et al.* (1999).)

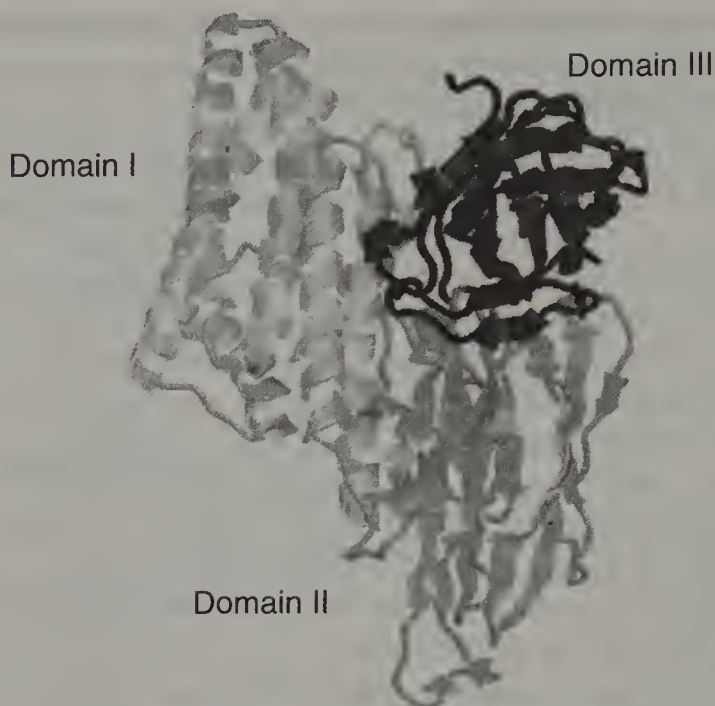


Figure 6.2 Ribbon model of Cry1Aa toxin molecule. The structure of the activated toxin is shown, demonstrating the three distinct domains. The α -helical cylinder that forms domain I is involved in membrane insertion and pore formation. Domains II and III are involved in recognition and binding to receptors in the insect midgut wall. (Redrawn with permission from de Maagd, *et al.* (1999).)

binding is to open cation-selective pores in the membrane. The flow of cations into the cells results in osmotic lysis of the midgut epithelium cells, causing their destruction. Thus, the δ -endotoxins are extremely toxic and can be lethal to susceptible insect larvae at relatively low concentrations. On the other hand, their toxicity to other animals (e.g. mammals) is extremely low.

The conditions in the insect larva midgut vary according to insect class. The midgut of Lepidoptera and Diptera is mildly alkaline, whilst the coleopteran gut is generally either more alkaline or acidic. These different conditions favour the solubilisation and activation of different Cry subfamilies. Furthermore, the specificity of the interaction between the endotoxin and the midgut receptor means that individual Cry proteins are active against particular insect larvae. Table 6.2 indicates the specificity of some of the best characterised groups of Cry proteins.

The use of 'Bt' as a biopesticide

Preparations of *B. thuringiensis* spores or isolated crystals have now been used as an 'organic' pesticide for half a century. The isolated crystals have a

limited persistence on foliage of a few days, whilst the spore preparations are effective for about 40 days on foliage and up to 2 years in soil. Neither method of application has a particularly effective penetration, with regard to concealed surfaces and organs of the plant, or against sap-sucking insects—i.e. they are not systemic. However, the agronomic experience of using *B. thuringiensis* spores or isolated crystals as a biopesticide, coupled to the associated safety data and regulatory approvals, has been useful for the rapid development of the ‘*Bt*’ strategy to genetic modification.

Note that *Bt* is generally used as a shorthand for a crop transformed with a *cry* gene (hence *Bt* cotton, etc.), and also for the Cry proteins (hence *Bt* protein). This can be confusing, bearing in mind that we have now met four different terms for effectively the same group of proteins—ICP, δ -endotoxin, Cry and now *Bt*.

***Bt*-based genetic modification of plants**

Although the GM approach to using the *cry* genes to obtain pest resistance in plants is conceptually simple, it does provide an object lesson in the detailed molecular biology that may be required to achieve high levels of expression of a bacterial gene in a transgenic plant. This goes beyond the obvious requirements for plant promoter and terminator sequences to regulate transcription as described in previous chapters. The first attempts to express Cry1A and Cry3A proteins under the control of the CaMV 35S or *Agrobacterium* T-DNA promoters resulted in very low levels of expression in tobacco, tomato and potato plants. It was realised that the prokaryotic gene sequence itself would need to be extensively modified in order to obtain high levels of stable expression. These modifications are described in detail in Chapter 4 (Box 4.3). The eventual result was that expression was enhanced by 100-fold to give much better levels of expression, of the order of 100 ng *Bt* protein per mg total protein. Subsequent laboratory tests showed that the effect of producing a specific Cry protein at this level was to provide a considerable degree of protection against damage by susceptible insect larvae. Further confirmation of the effectiveness of these plants in providing protection against insect damage in small-scale field trials resulted in the first *Bt* crops gaining approval for commercial planting in the USA in the mid-1990s.

The success of the *Bt* approach led to the development of *Bt* crops by several of the major biotechnology companies involved in crop protection (Table 6.4). This table demonstrates the point that the specificity of Cry proteins permits the targeting of specific pests by particular transgenes, and that different crops may have different *cry* genes inserted. We will see later that the nature of the *cry* gene construct can be important for the success or failure of a particular transgenic crop, and that some of the lines shown in Table 6.4 have already been discontinued. Indeed, maize and cotton are the only *Bt*

Table 6.4 Commercialisation of *Bt* technology

Company	Trade name	Bt protein	Crops	Insect pests
Monsanto	New-Leaf	Cry3A	Potato	Colorado beetle
Monsanto	Bollgard	Cry1Ac	Cotton	Tobacco budworm, cotton bollworm, pink bollworm
Monsanto	YieldGard	Cry1Ab	Maize	European corn borer
Novartis	YieldGard, Knockout			
Mycogen	NaturGard			
DeKalb	Bt-Xtra	Cry1Ac	Maize	European corn borer
Aventis	StarLink	Cry9C	Maize	European corn borer
Mycogen	Herculex 1	Cry1F	Maize	European corn borer
Pioneer				
Monsanto	pending	Cry3Bb	Maize	Corn rootworm larvae

crops that are currently commercially grown in the USA, since the New-Leaf *Bt* potato has now been discontinued.

CASE STUDY 1 **Resistance of *Bt*-maize to the European corn borer and other pests**

The European corn borer (*Ostrinia nubilalis* or ECB) is a major pest of maize. As the name suggests, the larvae damage maize crops by tunnelling into the central pith of the stalks and ears. However, despite its name, it is not confined to Europe, and causes considerable damage to the maize crop worldwide. Developing an ECB-resistant maize has therefore become one of the main targets of the agricultural biotechnology industry. Table 6.4 shows that several different companies have each produced a *Bt* maize line which is resistant against ECB. The rate of adoption of *Bt*-corn has been rapid in the USA, growing from <5% of the crop acreage in 1996 to ~25% in 2000.

The development of commercial lines involves testing a large number of transformed plants to find the optimal line in terms of the quality of the desired trait, the copy number and stability of integration of the transgene, as well as the stability, level and pattern of transgene expression. Regulatory approval is normally given only for the varieties ultimately descended from a single, fully characterised transformation event.

Three different transformation events with the *cry1Ab* gene (176, Bt11 and Mon810) have been developed by different companies and subsequently licensed to, and marketed by, a number of others. In all three cases, a truncated *cry1Ab* gene with modified codon usage optimised for maize expression was used. The transgene in the Bt11 (YieldGard,

Novartis) and Mon810 (YieldGard, Monsanto) events is regulated by a constitutive promoter. On the other hand, the Bt176 event (Knockout, Novartis; and NaturGard, Mycogen) was produced by bombardment with two separate *cry1Ab* constructs: one controlled by a maize PEP carboxylase promoter (specific for green tissue), and the other under the control of a promoter from a pollen-specific protein kinase gene. Varieties derived from the Bt176 event proved to be less effective against the second brood of ECB than the other two, and have now been phased out. Not surprisingly, the Bt176 event was also shown to produce the Cry1Ab toxin in pollen at much higher levels than the other two, with consequences for the potential toxicity of the pollen to non-pest lepidopterans (see Box 6.3). Other *cry* genes have also been tried, including *cry1Ac* (Bt-Xtra, DeKalb), *cry9C* (StarLink, Aventis) and *cry1F* (Herculex, Mycogen). The particular problems experienced by the StarLink lines are described in Box 6.1.

The current *Bt* maize hybrids have generally provided better control against ECB than could be achieved with a single, well-timed insecticide treatment. Bt11 and Mon810 also provide some control against corn earworm. On the other hand, Cry1F provides additional protection against the fall armyworm and western bean cutworm. However, these Cry1 proteins are not effective against other maize pests such as corn rootworm beetles, corn

BOX 6.1 The rise and fall of StarLink corn

The Aventis Bt *cry9C* maize event, marketed as StarLink, has been the topic of considerable controversy in the USA and Europe. The problem arose when AgrEvo (subsequently Aventis) sought approval for registration of StarLink in the USA. There were doubts raised by the US Environmental Protection Agency about the suitability of the Cry9C protein for human consumption, because it is more stable in acid than other approved Cry proteins (hence more slowly digested in the stomach), and because of a lack of data about allergenicity. AgrEvo accepted a limited and conditional approval for StarLink in order to get the product launched into what was a very competitive marketplace for GM maize. Under the terms of this limited registration, StarLink products could only be used for animal feed, and had to be prevented from contaminating food for human consumption. AgrEvo proceeded with the commercial planting, given that the bulk of US maize is used for animal feed and industrial purposes, whilst carrying out further allergenicity tests on the protein. Unfortunately, in 2000, before StarLink products had been approved for human consumption, traces of StarLink corn were identified in taco shells manufactured by Kraft Foods. There was an immediate recall of these products, followed by costly recalls of other maize products from other suppliers.

The case raises a number of issues around the decision to give conditional approval, and the mechanisms required for ensuring that the growers and processors of GM products complied with the conditions placed upon the biotechnology company that registered the transgenic crop. The case has also demonstrated the requirement to investigate the allergenicity of GM products, and highlighted the need for test procedures to detect the presence of GM materials in non-GM products.

rootworm larvae or spider mites. Maize hybrids containing the *cry3Bb* gene are being tested by Monsanto and may soon appear for the control of corn rootworm larvae.

The problem of insect resistance to *Bt*

One of the problems encountered by the accelerating use of *Bt* technology to provide insect resistance has been the equally rapid appearance of resistant pests. This problem initially attracted widespread attention during the first commercial growing season of the *Bt* cotton crop (Box 6.2).

The mechanism of insect resistance can be related directly to the specific binding involved in the mechanism of action of the Cry proteins. It may only require a small number of significant mutations in the insect gene coding for the receptor protein to greatly reduce the binding of a particular Cry protein. This is one of the consequences of single gene:gene interactions that will be discussed again in the subsequent chapters on disease resistance. Thus, insect pests resistant to the *Bt* crop could, in theory, appear within a few generations, and repeated growing of *Bt* crops in the same area would provide the selective advantage to accelerate the appearance of a resistant pest population. There are a number of strategies for countering the build-up of insect resistance.

One way is to think of additional GM approaches to tackle the problem caused by the first GM approach. For example, it is possible to use more than one transgene—a process called ‘pyramiding’, in which transgenes are successively ‘stacked’ by conventional crosses between different transgenic lines. In order to avoid cross-resistance to two different *Bt* genes, it may ultimately prove better to pyramid *Bt* genes with unrelated resistance genes, such as proteinase inhibitor genes (see Table 6.6), since it is considered highly unlikely that resistance to two completely different genes will arise in the same organism at the same time.

Another GM approach is to further enhance the effectiveness and range of activities of *cry* genes by, for example, domain engineering to produce chimeric proteins (see Box 6.2, for example). However, it is still important to realise that *Bt* and other GM approaches to insect resistance cannot be viewed as a ‘magic bullet’ to permanently eliminate the threat of insect damage. Rather, the GM crops just alter the balance of all the interactions between plant and environment that occur in the field, and the build-up of resistance in the insect population must be managed by good agricultural practices.

The current practice of ‘integrated pest management’ (IPM) takes into consideration the wider context of a pest/crop interaction (natural predators, adjacent plant species, crop rotation and chemical sprays). For example,

BOX 6.2 The development of resistance to *Bt* cotton

The cotton crop in the USA is prone to damage by lepidopteran larvae, particularly cotton bollworm (*Helicoverpa zea*), pink bollworm (*Pectinophora gossypiella*) and tobacco budworm (*Heliothis virescens*). The development of insect-resistant cotton has been one of the major targets for US plant biotechnology.

Following a considerable amount of laboratory and precommercial field testing, a single transformation event, Monsanto 531, was chosen as the event from which cotton varieties marketed by Monsanto as Bollgard are descended. Bollgard lines contain a synthetic *cry1Ac*-like gene driven by an enhanced CaMV 35S promoter (see Chapter 4). The proportion of Bollgard cotton grown in the USA has risen from about 12% acreage in 1996 to ~35% in 2000, and considerable information about its performance in diverse conditions and varieties has accumulated. It has been found that Bollgard cotton consistently provides a high level of protection against tobacco budworm and pink bollworm damage. On the other hand, protection against cotton bollworm is sufficient for moderate infestations, but applications of insecticide spray may be required for heavy infestations. In the first year of commercial-scale planting in 1996, there was an unexpected mid-season outbreak of cotton bollworm, requiring supplemental control with insecticidal sprays. This raised concerns that, in the haste to introduce the crop as widely as possible, there had been insufficient consideration given to the management of *Bt*-resistance in the cotton bollworm population. In particular, if the dose of Cry1Ac was insufficient to kill virtually all the cotton bollworm insects, then this could accelerate the development of a resistant population.

As part of the US commercial licencing agreement with Monsanto, farmers are required to follow the terms of a resistance-management plan. Resistance-management plans are required by the US Environmental Protection Agency (EPA) when a GM variety is approved for commercial growing. The conditional resistance-management plans for Bollgard cotton approved by the EPA in 1995 were updated in 2000 and amended for the 2001 season to allow three options for growers:

1. *80:20 external sprayed refuge*. For every 100 acres of Bollgard cotton, plant 25 acres of non-*Bt* cotton that can be treated with insecticides that control tobacco budworm, cotton bollworm and pink bollworm. This refuge should be planted within 1 mile of the edge of the Bollgard field. (See Figure 6.3)
2. *95:5 external structured unsprayed refuge*. For every 95 acres of Bollgard cotton, plant 5 acres of non-*Bt* cotton. This refuge cannot be treated with any insecticides specified for tobacco budworm, cotton bollworm and pink bollworm. The refuge must be managed similarly to the Bollgard cotton and be planted within 0.5 mile [0.8 km] of the edge of the Bollgard field.
3. *95:5 embedded refuge*. Plant at least 5 acres of non-Bollgard cotton for every 95 acres of Bollgard cotton, embedded as a contiguous block within the Bollgard field. This refuge may be treated with insecticides for the control of tobacco budworm, cotton bollworm and pink bollworm whenever the entire field is treated.

These three options are based on a target refuge size of ~5%. (If it is assumed that insecticide spraying in Option 1 will kill 75% of the insects in the non-*Bt*

BOX 6.2 Continued

cotton, the area of this sprayed refuge is 20% of the total acreage, so the proportion of susceptible insects in the refuge will be ~5% of the total.)

Overall, it is estimated that the amount of Bollgard cotton planted in the USA in 1998 accounted for a reduction in weight of 2 million pounds of insecticide sprayed. The economic benefits to growers have been estimated at nearly \$US50 per acre, with a yield advantage of ~10% over the non-transgenic crop. Thus, there is an incentive to all parties to implement the resistance-management plans to ensure that the pest resistance does not collapse due to the appearance of resistant insects.

A new line of transgenic cotton, Bollgard II, is now under development, via the stacking of a second *Bt* gene (*cry2A*) to provide protection against beet armyworm and fall armyworm. In addition, an improved Cry1Ac protein is under development by replacement of domain III with that of Cry1Fa to increase the spectrum of insects controlled by the chimeric protein.

selective usage of *Bt* crops (rotating *Bt* crops with non-*Bt* crops, avoiding growing different *Bt* crops within the migration distance of pests common to both, etc.) may prevent the build-up of resistance in the insect population, which would otherwise occur where there is a continuous selective advantage created by repeated use of *Bt* crops. One example of pest resistance management favoured for managing the build-up of *Bt* resistance is the ‘high dose/refuge’ approach (Figure 6.3). In this scheme, transgenic crops expressing a high dose of *Bt* protein are grown alongside a smaller ‘refuge’ of non-transgenic crop, or any other plant that the insect pest feeds upon. The high dose is important to ensure that only homozygous resistant insects would be able to tolerate feeding on the GM crop. Homozygous resistant insects are likely to be rare in the early stages of the development of resistance. The refuge ensures the presence of a much larger population of susceptible insects in the vicinity of the *Bt* crop than the homozygous resistant insects. Therefore, most mating events between insects carrying one or two resistance genes will probably be with a homozygous susceptible insect, producing heterozygous resistant offspring, which cannot survive on the crop.

Box 6.2 gives an example of refuge options approved by the US Environment Protection Agency for the growing of Bollgard cotton in 2001. It is obviously critical that the approved refuge strategy is correct and monitored, and that the growers are fully informed and comply with the requirements of the resistance management plan. Any deviation from the planting scheme could create the conditions for insect resistance to become established. Another problem resulting from the refuge strategy is that it may provide additional opportunities for the mixing of transgenic and non-transgenic

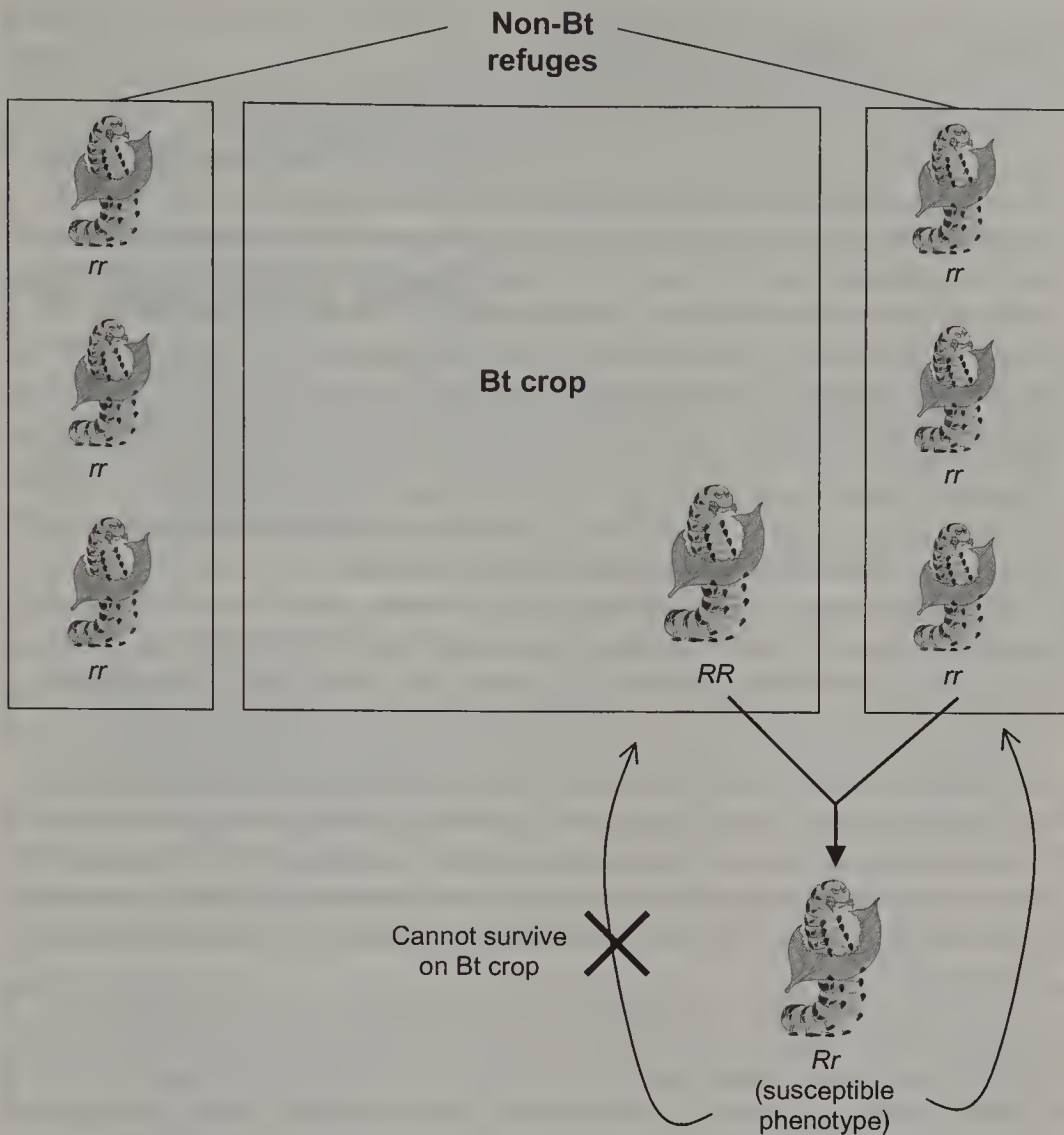


Figure 6.3 High dose/refuge-resistance management scheme. The rationale behind the high dose/refuge strategy is demonstrated. The high dose is required to maintain resistance against heterozygote insects carrying one copy of a resistance gene R . Thus, only homozygote RR insects will be able to survive on the crop. The refuge provides a mechanism for ensuring that the proportion of RR insects is kept very low. The majority of RR insects will mate with the much larger population of rr insects in the refuge, producing Rr -susceptible progeny.

materials. The problems caused by contamination of non-GM foods with GM materials are discussed in Box 6.1, and in Chapter 12.

The environmental impact of *Bt* crops

Whilst the first environmental concerns regarding *Bt* crops related to the build-up of resistance in the pest population, a separate issue was brought to the fore by a report that appeared in 1999 indicating that pollen from *Bt* maize might be toxic to the larvae of the Monarch butterfly. This particular

BOX 6.3 The Monarch butterfly affair

One of the major controversies to hit *Bt* crops was the publication in 1999 of a paper in the journal *Nature* by scientists at Cornell University, showing that the larvae of Monarch butterflies fed on milkweed leaves covered in pollen from *Bt* maize did not grow as well as those on control leaves. The Monarch larvae feed only on milkweed leaves, which are commonly found in cultivated areas and other habitats, so even though they are lepidopterans, and therefore susceptible to the Cry proteins produced in *Bt* maize, there should be no direct risk from them eating the maize plants. However, there is the potential for them to come into contact with pollen from GM maize, and this paper reported an attempt to test the toxicity of GM maize pollen to the larvae. The method was to dust pollen from *Bt* maize on to milkweed in a laboratory and feed these to Monarch caterpillars. The larvae fed on *Bt* pollen-covered leaves ate less and grew more slowly than controls on leaves dusted with non-GM pollen, or with no pollen.

The paper sparked off considerable public concern, given that the Monarch butterfly is one of North America's most colourful and familiar natives. Opponents of GM crops made the most of the report and claimed that it demonstrated there were major environmental problems with *Bt* crops. On the other hand, the proponents of GM technology criticised the methodology and dismissed the paper as just a preliminary finding. A major research collaboration between six groups in the USA and Canada was funded to research the question of *Bt* pollen toxicity in more detail, and to investigate the likelihood of exposure of monarch caterpillars to *Bt* maize pollen under natural conditions. The findings of the six groups were published together in the *Proceedings of the National Academy of Sciences USA* in 2001.

One of the conclusions of the group was that there was a significant difference between the levels of expression of *Bt* in pollen from event Bt176 (see Table 6.4) and other events containing the same *cry1Ab* gene ($1.1\text{--}7.1\ \mu\text{g g}^{-1}$ in Bt176, compared to $0.09\ \mu\text{g g}^{-1}$ in the other two events). Given the pollen-specific promoter used to regulate the expression of one of the *cry1Ab* constructs in Bt176, such a significant difference in protein levels in the pollen is not surprising. There was also a correspondingly marked difference in the toxicity of pollen from different *Bt* maize plants. Pollen from Bt176 had an LD_{50} of the order of $100\text{--}400\ \text{grains cm}^{-2}$ when tested on Monarch butterfly larvae, and $600\ \text{grains cm}^{-2}$ on Black swallowtail butterfly larvae, whilst the other two showed no effect at concentrations $>1600\ \text{grains cm}^{-2}$ (the highest concentration of pollen actually measured in a maize field).

Since Bt176 is no longer available in US maize varieties, the overall conclusion is that the risk to Monarch butterfly populations from current *Bt* maize varieties is low.

report and its consequences are described in Box 6.3, and, to some extent, the particular fears about the Monarch butterfly have now been allayed. The affair does still beg the question of how event Bt176 was approved for commercial growing, given the deliberately high levels of expression of the *Bt* protein in the pollen of these plants. This highlights the need to assess the risks and monitor the environmental impact of GM crops more thoroughly than

perhaps has been the case. However, this point also applies to the positive effects of these crops. For example, to what extent have non-pest insect populations benefited from the reduction in chemical pesticides that have certainly resulted from the adoption of *Bt* cotton?

The 'Copy Nature' strategy

Some of the problems encountered with *Bt* have led some scientists to contrast the single, highly toxic, heterologous gene approach epitomised by *Bt* with what has been termed a 'Copy Nature' approach (Boulter 1993). The strategy involves a rational approach to the development of pest-resistant crops and is characterised by the following steps:

1. *Identification of leads.* The point here is to look in nature for plants that show resistance to insect pests. These leads can be found in the literature, world seed collections or observation in the field. The identification of resistant plants provides the starting point for the discovery of plant genes that could confer resistance to insect damage.

2. *Protein purification.* The next stage of the process is the purification of proteins with insecticidal properties. Proteins are the primary target for this type of screen because they provide the most direct way to isolate corresponding genes. (It should be remembered that plants produce a large number of insecticidal secondary products, but it would be a much longer process to engineer the synthesis of a secondary metabolite in a transgenic plant than of an insecticidal protein. Starting from the discovery of a novel compound, the biosynthetic pathway would need to be determined, the key enzymes identified and the genes isolated in order to produce a transgenic plant.) Partial sequencing of the purified protein is one method for isolating the corresponding gene. Characterisation of the protein may also provide a means of identifying the gene in a model organism such as *Arabidopsis*, and using the known DNA sequence as a probe to screen for the orthologous gene in the plant of interest.

3. *Artificial-diet bioassay.* It is important to determine the activity of the isolated protein against the target insect pests by performing feeding assays in the laboratory. In other words, is the protein going to be an effective deterrent to insect feeding at achievable levels of expression in a transgenic plant?

4. *Mammalian toxicity testing.* Prior to any insertion of the gene into a crop plant, the toxicity of the protein against mammals (and hence potentially humans) should be tested. The point is that it would be a waste of time inserting a gene into plants that subsequently was found to cause concerns about food safety.

5. *Genetic engineering.* It is only at this stage that the isolated gene would be considered valuable for transfer to crop plants. The typical techniques for transformation technique, vector design and construct building are described in Chapters 3 and 4. One particular consideration for pest-resistant constructs

is the choice of promoter. Although it is possible to rely on a strong, constitutive promoter such as the CaMV 35S promoter, there are certain advantages to using either a tissue-specific promoter or a wound inducible promoter. This is also a 'Copy Nature' strategy, since many insect defence genes are produced site-specifically or are induced by insect damage. Expression of the insecticidal protein at the site of insect damage limits the exposure only to the insects causing the damage. It also reduces the biosynthetic burden on the plant, and reduces the amount of transgenic protein in those parts of the plant destined for human consumption. Some of the promoters used to drive the expression of insect resistance genes are shown in Table 6.5.

6. *Selection and testing.* After transformation, the selection of transgenic plants, confirmation of transformation, inheritance of transgenes in T1, T2, etc. generations and testing of expression levels need to be carried out. The effectiveness of the construct then has to be evaluated by insect feeding assays.

7. *Biosafety.* The effect of the transgene on crop yield, insect damage and the wider ecosystem should be properly evaluated in field trials, rather than after several years of commercial planting, as occurred with *Bt* crops (Box 6.3).

The aim of the 'Copy Nature' strategy has been stated as: 'insect pest control which is relatively sustainable and environmentally friendly'. The strategy recognises a complex interplay in biological communities between plants, animals, microbes, the soil and the physical environment. It is not just a case of crop plant vs. insect pest. The strategy also takes account of the fact that host-plant resistance to pests is universal and falls into two major categories:

1. *Horizontal resistance.* This type of trait is usually polygenic, i.e. it results from the cumulative effect of several minor gene characters. This form of resistance does not involve gene-gene matching (as is the case with Cry proteins and their receptors in the insect midgut, for example) and is usually durable.
2. *Vertical resistance.* In contrast, vertical resistance typically arises from one major gene with a high level of expression. The mechanism of resistance may well involve gene-gene matching. The important point to note is that this form of resistance normally occurs in plants to provide a buffer during short-term changes in the level and type of insect damage, and is unlikely to be durable.

CASE STUDY 2 Cowpea trypsin inhibitor

The first stage in the development of cowpea trypsin inhibitor (CpTI) as an anti-insect, pest-control mechanism came from the identification of strains of cowpea growing in Africa that were resistant to attack from a range of insect pests. The isolated insecticidal protein was found to be a trypsin inhibitor. In an artificial-diet bioassay, the protein was shown to be effective against Lepidoptera, Orthoptera and Coleoptera. Further tests showed that this protease inhibitor did not affect mammalian trypsins and was not toxic to mammals. The gene

Table 6.5 Promoters used with insect resistance genes

Promoter	Origin	Expression site	Insecticidal protein	Plant
Mannopine synthase ^{TR}	<i>Agrobacterium</i> Ti plasmid	Most plant tissues	Cry1Ab	Tobacco, potato
Phytohaemagglutinin (PHA-L)	Bean	Seed	a-AI-Pv	Pea, adzuki bean, tobacco
CaMV 35S	Cauliflower mosaic virus	Most plant tissues	Most proteins	Most plants
Sucrose synthase (RSs1)	Rice	Phloem	GNA	Tobacco
Metallothionein-like (MT-L)	Maize	Root preferred	Cry1Ab	Maize
Phosphoenolpyruvate carboxylase (PEPC)	Maize	Green tissue	Cry1Ab	Maize, rice
Pollen-specific	Maize	Pollen	Cry1Ab	Maize
Tryptophan synthase α -subunit (trpA)	Maize	Pith preferred	Cry1Ab	Maize
Ubiquitin-1 (Ubi-1)	Maize	All plant organs	Cry1Ac	Rice
Proteinase inhibitor II (Pot PI-II)	Potato	Wound inducible	Pot PI-II, ipt	Rice, tobacco, tomato
rRNA operon (Prn)	Rice	Chloroplasts	Cry1Ac	Tobacco
Actin-1 (Act-1)	Tobacco	All plant organs	CpTI	Rice
Pathogenesis-related protein-1a (PR-1a)		Chemically induced	Cry1Ab	Tobacco

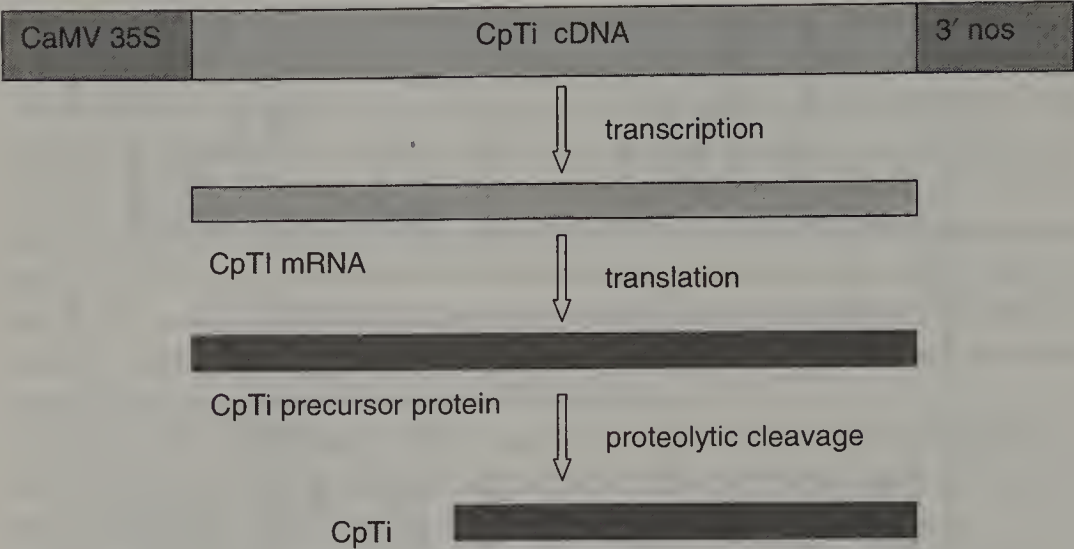


Figure 6.4 Construction of the CpTI expression vector for plant transformation. The CpTI cDNA from cowpea with 35S promoter and 3' nos terminator was inserted into a pBIN19 derivative. Expression of the gene in plants produces a CpTI precursor protein that is subsequently cleaved by specific proteases to yield the active CpTI protein.

was isolated and inserted into the transformation vector pROK 2 (a pBin19 derivative) (Figure 6.4).

For the initial trials, the gene was transferred into tobacco (see Chapter 3), and the transgenic plants produced were tested for their ability to withstand attack from a number of different lepidopteran pests of tobacco. It was found that the transgenic plants were visibly much more resistant to damage under laboratory conditions than were the control plants. There are different ways of quantifying the extent of resistance of a plant to insect damage—either in terms of measuring the amount of damage (loss of weight or leaf area), or assaying the effect of eating the transgenic plant on the insect. Both types of bioassay were used to show that lepidopteran pests (tobacco budworm, cotton bollworm and cotton leaf-worm) fed on the CpTI transgenic tobacco plants did less damage and grew less well than those on control plants (Figure 6.5).

Having tested the concept in the laboratory, it is important to confirm the findings in the field. The results of field tests of CpTI tobacco showed a significant level of protection against damage by lepidopteran pests, and a reduction in the number of feeding larvae (Figure 6.6). CpTI has subsequently been transformed into a range of crop plants including rice, potato, wheat, cotton, strawberry and pigeonpea (Table 6.6). However, none of these crops has yet been produced on a commercial scale.

CpTI is not the only gene that has been isolated from plants and used to engineer insect pest resistance. Plant resistance genes that have been tested include other protease inhibitors, α -amylase inhibitors and lectins. Table 6.6 shows the different genes that have been tried and the range of different transgenic crops that have been produced. However, very few of these examples

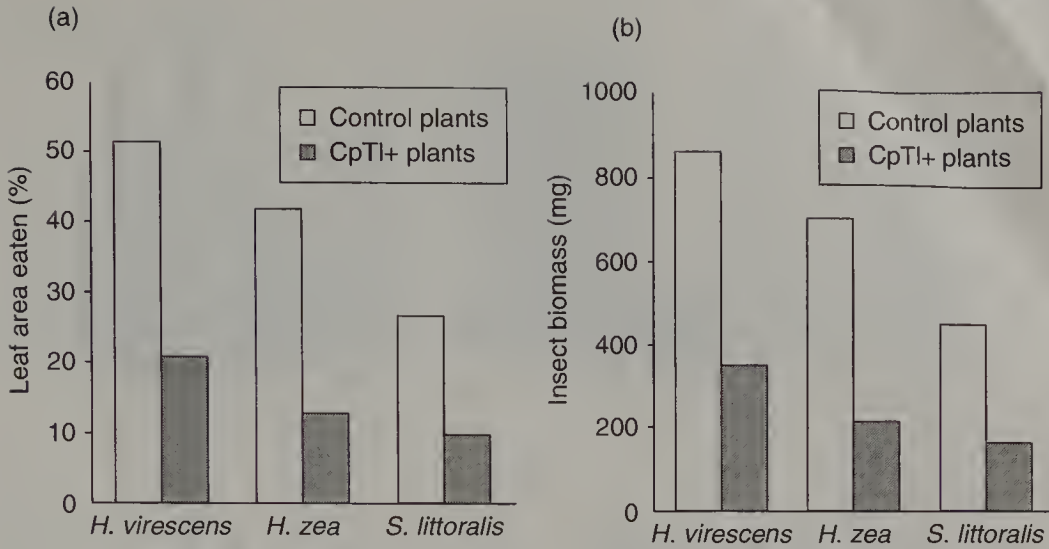


Figure 6.5 Bioassay data for feeding trials of lepidopteran pests on CpTI transgenic tobacco plants. The activity of three different lepidopteran pest larvae (*Heliothis virescens*, *Helicoverpa zea* and *Spodoptera littoralis*) on transgenic tobacco plants carrying the CpTI gene was assayed in the laboratory in two different ways: (a) the extent of damage to the tobacco leaves was determined by measuring the leaf area eaten; (b) the growth of the larvae on the tobacco was determined by measuring insect biomass. In both assays, there is a significant protective effect from the CpTI transgene compared to control plants. (Redrawn with permission from Gatehouse, A. M. R., *et al.* (1992).)

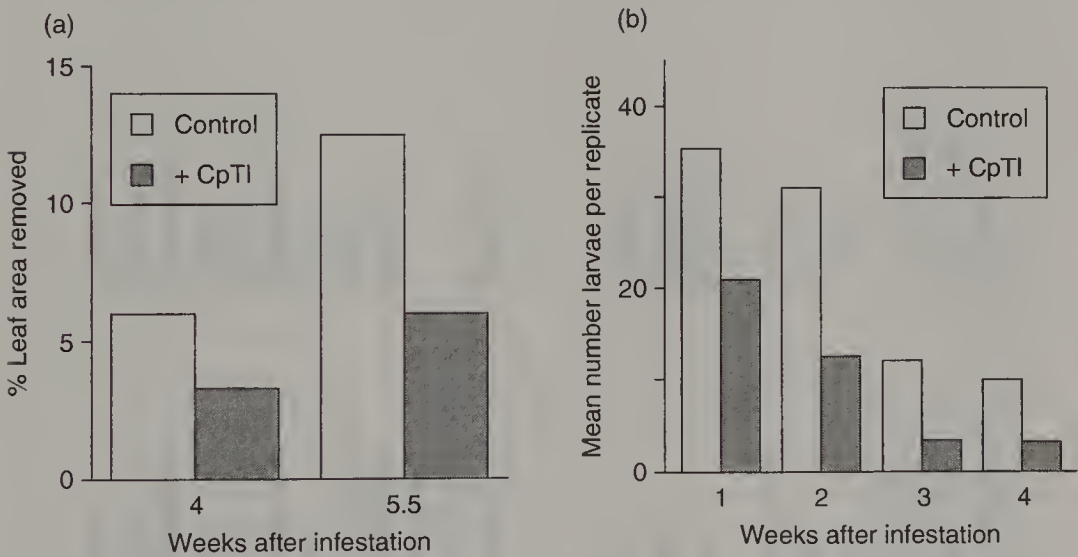


Figure 6.6 The insect resistance of transgenic tobacco carrying the CpTI gene in the field. The ability of transgenic tobacco to resist damage from *Helicoverpa zea* larvae in the field was assessed by two different assays: (a) the leaf area removed was determined; and (b) the number of larvae per plant was counted. There is a significant control of insect damage provided by the CpTI gene. (Redrawn from Gatehouse, A. M. R., *et al.* (1992).)

have progressed from the laboratory to the field trial and none are currently in commercial production. The 'Copy Nature' approach has considerable promise as a complementary method to current pest-management strategies. However, one of the factors which has slowed down the implementation of this strategy is that of food safety.

Table 6.6 Plant insecticidal genes used to engineer pest resistance

Plant gene	Encoded protein	Plant of origin	Target insects	Transformed plants
<i>Protease inhibitors</i>				
C-II	<i>Inhibited protease</i> Serine protease	Soybean	Coleoptera, Lepidoptera	Oilseed rape, poplar, potato, tobacco
CMe	Trypsin	Barley	Lepidoptera	Tobacco
CMTI	Trypsin	Squash	Lepidoptera	Tobacco
CpTI	Trypsin	Cowpea	Coleoptera, Lepidoptera	Apple, lettuce, oilseed rape, potato, rice, strawberry, sunflower, sweet potato, tobacco, tomato, wheat
14K-CI	Bifunctional serine protease and α -amylase	Cereals		Tobacco
MTI-2	Serine protease	Mustard	Lepidoptera	Arabidopsis, tobacco
OC-1	Cysteine protease	Rice	Coleoptera, Homoptera	Oilseed rape, poplar, tobacco
PI-IV	Serine protease	Soybean	Lepidoptera	Potato, tobacco
Pot PI-I	Proteinase	Potato	Lepidoptera, Orthoptera	Petunia, tobacco
Pot PI-II	Proteinase	Potato	Lepidoptera, Orthoptera	Birch, lettuce, rice, tobacco
KTi3, SKTI	Kunitz trypsin	Soybean	Lepidoptera	Potato, tobacco, rice
PI-I	Proteinase	Tomato	Lepidoptera	Alfalfa, tobacco, tomato
PI-II	Proteinase	Tomato	Lepidoptera	Tobacco, tomato

Table 6.6 Continued

Plant gene	Encoded protein	Plant of origin	Target insects	Transformed plants
<i>α-Amylase inhibitors</i>				
a-AI-Pv	α -amylase	Common bean	Coleoptera	Azuki bean, pea, tobacco
WMAL-I	α -amylase	Cereals	Lepidoptera	Tobacco
14K-CI	Bifunctional serine protease and α -amylase	Cereals		Tobacco
<i>Lectins</i>				
GNA	Lectin	Snowdrop	Homoptera, Lepidoptera	Grapevine, oilseed rape, potato, rice, sweet potato, sugarcane, sunflower, tobacco, tobacco
p-lec	Lectin	Pea	Homoptera, Lepidoptera	Potato, tobacco
WGA	Agglutinin	Wheat germ	Lepidoptera, Coleoptera	Maize
Jacalin	Lectin	Jack fruit	Lepidoptera, Coleoptera	Maize
Rice lectin	Lectin	Rice	Lepidoptera, Coleoptera	Maize
<i>Others</i>				
BCH	Chitinase	Bean	Homoptera, Lepidoptera	Potato
Peroxidase	Anionic peroxidase	Tobacco	Lepidoptera, Coleoptera, Homoptera	Sweet gum, tobacco, tomato
Chitinase	Chitinase			Oilseed rape
TDC	Tryptophan decarboxylase	<i>Catharanthus roseus</i>	Homoptera	Tobacco

Insect resistant crops and food safety

The wide range and type of genes shown in Table 6.6 raise understandable questions about their safety for human consumption, given that certain protease inhibitors and lectins are known to have toxic effects in mammals. One of the genes in the table, the snowdrop lectin GNA, was the first such gene to attract considerable attention following the suggestion that potatoes carrying this transgene might be responsible for changes to the gut lining of rats (Box 6.4).

BOX 6.4 The Pusztai affair

In the mid-late 1990s, tins of GM tomatoes and processed foods containing GM soya were clearly labelled and freely available on the shelves of UK supermarkets. Within the space of a year, most supermarkets had followed an apparently huge swing in public opinion against GM foods, and had removed all genetically manipulated products from their shelves. What had happened to cause this shift in the UK public's perception of GM food? One of the most influential factors was the announcement on television by Dr Pusztai of the Rowett Institute, Aberdeen, Scotland, in 1998, that GM potatoes caused intestinal changes when fed to rats.

There was considerable confusion about the nature of his initial claims, which were not immediately published in the peer-reviewed literature. In the storm of publicity that followed, Dr Pusztai was removed from his job. The Royal Society reviewed the unpublished work at the Rowett Institute, and published a report stating that the work was flawed. However, this did little to allay public concerns about food safety and the reliability of scientific advice in the aftermath of the BSE (bovine spongiform encephalopathy) crisis in the UK. By the time the work was subsequently published as a research letter in *The Lancet* in 1999, UK public concerns about GM food (and the environmental impact of GM crops) had reached a peak, and the supermarket shelves were being cleared of GM products.

The report itself concluded that feeding transgenic potatoes carrying the snowdrop lectin GNA (*Galanthus nivalis* agglutinin) gene to rats resulted in cell proliferation of the gastric mucosa. This effect could be attributed to the presence of GNA, since GNA added to normal potatoes had the same effect. However, increases in crypt length in the jejunum and decreases in caecal mucosal thickness were suggested to result from the genetic transformation of the potato, rather than the presence of GNA *per se*. Subsequently, it was suggested that the viral origin of the CaMV 35S promoter could, in some way, be responsible for this transformation-related toxicity. In addition to the Royal Society report, a number of critiques of the study have been published, pointing to problems of experimental design (lack of comparability between test and control diets, overall protein deficiency in all the diets), technique (measurement of crypt length, no actual measurement of cell proliferation) and amount of data (no time courses, no dose-response curves). However, finding a study to be inadequate is not the same

BOX 6.4 Continued

as proving its findings to be wrong, and clearly further investigation of this issue was required. However, to date there have been no convincing reports of a change in toxicity of plant material as a direct result of the act of genetic transformation, or of the integration of the CaMV 35S promoter. Nevertheless, the Pusztai affair catalysed a reappraisal of the adequacy of food safety testing and regulations regarding GM foods. This topic will be covered in more detail in Chapter 12.

Summary

In this chapter, we have looked in detail at the *cry* genes of *Bacillus thuringiensis* and their exploitation for the production of insecticidal proteins in transgenic crops. The current situation in Northern America is that two major *Bt* crops, maize and cotton, have been approved for commercial planting since 1996. The widespread growing of these crops following approval has highlighted a number of issues, some of which have generated considerable controversy. Two major environmental concerns have been considered: the management of insect resistance to *Bt*, and the effects of *Bt* crops on non-pest insects. Two food safety issues have also emerged. The StarLink corn affair raises a series of questions about regulatory approval, allergenicity testing and the contamination of non-GM foods. General issues of food safety testing and assurance are also raised by the use of insecticidal proteins from plants in transgenic crops. The different approach to pest resistance offered by such proteins has been contrasted with *Bt* crops in terms of a 'Copy Nature' strategy.

Further reading

Insect pests

Corn pests BASF—web-link 6.1: [http:// basfagproducts2.com/insectguide/](http://basfagproducts2.com/insectguide/)

Images of plant pests. PestWeb—web-link 6.2:

<http://www.pestweb.com/insectimages/index.cfm>

Insect-resistant crops

Estruch, J. J., Carozzi, N. B., Desai, N., Duck, N. B., Warren, G. W. and Koziel, M. G. (1997). Transgenic plants: an emerging approach to pest control. *Nature Biotechnology*, 15, 137–41.

Schuler, T. H., Poppy, G. M., Kerry, B. R. and Denholm, I. (1998). Insect resistant transgenic plants. *Trends in Biotechnology*, 16, 168–75.

Cry genes

Bacillus thuringiensis Toxin Nomenclature—web-link 6.3:

http://www.biols.susx.ac.uk/home/Neil_Crickmore/Bt/

Crickmore, N., Zeigler, D. R., Feitelson, J., Schnepf, E., Van Rie, J., Lereclus, D., Baum, J. and Dean, D. H. (1998). Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. *Microbiology and Molecular Biology Reviews*, 62, 807–13.

Insect pest susceptibility to Cry proteins

The *Bacillus thuringiensis* Toxin Specificity Database—web-link 6.4:
<http://www.glfsc.cfs.nrcan.gc.ca/Bacillus/btsearch.cfm>

Bt crops

De Maagd, R., Bosch, D. and Stiekma, W. (1999). *Bacillus thuringiensis* toxin-mediated insect resistance in plants. *Trends in Plant Science*, 4, 9–13

Peferoen, M. (1997). Progress and prospects for field use of *Bt* genes in crops. *Trends in Biotechnology*, 15, 173–7.

Transgenic crops: an introduction and resource guide—web-link 6.5:
<http://www.colostate.edu/programs/lifesciences/TransgenicCrops/>

US Environmental Protection Agency—web-link 6.6:
<http://www.epa.gov/pesticides/biopesticides>

Bt cotton

Perlak, F. J., Oppenhuizen, M., Gustafson, K., Voth, R., Sivasupramanian, S., Heering, D., Corey, B., Ihrig, R. A. and Roberts, J. K. (2001). Development and commercial use of Bollgard cotton in the USA—early promises versus today's reality. *Plant Journal*, 27, 489–501.

Union of Concerned Scientists—web-link 6.7:
<http://www.ucsusa.org/food/0biotechnology.html>

Bt corn

Bt Corn & European Corn Borer. University of Minnesota Extension Service—web-link 6.8: <http://www.extension.umn.edu/distribution/cropsystems/DC7055.html>

Peairs, F. B. (2001). Managing corn pests with *Bt* corn: some questions and answers—web-link 6.9: <http://www.colostate.edu/programs/lifesciences/TransgenicCrops/BtQnA.html>

StarLink corn

The StarLink case: issues for the future—web-link 6.10:
<http://www.pewagbiotech.org/research/starlink/starlink.pdf>

Monarch butterfly

Losey, J. E., Rayor, L. S. and Carter, M. E. (1999). Transgenic pollen harms monarch larvae. *Nature*, 399, 214.

Sears, M. K., *et al.* (2001). Impact of *Bt* corn pollen on monarch butterfly populations: a risk assessment. *Proceedings of the National Academy of Sciences USA*, 98, 11937–42.

Shelton, A. M. and Sears, M. K. (2001). The monarch butterfly controversy: scientific interpretations of a phenomenon. *Plant Journal*, 27, 483–8.

Protease inhibitors and other plant genes

Boulter, D. (1993). Insect pest control by copying nature using genetically engineered crops. *Phytochemistry*, 34, 1453–66.

Gatehouse, A. M. R., Boulter, D. and Hilder, V. A. (1992). Potential of plant derived genes in the genetic manipulation of crops for insect resistance. In *Plant genetic manipulation for crop protection* (ed. A. M. R. Gatehouse, V. A. Hilder and D. Boulter), pp. 155–181. C. A. B. International, Wallingford UK

Gatehouse, A. M. R., Shi, Y., Powell, K. S., Brough, C., Hilder, V. A., Hamilton, W. D. O., Newell, C. A., Merryweather, A., Boulter, D. and Gatehouse, J. A. (1993). Approaches to insect resistance using transgenic plants. *Philosophical Transactions of the Royal Society, London*, 342, 279–86.

Lawrence, P. K. and Koundal, K. R. (2002). Plant protease inhibitors in control of phytophagous insects. *Electronic Journal of Biotechnology*, 5, 1–9.

Pusztai affair

Commentary (1999). Genetically modified foods: ‘absurd’ concern or welcome dialogue? *Lancet*, 354, 1314–16.

Ewen, S. W. B. and Pusztai, A. (1999). Effect of diet containing genetically modified potatoes expressing *Galanthus nivalis* lectin on rat small intestine. *Lancet*, 354, 1353–4.

Royal Society (1999). Review of data on possible toxicity of GM potatoes—web-link 6.11: <http://www.royalsoc.ac.uk/>

Royal Society (2002). Genetically modified plants for food use and human health—an update—web-link 6.12: <http://www.royalsoc.ac.uk/>

7

Plant disease resistance

Introduction

With the world's population continuing to rise, the major target of modern agriculture has to be a sustainable increase in yield that keeps pace with the increasing number of mouths to feed. This may be achieved by several approaches, not least by a significant reduction in pre- and postharvest losses due to disease. This chapter will describe some of the progress that has been made by plant biotechnology in combating plant diseases. It will briefly review the types of diseases, the organisms involved and the costs due to these diseases, and then move on to consider the mechanisms of disease resistance found in normal plants and how these are being enhanced using genetic manipulation.

Existing non-GM approaches

At present, many diseases are controlled by the large-scale use of expensive chemical treatments that kill either the pathogen or the vectors that carry them. At present \$US700 000 000 is spent on fungicides in the USA annually. These chemicals may, by their very nature, be detrimental to human health and the environment. An example of this is the fumigant methyl bromide that has been used to control pathogens, insects and nematodes. In the USA, some 30 000 tons of the chemical have been used each year to fumigate the soil prior to planting crops, to fumigate harvested crops during storage and prior to export and import (to kill any potential pathogens). Methyl bromide has been identified as a chemical that contributes to depletion of the ozone layer, and as such its use is now being phased out. This has left the agricultural industry with the task of developing methyl bromide-free farming practices. In this chapter, we deal with some examples of strategies being developed to reduce the use of these chemicals (see also Chapter 6).

It is difficult to quantify the damage done by all the different plant diseases. It has been estimated that crop losses in the USA alone cost some \$US33 000 000 000 (\$33 billion) per annum. In human terms, the loss of a

crop by a subsistence farmer is as damaging, if not more so, than the loss of several thousand acres in the mid-west of the USA. These huge losses are partly due to the use of monocultures which encourage epidemics of pests and diseases. This is exemplified by the case of Southern corn leaf blight caused by the fungus *Bipolaris maydis*. Maize hybrids, the main form of the crop, are produced by breeding female, cytoplasmic male sterility lines (CMS) with male pollinators, under controlled conditions. By the late 1960s, maize varieties developed from the same CMS line (CMS-T) were being used in large areas of the USA. Unfortunately, this germplasm was sensitive to the blight. Large areas of maize crop were decimated by what had previously been a minor form of leaf blight, but was now a very virulent form of the disease. Although this is an extreme case, it highlighted the problems of monocultures. Recent work has begun to explain why dispersed plant regimes are better than monocultures. One particular point, which is discussed later in this chapter in the section on inducible resistance, is that wild or mixed populations have a better chance of surviving attack from a virulent pathogen because they are likely to display polymorphisms in their disease resistance systems. Monocultures are less likely to do so. There are two other important points against monocultures: (1) that different plant types add to the diversity of environmental conditions, making them less beneficial for particular pathogens; and (2) that the increased distance between plant genotypes dilutes the inoculum of a given pathogenic race as it is dispersed between compatible host varieties.

Agricultural seed companies have, with some success, developed disease-resistant lines. In the past this may have taken many years of intensive breeding programs to produce the plants, which may not necessarily have all the desired properties or the required level of resistance. Box 7.1 gives an example of a product produced by conventional breeding.

In the past, breeding strategies relied on basic observation of the phenotype

BOX 7.1 Traditional breeding—a long haul

It can take many years to breed crops with desirable traits. In the year 2000, Cornell University announced the development of a potato, New York 121. This potato is resistant to the Oomycete *Phytophthora infestans* (the cause of late blight), golden nematodes, scab and potato virus Y. The history of this potato dates back 30 years when the breeding programme first started with seeds obtained from the Andes mountains. The Cornell group developed varieties that were adapted to the New York region and had resistance to certain viruses. One of these lines was used as the mother of New York 121. This was crossed with lines that had been developed from seeds obtained from Peru with resistance to golden nematode. It took 9 years to develop and test the progeny New York 121. The potato has one major drawback, it is good for boiling but not for French fries.

in mature plants, but now they are being enhanced by the use of markers and molecular biology techniques. In the major crop plants, restriction fragment length polymorphisms (RFLP) and quantitative trait loci (QTL) maps, augmented with the results of the genome sequencing projects and PCR (polymerase chain reaction) -based molecular markers such as RAPD (random amplified polymorphic DNA), AFLP (amplification fragment length polymorphism), microsatellites and STS (sequence tagged site) provide a battery of very powerful tools that allow genes/traits to be followed during the breeding process. Recently marker-based breeding has been used to 'pyramid' multiple disease resistance genes into plants. These combined approaches offer rapid and cleaner routes to broad-spectrum, durable disease resistance, which could lead to a significant improvement in yield with a concomitant reduction in chemical usage. Marker-assisted breeding strategies have a lot of merit. However, they depend upon the existence of resistance genes in the germplasm and the molecular information being available.

Transformation approaches have significant value in this area. To understand the approaches being used and the opportunities available we need to consider the range of plant diseases, the organisms involved and the natural mechanisms of disease resistance that can be optimised by gene manipulation.

Plant-pathogen interactions

Interactions between plants and microorganisms can be viewed simply as being of four different types:

1. The microorganisms form symbiotic relationships with the host, these include bacteria like *Rhizobium* spp. and the rhizo-fungi.
2. The microorganism may cause disease in the host plant.
3. The host plant may be resistant to the pathogen, and no infection develops.
4. The host plant shows some tolerance to infection, the pathogen is able to grow and replicate but symptoms of infection are minimal.

Within this broad sweep of interactions is a whole range of others that are now beginning to be understood at the molecular level. The pathogens themselves are of two basic types: (1) necrotrophs, i.e. organisms that kill the host and feed on the contents, and which are often associated with toxin production; and (2) biotrophs, i.e. those that require a living host to complete their life cycle. Bacteria and fungi may be of either type, whilst viruses are obligate biotrophs. In the following sections a brief summary of the main types of organisms and their effects will be given.

Prokaryotes

Amongst the prokaryotes, the most economically important pathogenic organisms are members of the phytoplasma group (previously known as mycoplasma-like organisms) and organisms from a small number of bacterial genera. The phytoplasma belong to a group of prokaryotes that lack cell walls and survive as either saprophytes (organisms that live on dead organic matter which they help to break down) or as intracellular pathogens. Diseases include aster yellows, in which the phytoplasma is spread by the aster leafhopper. The pathogen infects over 300 host plant species, including flax, canola, wheat, barley and potato. Killing the leafhopper vector with chemicals controls the disease.

The main groups of bacterial plant pathogens include members of the genus *Agrobacterium*, which cause crown galls and hairy roots by transferring T-DNA into plant cells (see Chapter 3). Some pathogens can enter the host through natural openings such as stomata, but, like many other pathogens, *Agrobacterium* enters the plant through wound sites. Other bacterial pathogens are found amongst the *Corynebacterium*, *Erwinia*, *Pseudomonas*, *Xanthomonas* and the *Streptomyces* genera. Many of the symptoms caused by the different infestations of bacteria are similar: spots, galls, cankers (localised death of an organ), blights (rapid discoloration, wilting, death), soft rots (caused by secreted enzymes that break down plant cell walls and reduce storage time) and wilts (caused by bacteria blocking the vascular tissues). These infections can be localised and have a minimum commercial effect, disfiguring the fruit or leaves, or they can have major effects on yield or commercial value. For instance, the bacterial disease fireblight, caused by *Erwinia* spp. can kill whole fruit trees in a single season.

Fungi and water moulds

The vast majority of cellular pathogens of plants are found amongst the fungi. Whilst not directly affecting human health, the losses of food caused by fungal diseases have had profound effects on humanity. There are over 100 000 known species of fungi. Most are examples of saprophytes, but the fungi also include symbionts, such as the mycorrhiza, and perhaps more than 8000 species that cause plant diseases. Modes of entry include wound sites and natural openings such as stomata, and the enzymes that break down dead organic matter can be used by pathogens to gain entry into living plants by degrading cell wall macromolecules. There are too many fungal pathogens to go into any detail of the diseases they cause or their modes of action, but some specific examples are given in Table 7.1. Included in Table 7.1 is *Phytophthora infestans* (which is a classic example of an Oomycete-group pathogen), the organism responsible for the potato blight that caused the great Irish famine of 1846. Until recently, Oomycetes have been classified as fungi; however, biochemical

and molecular studies indicate that these ‘water moulds’ should now be classified separately. The cell walls of these organisms are composed of cellulosic compounds and glycan, not chitin, and the nuclei within the hypha (filaments) are diploid, not haploid, as in fungi.

Viruses

Biotechnological approaches to plant virus resistance will be dealt with in Chapter 8. Here it is important to note that viruses constitute a major group of pathogens, and that many of the natural host defences discussed below are induced by the cellular and viral pathogens.

Some effort has been made in recent years to classify plant viruses (families, genus, species, etc.). Full details of 950 plant viruses, their hosts and the diseases they cause can be found at Plant Viruses Online (web-link 7.1). The vast majority of plant viruses are single-stranded, positive-sense RNA viruses, but other types such as circular double-stranded DNA viruses exist. Interestingly, one of the examples given in Table 7.1, tungro virus disease, often involves two viruses: rice tungro bacilliform *badnavirus* (a DNA virus) and rice tungro spherical *waikavirus* (an RNA virus). This *badnavirus* causes the tungro symptoms (yellow–orange leaves and stunting), but it depends on the

Table 7.1 Examples of pathogens causing serious economic loss

Pathogens	Major crops infected (disease caused)	Notes
Fungal and water moulds		
<i>Bipolaris maydis</i>	Maize (southern corn leaf blight)	A minor pathogen until a major outbreak occurred in 1969/70 on maize lines with certain cytoplasmic male sterility background
<i>Fusarium</i> spp.	Wheat, barley (scab) Cotton (wilt)	Multibillion dollar losses worldwide Multimillion dollar losses in USA annually
<i>Phytophthora</i> spp.	Potato (blight)	<i>Phytophthora</i> caused the Irish potato famine of 1846
	Cocoa (e.g. black pod)	Loss >450 000 tons annually in Africa/ Brazil/Asia corresponds to \$430 m
	Soybean (e.g. root rot)	15% of Ohio land area is covered with a crop worth \$1b. Losses reach \$120 m annually Total crop losses due to <i>Phytophthora</i> spp of >\$10 b annually

Table 7.1 Continued

Pathogens	Major crops infected (disease caused)	Notes
<i>Puccinia graminis</i>	Wheat (rust)	A major problem controlled with chemicals. Losses could be in the order of billions of dollars annually
<i>Rhizoctonia solani</i>	Many major crops (rots, scabs)	Serious fungal pathogen
Viral		
Barley yellow dwarf <i>luteovirus</i> (BYDV)	Major small-grain cereals	Annual losses in USA estimated at >\$300 m
Cacao swollen shoot <i>badnavirus</i> (CSSV)	Cocoa (swollen shoot)	Losses in Africa of >\$50 m
Cassava African mosaic <i>bigeminivirus</i>	Cassava	Losses attributed to this and other cassava viruses of >\$2 b annually
Rice tungro virus complex	Rice	Losses in SE Asia of >\$1.5 b
Tomato spotted wilt <i>tospovirus</i> (TSWV)	Peanuts (spotted wilt disease)	Losses of \$20–40 m annually in Georgia, USA
Bacterial		
<i>Agrobacterium</i> spp.	Range of dicot plants	See Chapter 3
<i>Pseudomonas syringae</i>	Beans (blight, brown spot)	Florida's climate makes the \$2 b-worth of susceptible crops particularly vulnerable to the disease and losses can run into many millions
<i>Xanthomonas campestris</i>	Tomato, pepper, beans (bacterial spot)	
<i>Xanthomonas fragariae</i>	Strawberries (angular leaf spot)	

Costs are in \$US; m, million; b, billion = 1000 million.

waikavirus for transmission by leafhoppers (the *waikavirus* also increases the severity of the tungro symptoms), whereas the *waikavirus* alone causes mild stunting.

Natural disease resistance pathways—overlap between pests and diseases

It would be wrong to give the impression that plants have no resistance against pathogen attack, for it is clear they do. They do not have an immune system that produces specific cells to attack invading microbes, rather they

have adopted more general defence systems. There is also a lot of overlap between the plant's response to pathogens and plant pests. The general cellular damage caused by both can act as a signal to trigger general defence systems. There are advantages to the plant having such general systems because pests often act as vectors for pathogens. As we shall see below, plants also have systems that respond to specific signals. However, much of the recent molecular work has been done with model plants such as tobacco and *Arabidopsis* and it is possible that the observations are not applicable to all plants. In this section, four different levels of defence will be considered:

- (1) anatomical defences;
- (2) pre-existing protein and chemical protection;
- (3) inducible systems;
- (4) systemic responses.

Anatomical defences

As indicated earlier, many pathogens have to invade plants through wounds. This is because plants have developed morphological and structural systems that preclude pathogen access to living cells (the first line of defence). These can be thick layers of cuticle, bark, waxes, etc. Once this defence is breached then cascades of defence systems come into play.

Pre-existing protein and chemical protection

The second line of defence is made up of a range of antimicrobial proteins produced by the plant during growth and development, these include defensins and defensin-like proteins. The defensin proteins are similar to those found in insects and mammals where they play an important role in defence against infectious agents. Their structure has a conserved three-dimensional folding pattern, which suggests they represent a superfamily of peptides that pre-dates the divergence of plants and animals. Some defensins cause increased branching in fungi, while others simply slow growth. They are frequently associated with seeds at the time of germination, when they may be released into the environment and create a microenvironment around the seed that suppresses fungal growth. Many of the large number of small chemicals made as secondary products may also have antimicrobial properties. These proteins and chemicals may simply deter pest or pathogen growth, or they may actually be toxic to them (see Box 7.2).

Inducible systems

The third level of defence is a switch to counterattack that relies on *de novo* protein synthesis. It would be costly for the plant to have its defence system

BOX 7.2 Secondary products as antimicrobials

Plant cells make huge numbers of small molecules as secondary products. Many of them, such as terpenes, phenolics and alkaloids have antipathogen properties. Presynthesised compounds have been given the title of 'phytoanticipins'. These may be directly toxic or become so following release from, for example, a conjugate state. For instance, plant glycosides are often hydrolysed following insect damage or pathogen invasion that releases vacuolar glycosidases. This enzyme reaction releases aglycones that are toxic to the invader and adjacent cells, producing a local 'fire break'. Antipathogenic chemicals synthesised *de novo*, on infection, are known as 'phytoalexins'. Some of these chemicals can be classified phytoanticipins in one plant and phytoalexins in others. Such is the case for the methylated flavanone sakuranetin, which is constitutively accumulated in black-currant, but is a major inducible antimicrobial in rice.

permanently switched on, so there are mechanisms in place to detect the infection and then turn on the defence system. These mechanisms will be discussed below.

When a pathogen arrives at its host and gains entry to living cells it may induce resistance to infection. This response can be divided into three parts. First, there may be a local response that involves interactions with molecules released by the pathogen (elicitors). The second, recognition-dependent disease resistance, is based upon an interaction between specific proteins produced by the pathogen (the avirulence gene product) and a protein produced by the plant (the resistance gene products), which are shown in Figure 7.1. Both these interactions may lead to a cascade of reactions that invoke the hypersensitive response (HR). The third part of the response is the induction of a systemic resistance (and even the passage of signals to other plants).

Elicitor response

The initial local reaction has several parts. In the first instance there is normally damage to the plant cell wall, causing the release of wall fragments such as pectic-oligomers. These act as signals—endogenous elicitors—which bind to specific receptors, setting off a cascade of reactions that lead to the induction of specific defence genes. These defence genes code for enzymes that synthesise structural components for cell wall thickening (to repair the damage), enzymes of secondary metabolism, lectins (multimeric sugar-binding proteins that agglutinate cells) and many so-called 'pathogenesis-related (PR) proteins' (see Table 7.2). PR proteins include chitinases and β -1,3-glucanases, protease inhibitors, non-specific lipid transfer proteins, ribosomal inhibitor proteins and various antimicrobial proteins. These antimicrobial proteins include defensins such as SN1, which is active against bacterial and fungal pathogens in potato. One important point of relevance to the use of the genes in transgenic

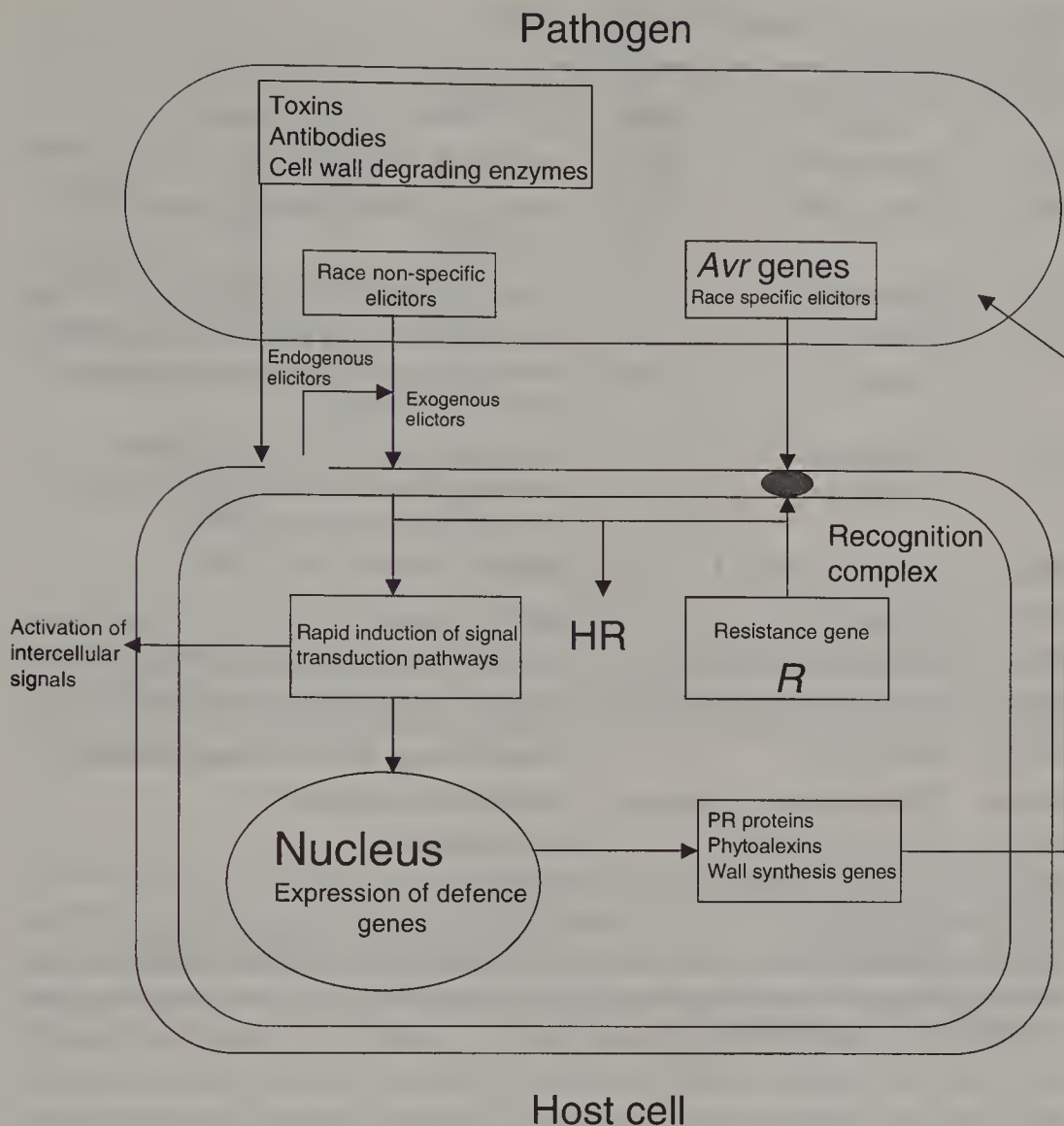


Figure 7.1 Cellular response to pathogen attack. Two of the major plant pathogen interactions are highlighted in this diagram. Host defence systems are induced by endogenous and exogenous elicitors and by the interaction between plant *R* (resistance) genes and the pathogen avirulence gene (*Avr*).

experiments, is that a number of the PR gene products have been identified as latex allergens. This is a family of proteins found in many plant species and responsible for serious medical conditions in many people who show allergenic response to latex products and a number of fruit species. More information can be found in the web links listed in the Further reading section at the end of this chapter.

The effect of synthesising the defence proteins depends, in part, on the pathogen. For instance, if fungal hyphae breach the cells defences then the chitinases and glucanases may cause some degradation of the pathogen's cell wall. This will lead to the production of chitin and β -1,3-glucan oligomers.

Table 7.2 Types of pathogenesis-related proteins (PR)

Family	Type member	Properties
PR-1	Tobacco PR-1a	Antifungal, 14–17 kDa
PR-2	Tobacco PR-2	Class I, II, III endo-beta-1,3-glucanases, 25–35 kDa
PR-3	Tobacco P, Q	Class I, II, and IV endochitinases, 30 kDa
PR-4	Tobacco R	Antifungal, <i>win</i> -like proteins, endochitinase activity, similar to prohevein C-terminal domain, 13–19 kDa
PR-5	Tobacco S	Antifungal, thaumatin-like proteins, osmotins, zeamatin, permeartins, similar to α -amylase/trypsin inhibitors
PR-6	Tomato inhibitor I	Protease inhibitors, 6–13 kDa
PR-7	Tomato P	Endoprotease
PR-8	Cucumber chitinase	Class III chitinases, chitinases/lysozyme
PR-9	Lignin-forming peroxidase	Peroxidases, peroxidase-like proteins
PR-10	Parsley PR-1	Ribonucleases, Bet v 1-related proteins
PR-11	Tobacco class V chitinase	Endochitinase activity
PR-12	Radish Ps-AFP3	Plant defensins
PR-13	Arabidopsis THI2.1	Thionins
PR-14	Barley LTP4	Non-specific lipid transfer proteins

Based on table in web-link 7.3; <http://dmd.nihs.go.jp/latex/defense-e.html> with permission of site owner/maintainer: yagami@nihs.go.jp

These compounds may act as signal molecules (exogenous elicitors) which bind to membrane receptors and re-enforce the induction of the defence systems. The cells may also produce phytoalexins (phenolic compounds or terpenes) that kill any pathogens as well as the cells in the vicinity of the infection, therefore limiting the spread.

Recognition-dependent disease resistance

Using a genetic analysis of the interactions between flax and flax rust, Harold Flor developed the gene-for-gene hypothesis more than 50 years ago. The idea stems from the observation that disease resistance requires two complementary genes. The pathogen carries the avirulence gene (*Avr*) which codes for a protein that is recognised by a specific receptor protein in the plant cell—encoded for by the resistance gene (*R*). This interaction induces the hypersensitive response (HR), which is manifested as a local necrosis that develops through a mitochondrial-associated NADPH-dependent oxidative burst and/

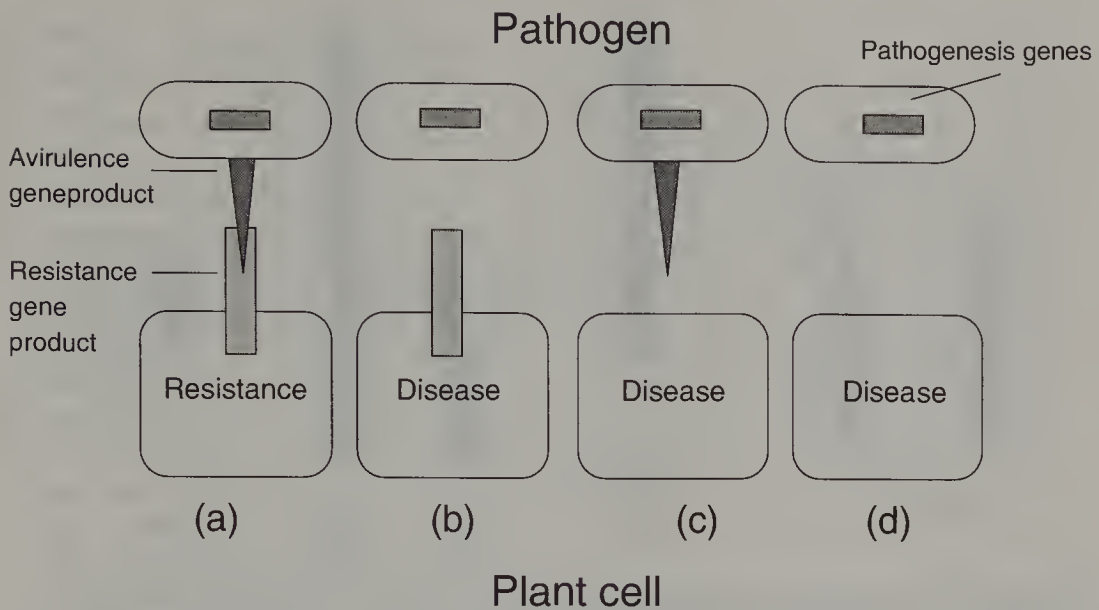


Figure 7.2 HR interactions. In (A) the plant cell contains the resistance gene (R) and the pathogen produces the avirulence gene product (Avr); the plant shows resistant to infection. In (B) the pathogen is not producing the Avr protein so the interaction results in disease. In (C) the host cell is not producing the R protein so disease results. In (D) neither the R protein nor the Avr protein is being produced so disease results.

or the release of phenolics and nitric oxide. The activation of signalling pathways also leads to the induction of many of the pathogenesis-related proteins. The most important feature of this recognition system is that if either of the proteins is absent then the pathogen will cause disease in the plant, this is depicted in Figure 7.2. In panel (a) both proteins are present so the defence systems are activated and the plant is resistant to the pathogen. In panel (b) the avirulence gene is absent, a local HR reaction is not induced and the plant develops the disease. In panel (c) the plant is missing the R gene, and in (d) both proteins are missing. Again the plant develops the disease in both cases.

The gene–gene system is one of the most important ways plants have of switching on resistance systems, and it has been shown to function for interactions between plants and aphids, nematodes, fungi, viruses and bacteria. Although little is known about the Avr proteins, it is postulated they might function as virulence factors, subverting cellular functions through interactions with plant-encoded pathogenicity targets. Many examples of the R proteins have now been identified: the annotation of the *Arabidopsis* genome sequence has indicated there are about 100 R loci distributed through the genome. They fit into five basic structural groups, which are shown in Figure 7.3. In the biggest group, the proteins are characterised by having a nucleotide-binding site and a leucine-rich repeat region (these are termed ‘NB-LRR proteins’). The plant R proteins show extensive similarity with mammalian and insect receptor proteins required for the onset of the innate

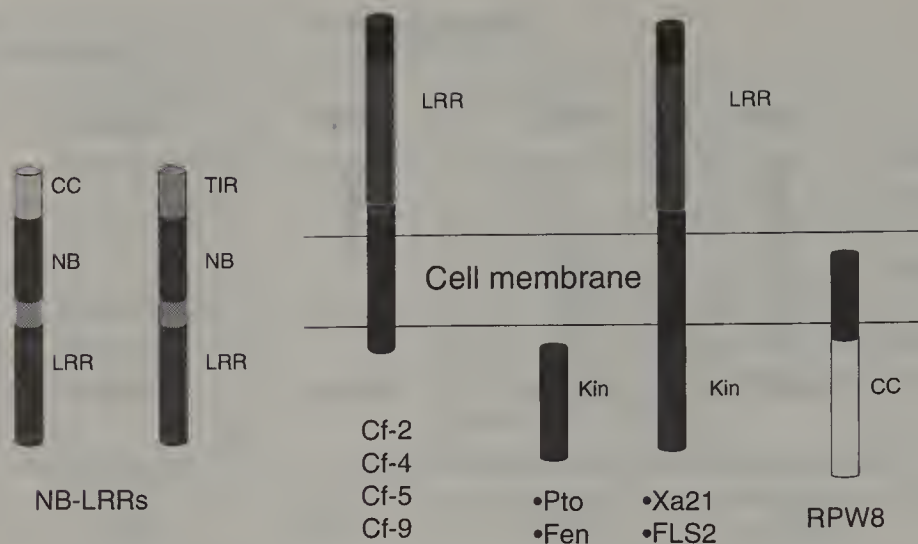


Figure 7.3 A representation of the location and structure of the five main classes of plant disease proteins. NB: LRRs and Pto are believed to be cytoplasmic, although they may be associated with membranes. The other types, (represented by Cf-2 , Xa21, and RPW8, etc., R proteins) are all membrane-associated. (Abbreviations not in the text: TIR, Toll-interleukin-resistance; Kin, kinase; CC, coil-coil domain.) (Redrawn with permission from Dangl, J. L. and Jones, J. D. G. (2001).)

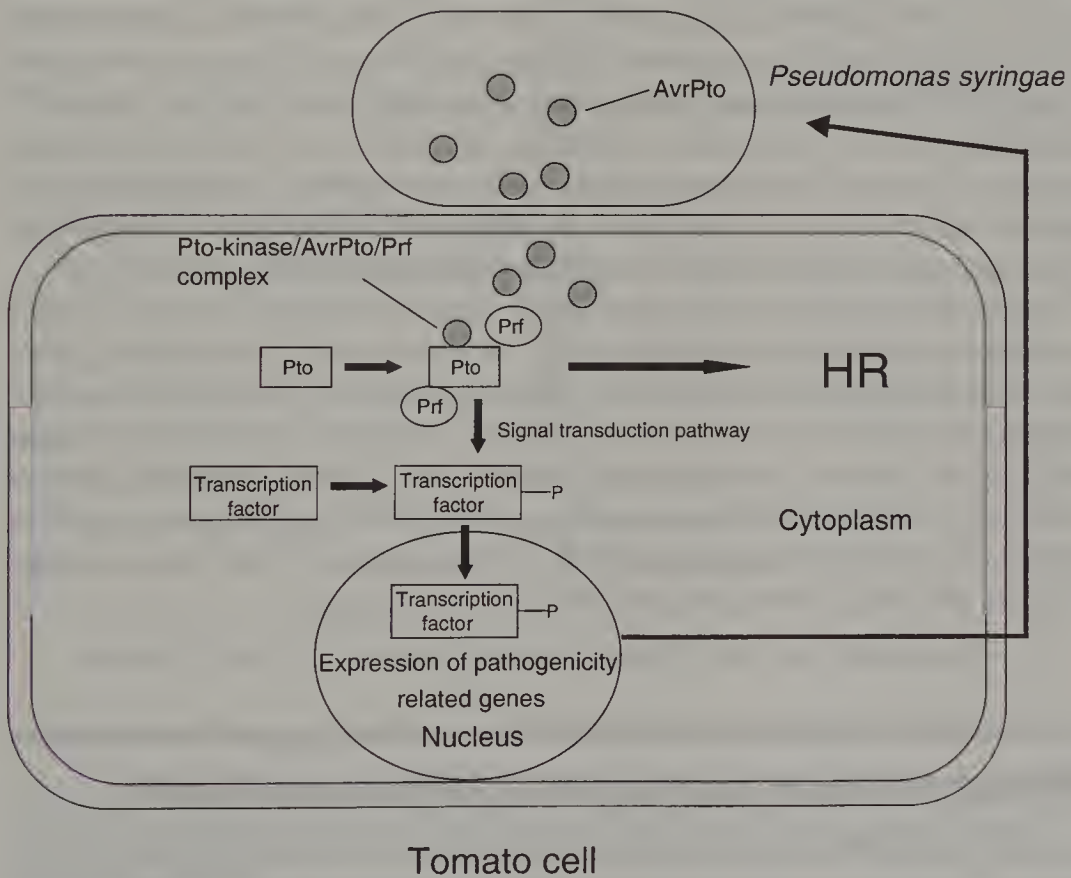
immune response and that are involved in the sensing of pathogen-derived factors called ‘PAMPs’ (for pathogen-associated molecular patterns).

The other four groups are structurally diverse, but examples have been found in which the R proteins also contain an LRR region or, as will be discussed below, the R protein interacts with another protein that contains the LRR domain. It is intriguing to consider that such a small number of R proteins can give the diversity required to deal with the large number of plant pathogens, and how they evolve to deal with evolutionary changes in the Avr protein. One factor seeming to favour plant defence mechanisms is that Avr proteins are key factors in the biology of the pathogen, and as such that they do not evolve at high rates. There is also evidence indicating that to meet with the changes that do occur, some variability in the plant R proteins comes from the fact that the LRR domain does mutate, within a highly variable region, and that mutations encoding new amino acids are selected for. The R-gene loci are very polymorphic in wild populations, so with each R-gene allele being present at a low frequency there is limited selection for virulence in the pathogen population. In conventional agriculture, with the use of monocultures, the balance between polymorphisms within the resistance genes is disrupted therefore making the crops potentially more susceptible to virulent infections. The mechanisms by which interactions between R and Avr bring about the HR response and induce the expression of the defence proteins are now being uncovered, and several of these points are discussed in Box 7.3.

BOX 7.3 AvrPto–Pto interaction

The Pto protein (an R protein) in tomato is a serine/threonine kinase that confers resistance to strains of the *Pseudomonas* genus carrying the *avrPto* gene. This is an interesting example for three reasons. First, the R protein is cytoplasmic, which suggests that the bacterial AvrPto protein must be introduced into the plant cell. The trafficking of pathogen proteins into the host cell is quite a common feature of pathogen attack. Second, Pto is a kinase, a common and direct mechanism for the initiation of a signal-transduction cascade. The third reason is that Pto lacks the LRR recognition domain. This domain is found in most of the R proteins identified to date. So how does Pto function? It has been shown that this reaction depends on a second plant protein Prf that does contain an NB–LRR domain. The Pto protein forms a complex with AvrPto and Prf, which initiates the HR response and a signal-transduction cascade (via the kinase activity) that leads to the expression of various defence proteins.

A MAP (mitogen-activated protein) kinase (MAPK) cascade has recently been identified for an *Arabidopsis* receptor protein FLS2, which recognises bacterial receptor proteins. The cascade includes a classic series of kinases (MAPKKKK, a



The Pto-kinase pathway. When *Pseudomonas syringae* producing the AvrPto protein infects a tomato plant, the protein can form a complex with the Pto-kinase and the Prf protein. The complex activates the HR response and a signal-transduction pathway that induces the expression of a range of pathogenicity related genes.

BOX 7.3 Continued

MAPKKK and a MAPKK) and a transcription factor involved in regulating the expression of the defence genes.

Several models have been proposed to account for these observations. One is that the Avr protein interferes with the R-protein's function in an innate non-specific defence system. In this scenario, the role of the NB-LRR domain (or in the Pto system, the Prf protein) is to recognise the malfunction of the innate defence system and guard the R protein by activating the cell's inducible defences, this is known as the 'Guard hypothesis'.

Systemic responses

The induction of local defence pathways may lead to the induction of inter-cellular signals that produce a systemic response, termed 'systemic acquired resistance' (SAR). Both avirulent and virulent pathogens may result in the induction of SAR, although it is usually a slower process in the case of virulent pathogens. SAR has two phases: the initiation phase and the maintenance phase. In the initiation phase, cells at the foci of the infection release signal molecules, typically salicylic acid (SA), into the phloem. These are transported to target cells in other parts of the plant where SAR genes (such as certain PR proteins, etc.) are expressed, thus giving the plant some level of resistance against infection. In the longer term, a maintenance phase is reached that may last for weeks or even the full life of the plant, in which there is a quasi steady-state resistance against virulent pathogens (see Figure 7.4).

Other states of disease resistance, such as induced systemic resistance (ISR), which are independent of salicylic acid, are also induced by avirulent pathogens through pathways that seem to include ethylene or jasmonic acid as messengers (see Figure 7.5). These may result in the induction of different classes of defence proteins. Evidence now suggests that signals carried by volatile chemicals, such as methyl-jasmonate, can travel between plants to warn of an attack. Many of the proteins involved in these pathways have been identified and are providing further insights into disease resistance.

Biotechnological approaches to disease resistance

A great deal of progress has been made in converting the fundamental information obtained on resistance into strategies for enhancing plant defence systems. As one would imagine, these range from strategies in which individual pathogens are targeted to approaches designed to enhance general resistance systems to inhibit a range of pathogens. Many have been tested on model systems, some have been introduced into crop plants, but none (except for virus resistance that is dealt with in Chapter 8) have so far been commercialised.

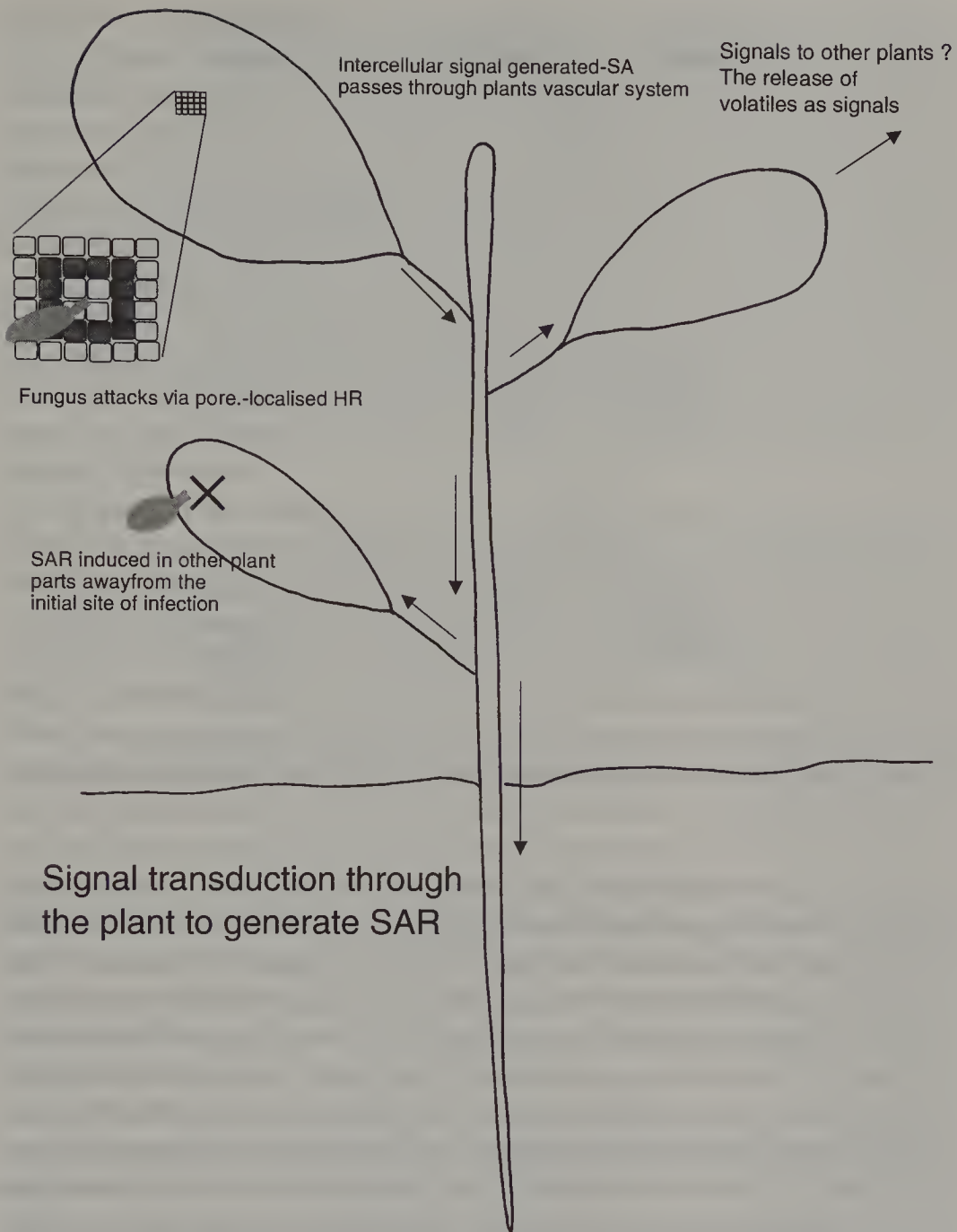


Figure 7.4 Induction of systemic acquired resistance (SAR). (SA, salicylic acid.)

Several of the systems used to genetically engineer plant resistance to bacterial and fungal pathogens will be considered in more detail in the following section of this chapter.

Protection against fungal pathogens

Initial studies for the enhancement of resistance have been based upon the plant's own defence systems, in particular the transformation of plants with genes encoding for PR proteins. To combat fungal pathogens, many plants

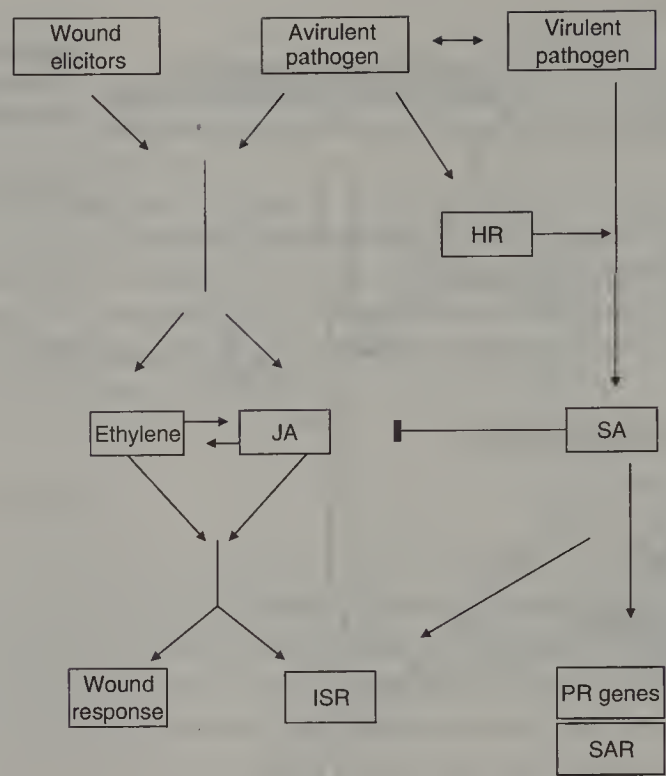


Figure 7.5 Network of disease response pathways. (JA, jasmonic acid; SA, salicylic acid; ISR, induced systemic resistance; SAR, systemic acquired resistance.) (Based with permission on Dong, X. (1998).)

have been transformed with genes that code for chitinase and glucanase enzymes. These degrade polymers in the cell walls of many, but not all, fungi, without affecting the host plant or any animals. The genes for these enzymes have been isolated from a number of sources, including plants (rice, barley), bacteria (*Serratia marcescens*) and even fungi (*Trichoderma harzianum*). The choice of enzyme used has partly depended on its availability and on its effectiveness against the pathogen being addressed, since successful pathogens are already likely to have resistance against plant systems. The 35S promoter (Chapter 4) has often been used to express these genes. Constitutive expression of the defence protein is designed to provide a barrier against the initial attack and not allow the pathogen to establish, as may be the case during the induction period of the host defence system. Wound inducible promoters (Chapter 4, Box 4.1), such as the potato prp1–1 promoter have also been used, as they mediate rapid and localised expression in response to pathogen attack. This type of promoter restricts gene expression to where the protein is required and may reduce any potential hazards associated with the proteins. As gene expression is only turned on during the infection therefore limiting gene expression, it means there is no serious drain on the plant’s biosynthetic capability nor a serious amount of yield reduction (yield drag).

The efficiency of these systems has been tested in many plants, and in many

cases some level of resistance/tolerance to specific fungal pathogens has been demonstrated. One of the earliest demonstrations of tolerance was exhibited by field-grown transgenic tobacco lines, containing a chitinase gene from the bacterium *Serratia marcescens*. Both intracellular and extracellular forms of the enzyme (glycosylated or unglycosylated forms) gave the plants increased tolerance to the pathogen *Rhizoctonia solani*. In another example, transgenic cucumber plants showed a range of responses to the grey mould (*Botrytis cinerea*) when transformed with a rice chitinase sequence (RCC2) driven by a 35S promoter. Interestingly, high-level resistance was observed in a number of lines where spread of the disease was blocked completely. However, the block in fungal attack occurred at different stages in the infection process. Some lines prevented hyphal penetration, whilst others allowed penetration but restricted the spread.

Perhaps the most ambitious work with chitinase genes has been carried out with fruit trees. Apple tree plants expressing either an exochitinase and/or an endochitinase isolated from the fungus *Trichoderma harzianum* have been produced using *Agrobacterium*-mediated transformation. A positive correlation was found to exist between resistance to scab produced by the fungus *Venturia inaequalis* and the level of expression of the proteins. If both genes were introduced into plants the proteins acted synergistically. One drawback was that expression of the endochitinase reduced plant growth. One important question that remains to be answered is how long the resistance will be maintained. A further disadvantage of the chitinase approach is that major Oomycete pathogens such as *Phytophthora* spp. do not have chitin in their cell walls. To overcome these difficulties, combinations of PR proteins have been used (described later) and other antimicrobial proteins have been tested. This is a sensible strategy anyway, as the use of more than one gene will help to prevent the development of resistance and provide longer term protection.

Glucanases, which are also one of the groups of PR proteins, have been used as a tool to enhance resistance to fungal infection. There are different forms of the enzyme with different substrate specificities, but one of the most widely used is a class-II β -1,3-glucanase from barley. When expressed in transgenic tobacco plants under control of the 35S promoter, the gene provided increased protection against the soil-borne fungal pathogen *Rhizoctonia solani*. In plants also expressing chitinase transgenes, there is a synergistic effect and resistance is further enhanced.

A third type of protein, ribosome inactivating proteins (RIP), has also been used as a defence against fungal infections. RIPs are enzymes that remove an adenine residue from a specific site in the large rRNA of eukaryote and prokaryote ribosomes, thereby inhibiting protein synthesis. The specificities of these proteins vary, but the critical point is that certain examples do not inhibit plant ribosomes. Type-1 barley RIP, expressed constitutively, has been used to provide resistance to fungal attack and has been tested in tandem with

the chitinase gene. Again, a synergistic effect was found against *Rhizoctonia solani* in transgenic tobacco plants.

Transgenic crops for food safety

Recently, transgenic approaches have been used to combat the effects of fungal infection rather than the infection itself. *Fusarium verticillioides* causes fusarium ear mould in maize and produces fumonisin mycotoxins, which are acutely toxic to certain livestock and are also carcinogenic. It has not yet been possible to incorporate ear-mould resistance into the maize crop, but it has been possible to incorporate either a fungal esterase gene or a fungal amine oxidase gene into maize and detoxify the maize to some extent.

Antimicrobial proteins

Many other genes that code for proteins with both antifungal and antibacterial activity have been introduced into plants with varying success. The proteins used have included lysozyme (this enzyme degrades chitin as well as peptidoglycan) which delayed fungal infection for a short time. Potato plants expressing lysozyme have been studied with respect to their resistance to *Erwinia carotovora*. Various thionin proteins (see Table 7.2) have been introduced into plants with mixed success, barley α -thionin expressed in transgenic tobacco was shown to increase resistance to *Pseudomonas syringae*, but other attempts did not enhance resistance.

An exciting area of development is in the area of defensins. These proteins are found in all living cells. They have various properties, but are small antimicrobial peptides (26–50 amino acid residues) that are ancient mediators of innate defences. Defensins display lytic activity by binding within microbial plasma membranes. This sort of attack may prove difficult for the pathogen to develop resistance to. As the integrity of the membrane is critical, it is unlikely that changes in the structure of the membrane could occur to prevent defensin infiltration. An alfalfa defensin gene, *alfAFP*, has been introduced into potatoes and its performance monitored in the field. The defensin was processed and secreted into extracellular spaces of leaves and roots in transgenic plants. In the greenhouse and in several field trials, the transgenic plants were shown to have a significant resistance toward the fungus *Verticillium dahlia*. The level of resistance was comparable to that found with a naturally resistant potato. However, they were also expected to be resistant to *Alternaria solani*, but this proved not to be the case.

An artificial defensin gene has been made that incorporates the *cecropin* and *melittin* genes, obtained from a giant silk moth and bee, respectively. This gene has also been introduced into potatoes, and transgenic plants challenged in tissue culture conditions with the bacterium *Erwinia carotovora*. Tubers were then tested for *E. carotovora* soft rot—during the course of the experi-

ment the transgenic tubers remained firm, whereas control plants rotted. This type of study also shows the power of transgenic approaches where genes with the required characters can be moved across kingdoms.

Some of the most impressive studies in this area have been carried out to combat the bacterial disease fireblight. Many of the most important commercial apple cultivars and rootstocks grown in Europe are sensitive to fireblight. This disease is caused by *Erwinia amylovora* which, once established, can kill a tree within a growing season. It is also a very serious disease of pears. A number of independent groups are addressing this problem by producing transgenic trees containing the genes for antimicrobial proteins. One strategy used has been the 'lytic' approach, to produce transgenic fruit trees (via *Agrobacterium*-mediated transformation) containing the gene for T4 lysozyme and genes for the insect antimicrobial proteins attacin E and cecropin. A second approach being used involves the inclusion of a bovine lactoferrin gene to compete with microbial siderophores and a depolymerase gene whose product is capable of degrading specific exopolysaccharides.

These experiments are at various stages, but field trials have been carried out with apple plants. It is claimed that plants expressing attacin E under the control of a constitutive promoter show increased field resistance to fireblight. Experiments were carried out in 1998 in which 2- and 3-year-old plants were inoculated with *E. amylovora* and the percentage of the current season's shoot-length blighted (SLB) tissue determined. Many of the transgenic lines showed less blight than the control plants. One line in particular had only 5% SLB compared with 56% in non-transgenic controls.

The examples given above highlight the use of antimicrobial proteins to interact directly with pathogens and stop infection. One can ask, though, will they ever be commercialised? It is likely some will, but so far these approaches have been disappointing, as even when using non-specific enzymes the spectrum of resistance achieved has been limited. This may be a problem, as one view is that a broad-spectrum of resistance in the plant lines is more likely to be acceptable to the farmer than if only single pathogens are being dealt with. The alternative view, as exemplified by the approach to fireblight, is that some diseases are so important that strategies that deal with these alone are acceptable. Other strategies have been tried in order to introduce broad-spectrum resistance; for instance, strategies to switch on resistance pathways have been developed.

Induction of HR and SAR in transgenic plants

From the information given above, it is clear that the most obvious approach for switching on a general resistance pathway would be to use the gene-to-gene interaction system. One possible implication of the Guard hypothesis (see Box 7.3) is that the R protein has a function in the cell that relates to the general resistance mechanism. This would suggest that an increase in the

production of the R protein will bring about an increase in the expression of the resistance systems. This hypothesis has been tested in a series of experiments in which the *pto* gene, under the control of the 35S promoter, has been introduced into tomato plants.

Close examination of three transgenic lines indicated that overexpression of the Pto protein activated a defence response in the absence of the Pto–AvrPto interaction. Cell death due to an HR response was highly localised to palisade mesophyll cells. The mesophyll cells also showed accumulation of autofluorescent compounds, callose deposition and lignification. Leaves also exhibited salicylic acid accumulation and increased expression of pathogenesis-related genes. The important outcome of this work is that this constitutive expression of the resistance systems is not species-specific. Plants that normally would be sensitive to *Pseudomonas syringae* pv *tomato*, without *avrPto*, were shown to be resistant. They were also tested with *Xanthomonas campestris* pv *vesicatoria*, a leaf pathogen that causes bacterial spot disease and *Cladosporium fulvum*, which causes leaf-mould disease. The transgenic plants were also more resistant to these pathogens.

Another approach that has been adopted to switch on the HR-related defence system has been devised, in which a *Phytophthora cryptogea* gene that codes for a highly active elicitor termed ‘cryptogein’ has been fused to a pathogen-inducible promoter from tobacco and introduced into transgenic tobacco. Under non-induced conditions (i.e. in the absence of the pathogen) no elicitor is made. On infection by the virulent fungus *Phytophthora parasitica* var *nicotianae* the cryptogein gene was expressed, and the elicitor-induced defence gene expression was detected around the infection site. Localised necrosis, similar to what would have been produced by non-virulent pathogens, occurred around the infection site. The plants also exhibited resistance to other fungi unrelated to the *Phytophthora* genus. Broad-spectrum resistance, at least to fungi, has been produced without the constitutive expression of the transgene.

Summary

This chapter has introduced the different types of plant pathogens and their effects. It is clear they cause major financial losses to the agricultural industry, despite the plants having defence mechanisms against them. The different levels of defence have been identified, but in western agriculture the resistance systems are inhibited by the nature of monocultures which limit the natural polymorphisms that occur in wild populations. Over the last 50 years or so, a great deal of information has been gathered on the mechanisms plants use to combat disease. This has allowed the development of a range of different genetic manipulation strategies to be developed.

Initially, the first attempts at enhancing resistance have involved the introduction of antifungal or antibacterial genes that are either pathogenesis-related (PR) proteins or

proteins identified as having antimicrobial properties. These proteins gave a limited spectrum of resistance. However, when attempting to deal with a very serious pathogen such as *Erwinia amylovora*, which causes fireblight, this is probably acceptable. To develop strategies for durable, broad-ranging systems of resistance, transgenic plants have been created which contain genes that switch on signal-transduction cascades. These approaches have been shown to give some benefit to the plant, but it may be at a cost. Switching on defence pathways constitutively can, because of the metabolic burden, lead to a yield loss. So far no plants enhanced for resistance to fungi or bacteria have been commercialised in the West. This may be due to a number of factors. Some of these are intrinsic to the technology, such as the limited range of resistance and yield drag. Some of the reasons are extrinsic, such as the anti-GM campaign, or the absence of drive from industry to commercialise the first generation of resistant plants. The molecular techniques being applied to traditional breeding strategies are now allowing resistant genes to be pyramided into the major crop plants. In this postgenomic era, the sequencing of the *Arabidopsis* and rice genomes (Chapter 1) has, in the first instance, allowed large numbers of disease-resistance genes to be identified, and transcription profiling (transcriptomics was briefly looked at in Chapter 1) will allow their interactions to be studied in detail. This will benefit both molecular marker-based breeding strategies and transgenic approaches. Breeding can only be successful if the resistance genes are already in the germplasm. Transgenic approaches will allow the transfer of resistance genes between species. One long-term target in the postgenomic era is to artificially re-establish polymorphisms by gene pyramiding numbers of resistance genes from resistant plants into sensitive crop species.

Further reading

Groups of pathogens

Plant Viruses Online—web-link 7.1: <http://image.fs.uidaho.edu/vide/refs.htm{CWS}>

Non-GM approaches to disease control

Chrispeels, M. J. and Sadava, D. E. (1994). *Plants, genes, and agriculture*. Publishers Jones and Bartlett, Boston.

Dixon, R. A. (2001). Natural products and disease resistance. *Nature*, 411, 843–7.

Disease resistance

Asai, T., Tena, G., Plotnikova, J., Willmann, M. R., chiu, W.-L., Gomez-Gomez, L., Boller, T., Ausubel, F. M. and Sheen, J. (2002). MAP kinase signalling cascade in *Arabidopsis* innate immunity. *Nature*, 415, 977–83.

Benhamou, N. (1996). Elicitor-induced plant defence pathways. *Trends in Plant Science*, 1, 233–40.

Dangl, J. L. and Jones, J. G. (2001). Plant pathogens and integrated defence responses to infection. *Nature*, 411, 826–33.

Dong, X. (1998). SA, JA, ethylene, and disease resistance in plants. *Current Opinion in Plant Biology*, 1, 316–23.

- Farmer, E. E. (2001). Surface-to-air signals. *Nature*, 411, 854–6.
- Feys, B. J. and Parker, J. E. (2000). Interplay of signalling pathways in plant disease resistance. *Trends in Genetics*, 16, 449–55.
- Lam, E., Kato, N. and Lawton, M. (2001). Programmed cell death, mitochondria and plant hypersensitive response. *Nature*, 411, 848–53.
- PR proteins as latex allergens—web-link 7.2: <http://dmd.nihs.go.jp/latex/index-e.html>
- van der Biezen, E. A. (2001). Quest for antimicrobial genes to engineer disease resistant crops. *Trends in Plant Science*, 6, 89–91.

Engineering disease-resistant plants

- Salmeron, J. M. and Vernooij, B. (1998). Transgenic approaches to microbial disease resistance in crop plants. *Current Opinion in Biology*, 1, 347–52.
- Stuiver, M. H. and Custers, J. H. H. V. (2001). Engineering disease resistant plants. *Nature*, 411, 865–8.

8

Reducing the effects of viral disease

Introduction

Virus infections of plants can cause cell necrosis, hypoplasia (retarded cell division and growth) or hyperplasia (excessive cell division or cell expansion) and these effects can lead to plants showing symptoms such as growth retardation, distortion, mosaic patterning of leaves, yellowing and wilting. In many cases the overall result of such infections is of minor consequence. Often viruses are endemic to a region and cause only moderate crop losses. However, as we have seen in the preceding chapter plant viruses can be a major problem to agriculture, with severe epidemics periodically causing losses adding up to billions of dollars. In this chapter there will be a brief introduction to the different types of viruses, some of their properties and the methods that have previously been used in agriculture to prevent the spread of virus infections. This will be followed by an in-depth discussion of several biotechnologically based strategies that have been developed to reduce the effects of viral infections in plants. The main approach used has been pathogen-derived resistance (PDR). This is where pathogen sequences are introduced into the host on the basis that expression at an inappropriate time, or level, or in an inappropriate form, would disrupt the ability of the virus to sustain an infection and, as will be shown later, the approach has been successful. This line of research has uncovered many interesting aspects of biology, some of which (like gene silencing) will be touched on in this chapter.

Types of plant viruses

There has been an initiative, instigated by the International Committee on the Taxonomy of Viruses, to formalise the viruses into families and genera and to standardise the nomenclature used. The names and acronyms used for viruses in this chapter are taken from the VIDE database, where this information is held (see Plant Viruses Online: web-link 7.1).

Table 8.1 Genome structure of plant viruses

Genome structure	Families and genera	Notes
RNA Single-stranded Positive-sense (act as mRNA directly)	Families: <i>Bromoviridae</i> , <i>Comoviridae</i> , <i>Potyviridae</i> , <i>Sequiviridae</i> and <i>Tombusviridae</i> Examples of unassigned genera: <i>luteo-</i> , <i>potex-</i> , <i>tobamo-</i> and <i>tobraviruses</i>	70% of known plant viruses. Both non-segmented and segmented (two or more RNAs in different virus particles)
RNA Single-stranded Negative-sense (RNA needs to be copied before it can act as mRNA)	Family: <i>Bunyaviridae</i> , genus: <i>tospovirus</i> Family: <i>Rhabdoviridae</i> , genus: <i>rhabdovirus</i> Unassigned genus: <i>tenuivirus</i>	<i>Bunyaviridae</i> possess a lipidic envelope in addition to their nucleocapsid
RNA Double-stranded	Family: <i>Reoviridae</i> , genera: <i>fiivirus</i> , <i>phytoreovirus</i> , <i>oryzavirus</i> Family: <i>Partitiviridae</i> , genera: <i>alpha-</i> and <i>betacryptovirus</i>	The plant members of the <i>Reoviridae</i> family have a genome consisting of 10–12 segments of RNA. Each has 1 ORF which produces a protein
DNA Double-stranded	Unassigned genera: <i>caulimovirus</i> (e.g. cauliflower mosaic <i>caulimovirus</i>) and <i>badnavirus</i>	The only plant viruses of this group are the caulimoviruses. Their genome consists of one ds-circular DNA molecule with specific ss discontinuities in both strands. It codes for 6 or 8 ORFs located on one strand only. DNA replication occurs by a process of reverse transcription (i.e. via an RNA intermediate) similar to that of animal retroviruses
DNA Single-stranded	Family: <i>Geminiviridae</i> , includes bigeminiviruses	3 groups, the only plant viruses possessing either one or two molecules of ss-genomic DNA. DNA replication occurs via ds-DNA intermediates. ORFs located on both the viral strand and its complement

ORF, open reading frame; ss, single-stranded; ds, double-stranded.

It is impossible to discuss all the different types of virus in any detail, but Table 8.1 highlights the main families and unassigned genera. The virus particle is composed of a nucleic acid genome (five types are indicated in Table 8.1), which can be a single or multipartite structures (segmented). The genome is surrounded by the capsid, or protein coat, forming the nucleocapsid. Plant viruses come in a range of morphologies, from rod shapes to isometric particles, and some have membrane envelopes around the nucleocapsid. The virus genomes code for similar ‘core’ proteins—the capsid or coat protein(s), replication related proteins (such as polymerases and helicases), capping enzymes/proteins and movement proteins (which facilitate the spread of the viruses). The functions of these and other virus proteins will be discussed in various sections of the chapter. Some viruses also have extra, small RNA segments associated with them, called ‘satellite RNA’. These can alter the pathology of the virus infection. Where viruses are mentioned in the text, but no specific background is given, further information can be obtained from the web links listed at the end of this chapter.

To highlight the organisation and different modes of gene expression patterns found with plant viruses, two members of the most common group, positive-sense RNA viruses, which demonstrate widely different modes of gene expression, will be used as examples in the next section.

RNA viruses

Tobacco mosaic tobamovirus (TMV)

Figure 8.1 shows the genome structures and expression patterns of Tobacco mosaic *tobamovirus* (TMV). TMV is perhaps the longest studied of all viruses, it has been the subject of intensive investigation for over 100 years. It has a genome composed of one, single-stranded linear RNA molecule (~6400 bases) that has a tRNA-like structure at the 3'-end. The genome codes for five translated open reading frames (ORFs), of these, at least three ORFs are non-structural proteins: a 183-kDa protein, the 126-kDa protein and the 30-kDa protein. The two larger proteins are involved in the replicase complex, the 126-kDa has a methyltransferase activity, which is involved in RNA capping, and a helicase function. The 183-kDa protein is a read-through product of the 126-kDa protein that also has a polymerase function. The 30-kDa protein is a movement protein (MP). The RNA also codes for a putative 54-kDa protein of unknown function, and the capsid or coat protein (CP of 17-kDa). Translation of the TMV open reading frames is quite complex (see Figure 8.1). On entry into the cell the two high molecular weight proteins, 126-kDa and 183-kDa, are translated directly from the full-length genomic RNA, which acts as messenger RNA. The 183-kDa protein is produced by read-through of the amber termination codon that terminates translation to produce the 126-kDa protein. In addition, there is a start codon within this read-through region and an open reading frame that could potentially encode the putative

54-kDa protein. Alternatively, a subgenomic RNA (I_1) may code for this protein. The replication complex copies the positive-sense RNA to form negative-sense RNA, which subsequently acts as a template for the production of the positive-sense RNA. Subgenomic messages are transcribed from two promoters within the negative-sense strand, and these act as templates for the synthesis of the MP and the CP from the I_2 RNA and CP subgenomic (sg) RNAs, respectively (Figure 8.1).

Comoviruses (CPMV)

Figure 8.2 shows the genome structure and translation strategy for a comovirus cowpea mosaic *comovirus* (CPMV). The genome is segmented, it is split into two parts, RNA1 and RNA2. RNA1 has the coding sequences for the core polymerase, a protease, and a VPg (genome-linked protein) protein. The VPg is a protein that is attached to the 5'-end of the molecule

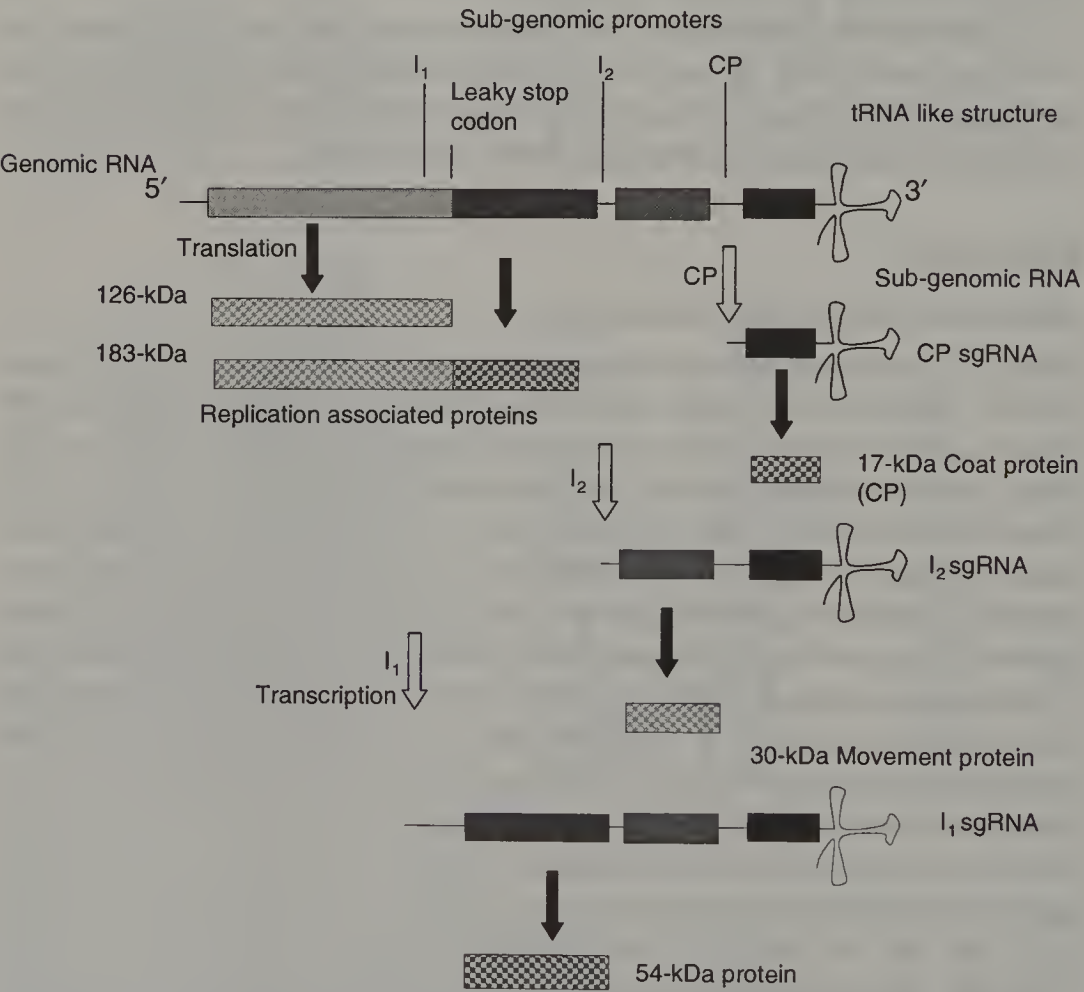


Figure 8.1 Transcription and translation of the TMV genome. Initial translation produces a replicase which then synthesises subgenomic mRNAs from (–) strand. Transcription, empty arrows; Translation, filled arrows. (Adapted with permission from Rybicki, E.—web-link 8.3: <http://www.uct.ac.za/microbiology/tutorial/RNA%20virus%20replication.htm> and Bustamante, P. I. and Hull, R.—web-link 8.1: <http://www.ejb.org/content/vol1/issue2/full/3/index.html>)

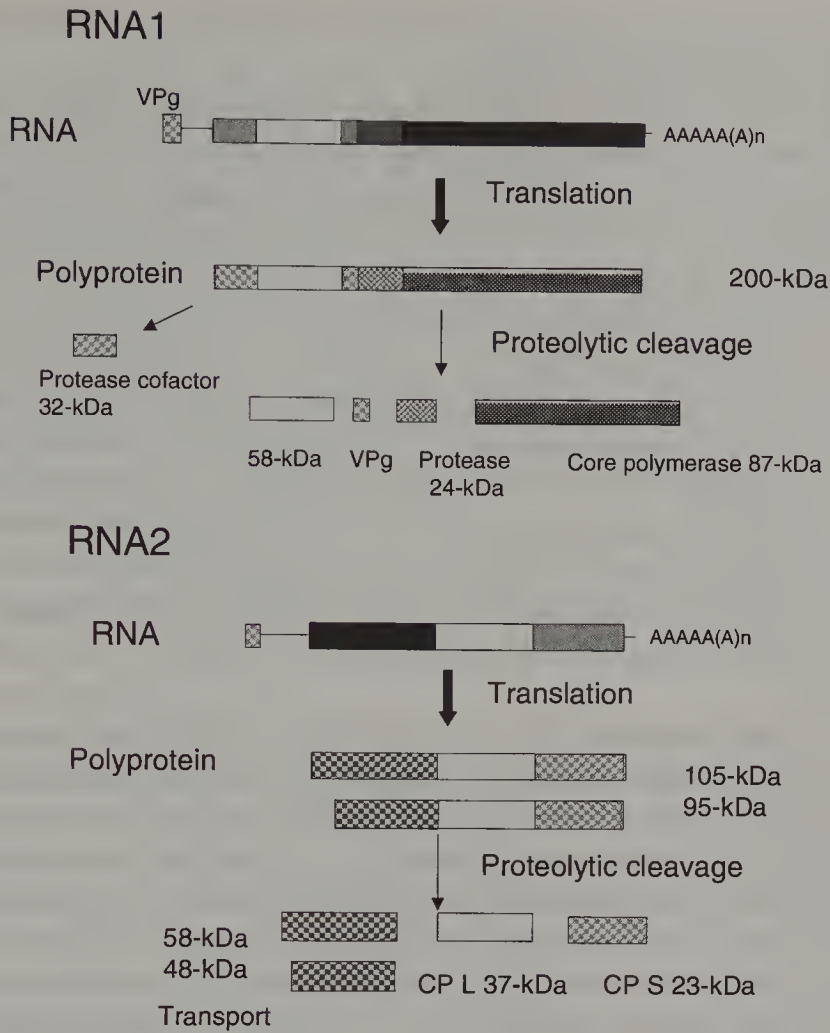


Figure 8.2 Translation of the CPMV genome. (Adapted with permission from Lomonossoff, G. P. and Johnson, J. E. (1991).)

and fulfils a cap function. The coding sequence for, in this case, the two coat-protein subunits are found on RNA2 along with the movement protein. On entry into the cell the genomes act as messages and can be translated directly. Large polyproteins are made that represent the entire coding sequence (Figure 8.2). These are then cut into the active proteins by specific proteases. Other members of the *Comoviridae*, such as nepoviruses, which will be discussed later in this chapter, demonstrate this type of genome structure and mode of gene expression.

Entry and replication—points of inhibition

As with any virus infection the virus has to enter the host cell to replicate. Individual viruses have specific mechanisms for their initial entry into the cell and subsequent movement through the plant. Many are transferred between

plants by vectors such as nematodes, various arthropods (such as aphids, leaf hoppers and whiteflies) and fungi.

Once inside the cell the viruses uncoat, and the viral genetic material is then translated and replicated. Replication is carried out by the replicase complex encoded for by the virus. The ultimate stage of the infection is the packaging of the replicated viral nucleic acid into capsids. One interesting point is that the nucleic acid molecules are packaged into the capsids separately, each RNA molecule is packaged into its own capsid. This has been shown for icosahedral particles of CPMV by gradient centrifugation, a technique whereby it is possible to size-fractionate separate populations on the basis of their RNA content. Also, electron microscopy (EM) has been used for rod-shaped viruses such as beet necrotic yellow vein *furovirus* (BNYVV), which has four RNAs in most isolates. In this case, it is possible to observe four different-sized populations of virus particles when viewed using the EM. The mechanisms involving translation and replication will depend on the types of virus; further information can be obtained from the web links listed at the end of the chapter.

For a systemic infection, the virus must move from cell to cell via the plasmodesmata and/or the vascular tissues. The exclusion limit of the plasmodesmata pore is normally too small to allow the passage of virus particles. The viral movement proteins may function by increasing the molecular size exclusion limit of the plasmodesmata allowing passage of the virus. Different viruses exhibit different modes of movement, some are transferred as RNA, which may be coated with capsid or movement proteins, or they may be transferred as virus particles. For instance, with TMV the 30-kDa movement protein modifies the plasmodesmata to allow the passage of genomic RNA coated by the 30-kDa protein to pass from cell to cell. With comoviruses such as cowpea mosaic virus (CPMV) the 58/48-kDa proteins form tubular structures through the plasmodesmata which allow the passage of intact virus particles from one cell to another. Other mechanisms may exist for transport into the vascular tissue.

How has industry dealt with viruses?

Farmers and the agricultural industry have dealt with the problem of viruses in a number different ways, based on whether the viruses are normally present or not.

Where viruses are endemic, the direct chemical control of viruses is not normally an option. However, some control can be exerted over the spread of the virus with chemical control methods if it is transmitted by an insect or another vector, but often this only has a limited impact.

Some plant systems have benefited from another approach known as ‘cross-

protection'. Since the 1920s it has been known that some plants infected with a virus were apparently protected against infection by a related virus. One strategy therefore has been to deliberately infect plants with viruses that produce only mild symptoms but which generate resistance to infection by more virulent strains. Mild strains have been obtained for some viruses from areas where plants apparently showed some resistance to infection. When analysed, the plants were found to harbour less infective strains of the virus. These viruses were then used to protect other plants against infection. Cross-protection has been used with some success to combat disease caused by viruses such as citrus tristeza *closterovirus* (CTV), i.e. the most important disease of citrus trees (a \$US10 billion worldwide industry), papaya ringspot *potyvirus* (PRSV) and zucchini yellow mosaic *potyvirus* (ZYMV). A similar approach has been used with strains of cucumber mosaic *cucumovirus* (CMV) that have satellite RNA associated with them. The presence of the satellite RNA reduces the pathogenicity of the initial virus infection, and it has been shown that it also reduces the symptoms caused by other CMV related viruses that do not have the satellite.

There is, though, some hesitance amongst the farming community to adopt the cross-protection strategy completely as there are some worries about deliberately infecting plants and controlling the properties of the mild viruses that could revert to a virulent form. However, these cross-protection strategies form the basis for the main GM strategy discussed later in this chapter.

Another approach that is used is the production of virus-free plant stocks. This has been particularly important with plants such as potato, sweet potato and strawberry that are vegetatively propagated and are therefore clonal material. The symptoms of virus infections may not be readily apparent during vegetative growth, but they can have serious effects on yield and/or quality of the crop. Viruses such as potato leaf-roll *polerovirus* (PLRV) and potato Y *potyvirus* (PVY) can reduce tuber yield by 50–80%. Entire clonal populations of the potato stocks may become infected because of the mode of propagation, so great effort has been put into systems for the development of virus-free stocks. The most common systems that have been in use since the 1960s have been heat therapy and meristem tip culture (see Chapter 2). More recently, antiviral chemicals have also been used. In principle, combinations of these technologies can be used to produce virus-free stocks. These can then be used in the field, but they do not provide protection against subsequent infection.

Many countries try to operate a quarantine system to prevent the import of virus diseases into non-endemic regions. This has been a major part of the strategy adopted in the UK, and is highlighted by the long-term struggle to keep the sugar beet disease, rhizomania, out of the country. This disease is caused by beet necrotic yellow vein virus (BNYVV), a positive-sense RNA furovirus, that is carried in soil from one plant to another by spores of the fungus *Polymyxa betae*. It can reduce sugar yield by quite significant amounts. The disease has slowly moved through the sugar beet growing areas of the

world transmitted via contaminated soil on vehicles, tools and unwashed crops. Unfortunately, it has been difficult to identify contaminated locations before symptoms appear and thus allowing further spread to occur. One method of detection that is still being used is to overfly beet fields to look for changes in foliage appearance, a symptom of many virus infections. The major problem for detecting soil contaminated with BNYVV was that the assay used was a bait test, which is shown in Figure 8.3. Soil, potentially contaminated with BNYVV was brought into the laboratory and used as a medium to grow non-contaminated test plants to see if they developed the disease. The plants would then tested for the presence of the virus with ELISA (enzyme-linked immunosorbent assay). The system was time-consuming and insensitive. To overcome these shortcomings, molecular biology techniques such as hybridisation and PCR have been used to improve sensitivity and shorten the test time (see Figure 8.4).

Unfortunately, these efforts have not worked and Great Britain is no longer a ‘protected zone’ for rhizomania. In order to protect the sugar industry, sugar beet infected with the virus is allowed to be processed but is strictly managed to minimise any infected soil spreading the disease on to other land.

The most important approach to fighting BNYVV disease has been the

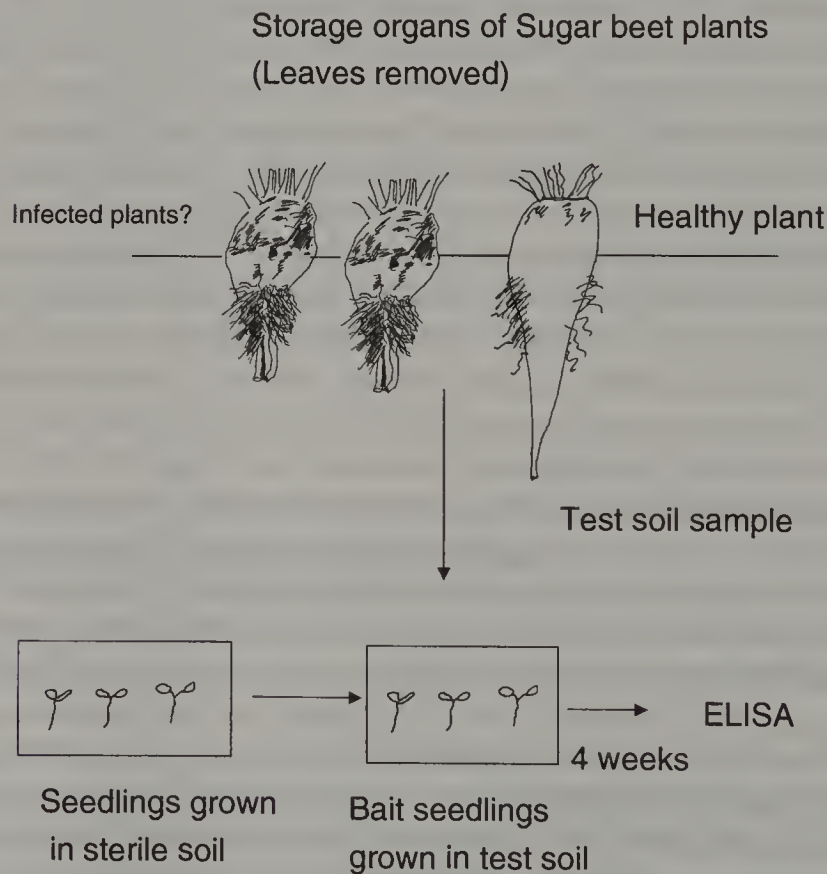


Figure 8.3 Soil bait test scheme.

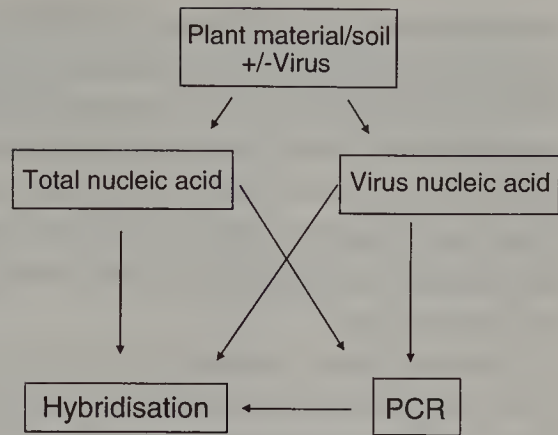


Figure 8.4 Molecular tests for viral nucleic acids.

development of crops with resistance. Previously, there had been some reticence about this strategy as in some cases the plants that were produced were tolerant, rather than resistant—the virus infected the plant, but the symptoms and virus titre were much reduced. While the crop was protected it could act as a massive reservoir of virus that could infect non-resistant varieties. Recently, it has proved possible to produce varieties that resist rhizomania infection, which produce comparable yields of sugar in non-infected and infected trials and show low levels of BNYSV within the roots. This low level of BNYSV is important, as it will help to slow virus build-up in affected fields and help to restrict spread to uninfected areas and the processing factories. The production of such varieties has been possible because resistance genes exist within related *Beta* species and it has been possible to breed them into ‘elite lines’. In many crops such resistance genes for specific diseases have not been identified. Therefore other alternatives need to be considered, in particular pathogen-derived resistance.

The transgenic approach—PDR

The first, and still the main, antiviral transgenic approach used has been that of pathogen-derived resistance (PDR), originally known as parasite-derived resistance. In PDR, pathogen genomic sequences are deliberately engineered into the host plant’s genome, on the non-scientific basis that the sequence may be expressed at an inappropriate time, in inappropriate amounts or in an inappropriate form during the infection cycle and thereby induce some form of resistance in the plant. This type of strategy has, in part, developed from experiments such as cross-protection studies, where the introduction of a minor viral pathogen induced resistance to a related but more virulent pathogen. However, many of the developments in this area have been made with no substantial knowledge base driving them forward. Indeed, it has been said that

the reverse is true, and the use of PDR has led to a lot of basic information on virus biology and disease resistance being discovered. The mechanisms of PDR, as we will see, are varied and complex; the presence of the pathogen sequence may directly interfere with the replication of the pathogen or may induce some host defence mechanism. It has become clear that these interactions are not simple. As more is learnt about the underlying mechanisms and the resistance genes involved, new strategies involving plant resistance genes will be developed.

Potential types of interaction between transgenic sequences and virus life cycles can be categorised into two groups: those involving the synthesis of viral proteins and those involving viral RNA. In this section we will discuss some of these strategies.

Interactions involving viral proteins

The first and most successful transgenic approach to viral resistance has been with the transgenic expression of the coat protein (CP) coding sequence. CP-mediated resistance was first reported with a TMV-tobacco model system in 1986. Subsequently, a large number of transgenic lines containing CP transgenes have been produced for a whole range of crop species and many different viruses, some examples are listed in Table 8.2. This is by no means a complete list, as there are far too many examples to include them all. Specific examples will be discussed in subsequent sections of this chapter. In many cases, some degree of resistance to virus infection has been reported; however, some variations in the level of resistance are found that can be related to transcriptional gene silencing, transgene position effects and the relationship between the coding sequence and target virus. Some of these points are discussed later.

There are several important features of the CP strategy. First, there is some level of cross-protection against infection by related viruses. As will be discussed later, resistance is not limited to viruses with the same nucleic acid sequence. Another important point is that, for TMV and several viruses from other groups, the level of resistance has been related to the level of transgenic coat protein produced. This is true for alfalfa mosaic *alfamovirus* (AlMV) and potato X *potexvirus* (PVX) but not for potato Y *potyvirus* (PVY) in which resistance was found in plants with no detectable level of transgenic coat protein. This suggests that several different resistance mechanisms may be at work. Much of the work with TMV has focused on understanding the mechanism of TMV coat protein-mediated resistance (CPMR). Different mutant CP genes have been made which are either unable to form virus-like particles (VLP), associate with other CPs or have increased stability when introduced into tobacco. These studies suggested that CPMR TMV-like virus was due to an interaction between the transgenic CP and invading virus. With the non-VLP-forming mutant CPs there was no resistance. With mutants that

Table 8.2 Examples of coat protein-mediated resistance

Plant	Source of coat protein gene	Virus resistance exhibited to
Alfalfa	AIMV	AIMV
Citrus	CTV	CTV
Papaya	PRSV	PRSV
Peanut	TSWV (N-gene)	TSWV
Potato	PVX	PVX, PVY
	PVY	PVY
	PLRV	PLRV
Rice	RTSV	RTSV
	RSV (N gene)	RSV
	RHBV (N gene)	RHBV
	RYMV	RYMV
Squash	CMV	CMV
	WMV 2	WMV 2
	ZYMV	ZYMV
Sugar beet	BNYVV	BNYVV
Tobacco	TMV	TMV, PVX, CMV, AIMV
	CMV	CMV
	AIMV	AIMV
Wheat	SBWMV	SBWMV
	BYDV	BYDV

List of virus names can be obtained from the web links: <http://image.fs.uidaho.edu/vide>

AIMV	Alfalfa mosaic <i>alfamovirus</i> —
BNYVV	Beet necrotic yellow vein <i>furovirus</i> —
BYDV	Barley yellow dwarf <i>luteovirus</i> —
CMV	Cucumber mosaic <i>cucumovirus</i> —
CTV	Citrus tristeza <i>closterovirus</i> —
PLRV	Potato leafroll <i>luteovirus</i> —
PRSV	Papaya ringspot <i>potyvirus</i> —
PVX	Potato X <i>potexvirus</i> —
PVY	Potato Y <i>potyvirus</i> —
RHBV	Rice hoja blanca <i>tenuivirus</i> —
RSV	Rice stripe <i>tenuivirus</i> —
RTSV	Rice tungro spherical <i>waikavirus</i> —
RYMV	Rice yellow mottle <i>sobemovirus</i> —
SBWMV	Wheat soil-borne mosaic <i>furovirus</i> —
TMV	Tobacco mosaic <i>tobamovirus</i> —
TSWV	Tomato spotted wilt <i>tospovirus</i> —
WMV 2	Watermelon mosaic 2 <i>potyvirus</i> —
ZYMV	Zucchini yellow mosaic <i>potyvirus</i> —

had increased CP stability in transgenic plants there was an increase in CPMR when challenged with virus particles. The proposed explanation for TMV CPMR was that the CP was actively inhibiting the uncoating of the infecting virus so preventing translation and replication. It was also suggested that the CP may interfere with the spread of some viruses. In addition, it was clear that transgenic TMV-CP did not stop infection when the plants were challenged with naked TMV RNA. Transgenic plants expressing the coat

proteins of other viruses such as PVX and AIMV did, however, display resistance to infection by their RNA. The important point is that, as indicated earlier, other viruses have different properties to TMV, and that other mechanisms exist for CPMR at the protein level and, as we will see later, at the RNA level.

Attempts at using other virus proteins have met with various degrees of success. The genes for replication related proteins (complete and partial sequences) have been introduced into plants with varying success. Where resistance has been observed it is generally specific to the virus from which the sequence was obtained. There is some debate over the mechanism of these effects, at the protein level there may be some interactions between the transgenic replicase protein and that coded for by the virus. However, it has also been suggested that the resistance is an RNA effect. This will be discussed later in this chapter. Attempts to use the transgenic expression of mutant movement proteins to block the spread of viruses have met with varied success. As an example, transgenic plants containing a defective TMV movement protein show resistance to several tobamoviruses, AIMV, cauliflower mosaic *caulimovirus* (CaMV) and several other viruses.

In the next section we will look at a case study on the practical aspects of isolating a CP coding sequence and expressing it in a plant.

CASE STUDY **Arabis mosaic virus (ArMV)**

Arabis mosaic nepovirus (ArMV) is one of a group of economically important nepoviruses (which are transferred by nematodes), that also includes grapevine fanleaf *nepovirus* (GFLV) and raspberry ringspot *nepovirus* (RRV). These are very important in the wine industry, causing significant reductions in grape yield and quality. As the virus is transferred by nematodes it was felt that developing resistant root stocks would provide adequate protection against the virus. However, as no suitable resistance markers are available in the germplasm for breeding programmes, several groups have looked at transgenic approaches to produce resistant lines. In this case study the production and laboratory testing of CP constructs in tobacco will be discussed. The aim of the project was to produce high levels of CP in the plant as there can be a positive correlation between CP expression levels and resistance. Therefore a strong promoter, transcription enhancers and a translation enhancer were used. The methods described in this case study were used for ArMV but they would be suitable for other viruses.

ArMV is a positive-sense RNA virus that has its genome on two separate molecules, RNA1 and RNA2. cDNA clones have been made to these sequences and it has been possible to locate the position of various protein coding sequences in the genome. The coat protein is coded for on RNA2, but, as we have seen for the members of the *Comoviridae*, it is initially synthesised as part of a polyprotein that is processed in the cell to release the capsid or coat protein (CP). The first step in producing the CP vector was to identify the precise loca-

tion of the CP gene sequence in the genome. This was done, in part, by using computer analysis to look at sequence homologies. The capsid of CPMV has two CP subunits, but with the nepoviruses there is a single CP subunit (64 kDa) and this is to be found at the 3'-end of RNA2 (Figure 8.5). The 3'-end of the CP gene was defined by the stop codon (see Chapter 1) of the open reading frame. To identify the 5'-end of the protein sequence was more difficult as it is located in the region coding for the polyprotein. However, it was possible to obtain the N-terminal amino acid sequence of the CP by sequencing CP that had been purified by way of various phase separations and centrifugation in linear sucrose gradients. This could then be compared with the sequence obtained by computer (virtual) translation of the RNA sequence and the 5'-end identified.

PCR primers (oligonucleotides) were then designed to amplify, using PCR, the complete coding sequence of the coat protein. The primers were complementary to about 30 nucleotides of the virus and they had several restriction endonuclease sites, which did not occur in the sequence, incorporated at the 5'-end of the oligonucleotides. The primer for the N-terminal end of the sequence also contained an AUG start codon adjacent to and in-frame with the CP coding sequence. PCR was carried out using RNA2 cDNA as the template. The amplified cDNA was then digested within the oligo-restriction sites to allow the amplified CP construct to be directionally ligated into an *Escherichia coli* cloning vector. The ligation was carried out so that at the N-terminus of the construct there was a CaMV promoter with several transcription enhancers and a TMV Ω sequence, to optimise the level of translation. At the C-terminal end there was a nopaline synthase-derived transcription terminator sequence (NS) terminator (see Chapter 4). The complete sequence was then digested out of the intermediary plasmid and ligated into the binary vector pBIN19 (pBIN19 is described in Chapter 4)—which was used for *Agrobacterium* transformation. Many binary vectors are now available that make the intermediate step in the *E. coli* vector unnecessary. This general scheme is shown in Figure 8.5.

The CP construct was introduced into tobacco using the protocol described in Chapter 3, and transformed plant material was selected for by incubating the plant material on a medium containing kanamycin. A number of transgenic plant lines were obtained and these were screened for protein expression levels using ELISA. Transgenic CP levels of the order of $20 \mu\text{g g}^{-1}$ fresh leaf tissue were found in various plants (a similar level of expression as compared to the TMV subgenomic CP promoter levels of proteins discussed in Chapter 10). Interestingly, microscopic examination of the transgenic plants indicated that the coat protein was forming virus-like particles (VLPs) in the absence of the full virus RNA genome. These plants were tested for resistance to ArMV by mechanical inoculation of the leaves either with whole virus, RNA or water as a control. The local necrotic lesions induced by the virus on the inoculated leaves were counted, and it was found that the transgenic plants had much fewer than the control plants. ELISA analysis was also used to assay the inoculated leaves (local infection) and the tips of new leaves (systemic infection) for the quantity of virus antigen present. The transgenic plants were found to have much lower levels of coat protein antigen present in the new leaf material compared to the levels found in infected control plants. The levels of antigen detected are shown in Figure 8.6. The data indicate quite clearly that in the transgenic plants the presence of the ArMV coat protein is

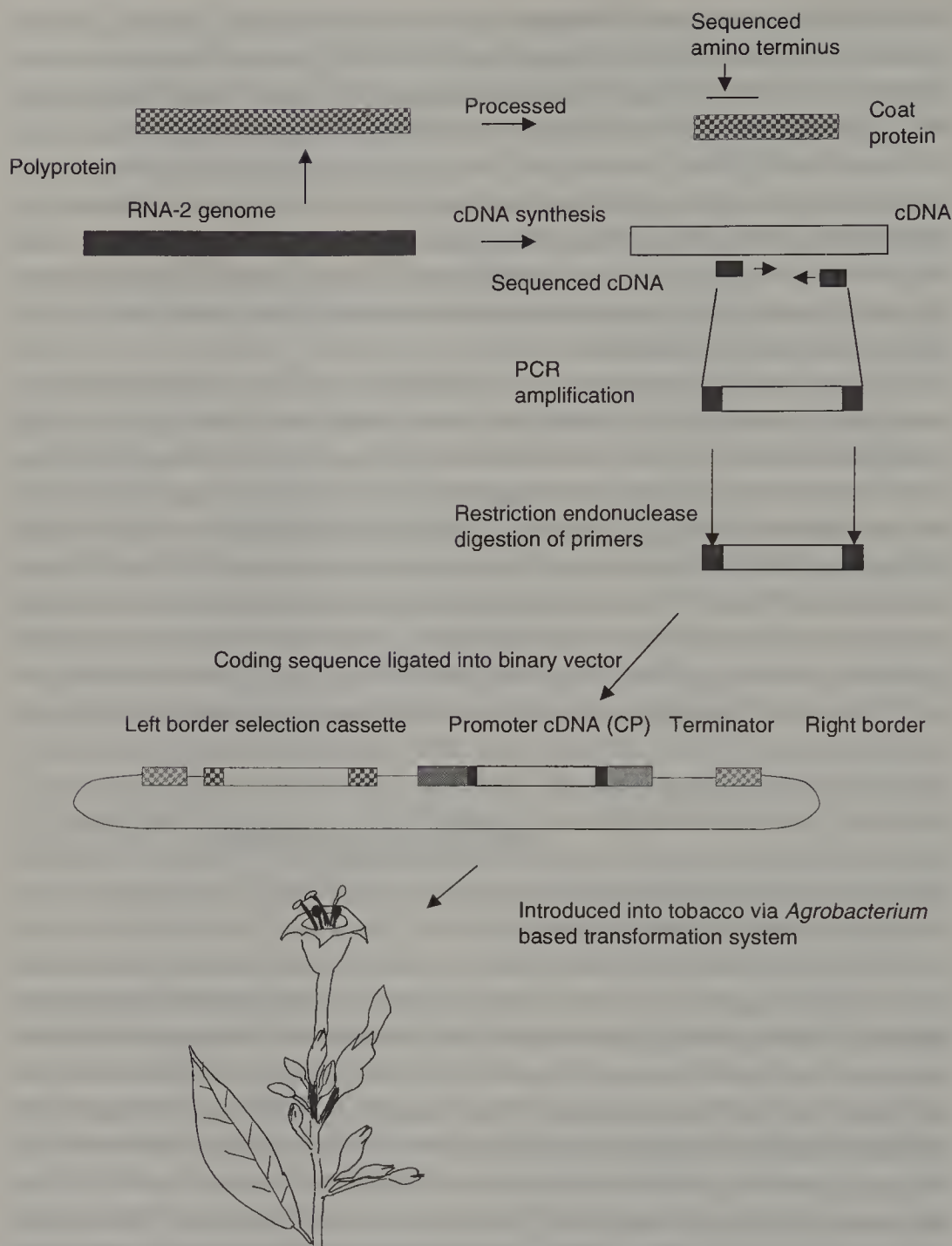


Figure 8.5 Strategy for the construction of ArMV transformation vectors and transgenic plants.

inhibiting the systemic infection process by both the virus particles and naked RNA (unlike the situation for TMV discussed earlier).

RNA effects

Some of the most interesting developments in the area of PDR relate to RNA effects. Several approaches have been developed that do not require protein

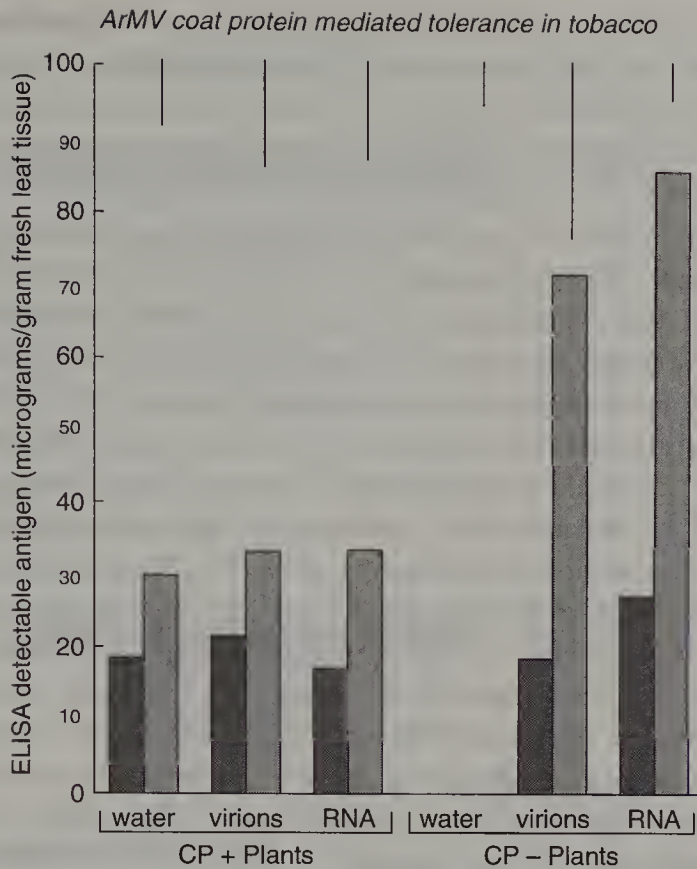


Figure 8.6 Quantities of ArMV coat protein, as detected by ELISA, in tip leaves (lighter shaded bars) and inoculated leaves (darker shaded bars) of plants expressing the ArMV coat protein (CP+ plants) and control plants not expressing the ArMV coat protein (CP- plants) inoculated with water, ArMV virions or ArMV-derived RNA. The 95% confidence limits for the tip leaf data are given by the lines at the top of the graph. (Redrawn with permission from Bertoli, D.J., *et al.* (1992) *Ann appl. Biol.* 120. 47–54).

synthesis, these include satellite RNAs, antisense RNA and ribozyme technologies. Perhaps the most important development in this area though is gene silencing. Studies where virus coding sequences that have been mutated and then introduced into plants led to the conclusion that high levels of RNA synthesis could lead to a post-transcriptional gene silencing (co-suppression) phenomenon. This area of research has gone some way to explain some of the modes of resistance obtained with CP studies, but it has also led to the development of alternative strategies for PDR.

Satellite sequences

Plant viral satellite RNAs are small RNA molecules that are unable to multiply in host cells without the presence of a specific helper virus. The satellite is not required for virus replication but may affect disease symptoms. In many cases the presence of a satellite RNA may reduce the effect of the virus, so one strategy has been to produce transgenic plants containing satellite sequences. In the section on cross-protection it was noted that cucumber mosaic

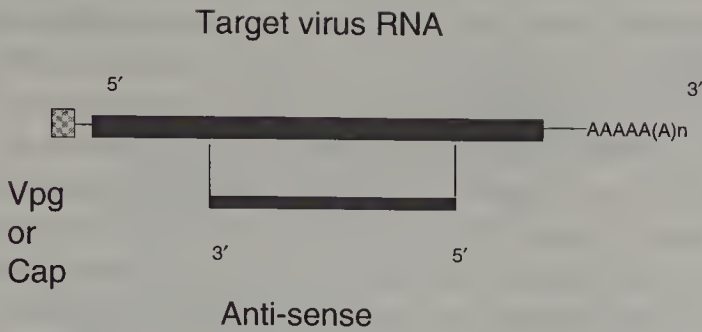
cucumovirus (CMV) symptoms were reduced when the virus was carrying a satellite and that this had been used as a commercial method for antiviral protection. Major studies have been carried out to see if a transgenic approach could benefit this approach. Transgenic tobacco and tomato plants expressing CMV satellite RNA were tested in the field in China (1990–1992). Although some reduction in disease symptoms was found, it was not strong enough to protect the plants. However, a strategy was developed in which satellite RNA PDR was developed in combination with CMV CPMR. It was found that the resistance was stronger than that of either CPMR or satellite-PDR alone. This appears to be a good approach as combinations of genes are likely to generate more durable resistance; but some studies indicate that while satellites RNAs generally have a neutral or attenuating effect on the helper virus, in some cases the severity of the disease can be increased. In one study, where the effect of satellite RNA on the pathogenicity of an ArMV isolate was investigated, of the 42 different plant species tested it was found that the presence of the satellite either had no effect or reduced the effect of the virus in most cases. However, the satellite worsened the effects of the virus in several leguminous species. The authors of this study comment on the potential risks of including the sequence in transgenic plants. A related approach has been the inclusion of small defective interfering RNAs and DNAs (DI). These are small sequence fragments generated during the replication of some viruses. The presence of the sequences redirects the replication machinery towards the DI, reducing the amount of virus. This strategy has some relationship to that of co-suppression.

Antisense and ribozymes

One simple strategy that has been attempted is the use of antisense RNA approaches (see Chapter 10), ribozymes are also included within this broad category. To produce an antisense-mediated resistance to viral disease in transgenic plants, constructs have been designed that express a negative-sense RNA molecule that will hybridise with the infecting virus sequence (Figure 8.7(a)) The antisense approach was designed to specifically interfere with virus replication.

Ribozymes are antisense catalytic RNA molecules capable of catalysing the cleavage of the target sense RNA sequence. The constructs are similar to those used for the antisense approach, except that a short catalytic sequence is embedded within a specific target region (Figure 8.7(b)). The aim is to both block replication by the formation of a double-stranded RNA:RNA hybrid and to cut a key region of the virus genome (such as within a replicase complex genes) before it is able to replicate. As both antisense and ribozyme technologies are based on sequence homology, they have limited cross-protection value unless the antisense sequence is designed against a very conserved region. These homology-based approaches have met with mixed success. There have been some reports that the antisense RNA does provide protection, but this may

(a) anti-sense



(b) ribozyme

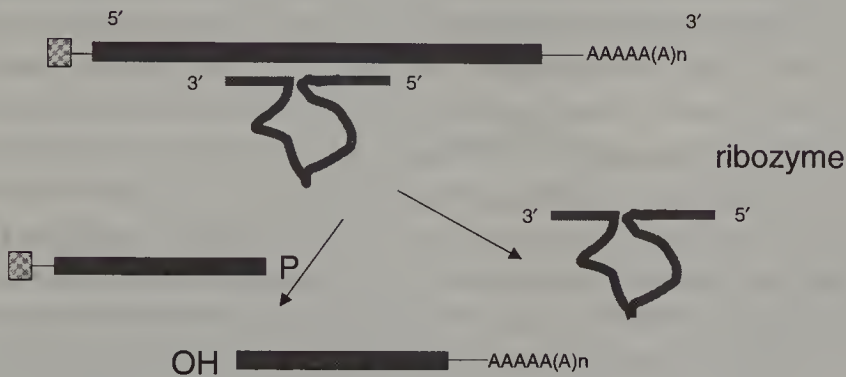


Figure 8.7 RNA approaches to PDR. (a) Antisense; (b) ribozyme.

be due to gene silencing (which is described in the next section). The ribozyme approach has generally not been very successful in plants as an antiviral tool.

Gene silencing/co-repression—the explanation for the RNA effects

The most important phenomenon associated with RNA-mediated PDR is post-transcriptional gene silencing (PTGS). PTGS was discussed briefly in Chapter 3 in the context of plant transformation. This phenomenon is found not just in plants but has been described in a wide range of organisms (fungi, nematode worms, humans) and seems to be an important system for preventing invasion by viruses and transposon sequences, as well as having some regulatory functions in the cell.

Two separate avenues of research in plants uncovered the phenomenon of gene silencing. In studies with transgenic petunia plants expressing introduced chalcone synthase genes, under the control of strong promoters (which were designed to produce deep-purple flowers), it was found that pigment production in certain regions of the petals was often turned off and

predominantly white variegated flowers were produced. This effect was termed 'co-suppression' since both the introduced gene and the endogenous genes were turned off. With plant viruses, the first evidence for PTGS came from studies with tobacco etch *potyvirus* (TEV). In these studies virus CP transgenes, with in-frame stop codons so they were unable to produce full-length proteins, were found to give higher levels of protection against TEV infection than wild-type CP transgene constructs. Other related observations came from experiments performed with a mutated tomato spotted wilt *tospovirus* (TSWV) nucleocapsid protein, potato Y *potyvirus* (PVY) CP constructs and with various forms of the RNA-dependent RNA polymerase gene (RdRp) from potato X *potexvirus* (PVX) that were incapable of being translated into functional proteins. Initially it was thought the effects were related to an antisense effect, in which the positive-sense RNA from the transgene was binding to the negative-sense strand of the virus. However, it soon became clear that with the resistant lines no correlation could be found between a high level of RNA from the transgene and the high level of resistance. On the contrary, the resistant plants seemed to have low steady-state levels of the transgene RNA, although run-off experiments, where the actual level of gene expression is measured, indicated that transcription levels of the RNA seemed to be high. From these and other observations, a simple model for PTGS was developed. It is believed that the cell has a mechanism by which it can monitor RNA levels, it scans for unusual amounts of double-stranded nucleic acids, such as in a viral infection where positive and negative hybrids may form during the replication process, and above a certain critical level the RNA is rapidly degraded. In the context of PDR and virus infection, it is believed that plants showing good resistance due to the RNA effect are the ones in which the transgene RNA levels are very close to, or have exceeded, the critical level. The transgene RNA is being expressed at a high level in these plants, but is being rapidly degraded. So on infection by the same or a very closely related virus, the cell is already primed to degrade the invasive RNA, which it does, therefore stopping the infection. While this 'critical level' model explains many of the results, there are some problems. It has been modified to allow for the possibility that it may not necessarily be the quantity of the dsRNA that is produced, but may, in part, be its quality that triggers destruction. It has been proposed that the cell can screen for dsRNA molecules that may contain aberrant RNA. How these are actually produced is still not clear, but they may result from 'cryptic' transcription start sites (resulting in antisense RNA), transcription through inverted repeats or through the activity of an RNA-dependent RNA polymerase. This model can also be used to explain the co-repression phenomenon observed in transgenic plants.

Great progress has been made in our understanding of the mechanism of PTGS, partly due to the phenomenon being present in other eukaryotes. An overview of the current thinking is given in Figure 8.8. The dsRNA is digested into 21–23 nucleotides (19–21 base duplexes with 2 base overhangs at their

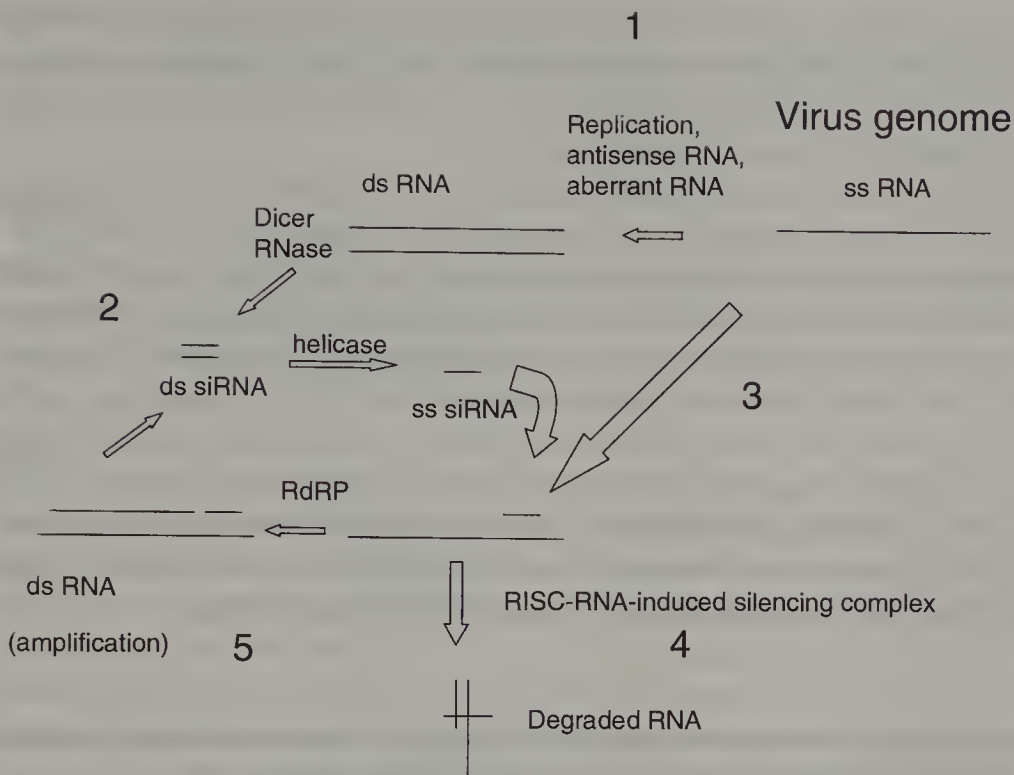


Figure 8.8 A model for the mechanism of post-transcriptional gene silencing (PTGS). (1) Double-stranded RNA (dsRNA) forms in the plant cell. (2) The dsRNA is recognised by the Dicer complex and is degraded to single-stranded (ss) siRNA. (3) Invading virus is detected entering the cell and is (4) degraded by the RISC complex. (5) Alternatively, the RNA is part of an amplification complex. (Adapted with permission from Melcher, U.—web-link 8.2: <http://www.opbs.okstate.edu/~melcher/mg/MGW3/MG33322.html>)

3'-ends), small interfering RNAs (siRNAs—which are also known as 'guide RNAs'). This is carried out by a RNA III double-stranded specific ribonuclease (in *Drosophila* this is known as Dicer). These siRNA duplexes then bind to a complex, which is believed to contain a different nuclease, to form the RNA-induced silencing complex (or RISC). An associated ATP-dependent helicase then unwinds the duplexes and the RISC then targets homologous single-stranded RNA transcripts, via base-pairing interactions, and cleaves the RNA. It has been suggested that some of the short RNA molecules (possibly a different size range) act as primers for the synthesis of more double-stranded molecules, thus introducing an amplification step in the process. It is also believed that the short siRNAs may act as systemic signals that are transported throughout the plant thereby inducing systemic gene silencing.

One question that comes from these observations is that if plants contain effective defence mechanisms how do viruses manage to cause an infection? Part of the answer to this question is that virus proteins that can act as suppressors to gene silencing have been identified. In several potyviruses the HcPro protein has been identified as having this function. Similarly, in CMV the 2b protein

has been shown to act as a suppresser of gene silencing. It is also interesting to note that these proteins function in different ways. HcPro blocks the maintenance of already established gene silencing. The 2b protein, on the other hand, suppresses the initiation of gene silencing.

It is likely that many of the effects seen with CP constructs are due to gene silencing. Recently, there have been deliberate attempts to use PTGS as a mechanism to develop resistance. One group has inserted an inverted repeat of CMV cDNA into tobacco. On transcription, this generates an RNA sequence with the ability to form intramolecular double-stranded molecules. These have been shown to induce resistance to CMV by switching on the PTGS system. This was confirmed by the presence of the short siRNAs discussed earlier. In another recent study where nepovirus sequences (coding for several proteins, including the movement protein) were introduced into grapevine root stocks as inverted-repeat constructs, in order to induce PTGS, little effect was observed. Exactly why, is not clear at this stage.

What has been commercialised in the West?

Yellow squash and zucchini

The Asgrow seeds company markets several varieties of squash and zucchini under the names Independence II, Liberator III, Freedom III and Destiny III. These lines are resistant to three important viral diseases: zucchini yellow mosaic *potyvirus* (ZYMV), watermelon mosaic 2 *potyvirus* (WMV-2) and cucumber mosaic *cucumovirus* (CMV). These are all serious pathogens of cucurbits, causing leaf mottling, yellowing, dwarfing and deformed fruit. Conventional resistance routes have been ineffective; the genetically engineered strains provide better levels of resistance, although they are susceptible to variations in the pathogen strain and the environment. There are two interesting points about these lines. The constructs used for some of the strains contain three separate CP coding sequences for the viruses. They also benefited from the serendipitous loss of linkage between the selectable marker and the CP constructs. This allowed marker-free plants to be produced by simple breeding strategies.

Papaya

Several different genetically engineered papaya lines have been produced and marketed commercially. The amazing thing about the production and commercialisation of these is the rapid time scale in which it was brought about.

Beginning in 1992, a very serious outbreak of papaya ringspot *potyvirus* (PRSV) had reduced Hawaii's papaya production by over 40% within 5 years.

Workers at Cornell University and in Hawaii produced two genetically engineered lines, SunUp and Rainbow, that have been available to farmers since 1998.

Potato

Monsanto marketed NewLeaf potato lines that had both *Bt* resistance and resistance to several virus lines (potato leafroll *polerovirus*, potato Y *potyvirus*). Due to a low take up of these lines they have been withdrawn from sale.

Risk

The commercialisation of the CP lines has been a particular target for opponents of GM technology who have attacked them over the problems discussed in Box 8.1 (transcapsidation and recombination). They have also raised issues over the presence of the CP sequence giving the plants some growth advantage, in essence turning them into weeds. The same argument has been used for situations where the CP gene may be transferred to wild relatives. As some of these plants are in the field, the big experiment has begun and time will tell. The durability of disease resistance will also now be tested.

BOX 8.1 The risk of producing a new virus

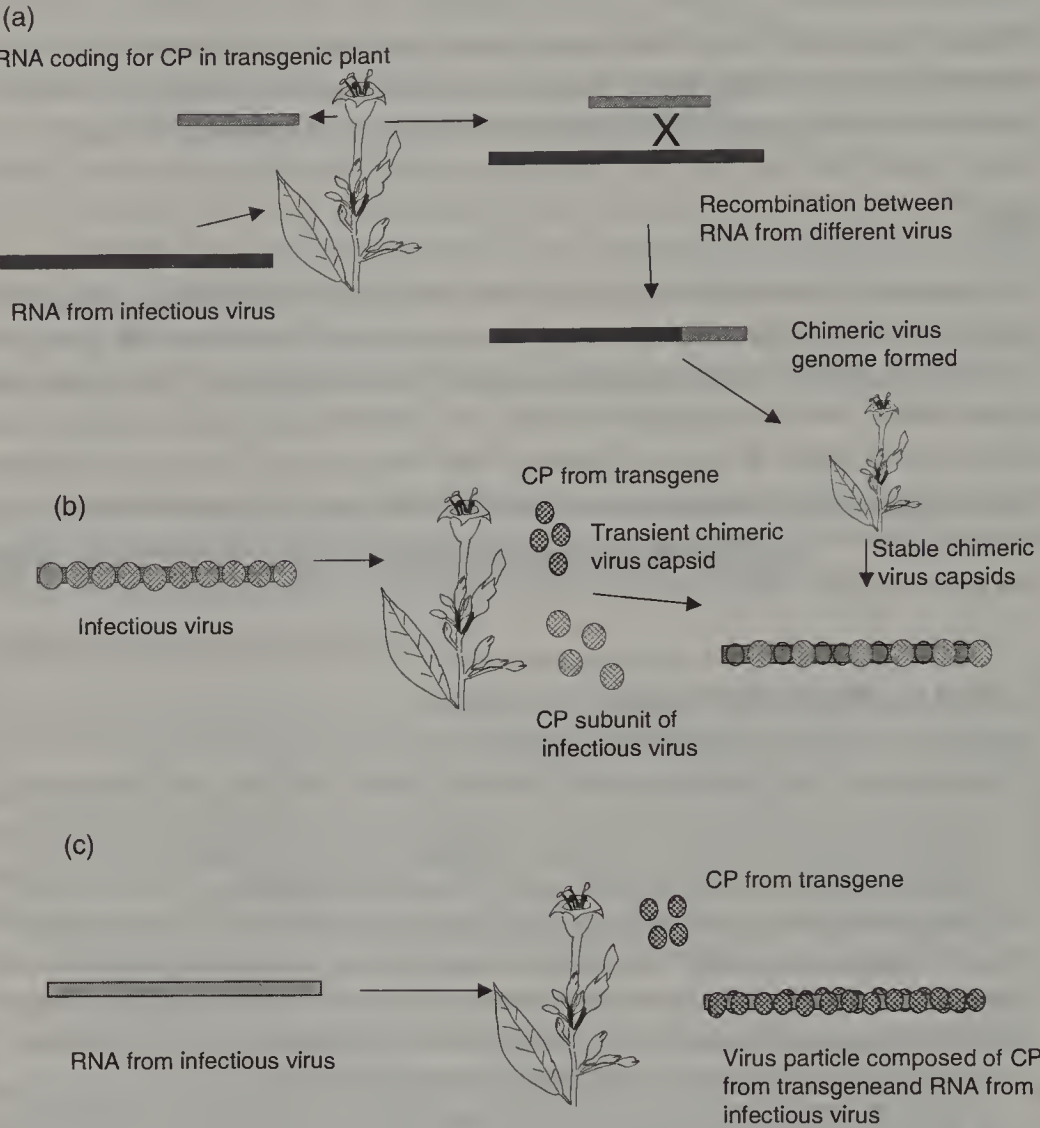
It seems that CPMR works very well for many of the constructs, but deliberate expression of both RNA and protein, to bring about the coat protein (CP) effect, leads to two levels of interaction where problems could be encountered. There is a danger of RNA recombination (Figure (a)) and when expressing a coat protein, a serious consideration has been whether a phenomenon known as ‘transcapsidation’ (which is also called ‘heterologous encapsidation’, ‘transencapsidation’ or ‘genome masking’) could occur. The danger is that the CP protein produced from the transgene could form chimeric particles with invading viruses CP (Figure (b)) or package the nucleic acid of compatible viruses (Figure (c)).

This is important because coat proteins may affect a range of virus properties, e.g. protection of the nucleic acid, vector transmission and specificity, systemic invasion and symptom expression.

Many studies have been carried out to investigate the possibility of transcapsidation. Chimeric particles were first demonstrated in a study where a plant producing transgenic TMV coat protein was challenged with a mutant strain defective in CP production. The transgenically produced CP was able to package the defective RNA. Phenotypic mixing has also been detected in transgenic potato plants producing the coat protein of PVY strain N. When the plants were infected with PVY strain O, as many as 75% of the PVY-O particles were found to contain transgenic PVY-N coat proteins. Mixing of coat proteins has also been shown to occur between two unrelated viruses, as was the case when transgenic plants producing the coat protein of alfalfa mosaic *alfamovirus* (AlMV) were infected with cucumber mosaic *cucumovirus* (CMV).

BOX 8.1 Continued

The biological significance of these observations is still under debate as many other studies have been carried out that fail to show transcapitation. However,



Recombination, transcapitation and chimeric particles.

perhaps the most important type of observation as far as plant biotechnology is concerned, is that CP mixing has been demonstrated to restore aphid transmissibility to a non-aphid transmissible isolate of zucchini yellow mosaic *potyvirus* (ZYMV). This happened after phenotypic mixing with a transgenic CP of plum pox *potyvirus* (PPV). Again, contradictory results have been obtained from field studies. Curcubit plants expressing the CP of a highly aphid-transmissible strain of CMV were inoculated with non-aphid-transmissible strains of the same virus. No aphid-mediated transmission of the second strain to control plants was detected. Apparently, in this case, the transgenically produced CP did not mediate transfer.

BOX 8.1 Continued

It is possible that the synthesis of coat protein leading to transcapsidation could alter virus transmissibility, but it could be argued such phenomenon already happen in nature. In many cases this is probably true, although genetic engineering can produce situations that do not normally occur in nature. Problems may arise if, for instance, coat protein is produced in tissues where the virus is not normally found. This is the case for potato leafroll *polerovirus* in potato, which is normally expressed in vascular tissues during infections. However, with constitutive expression of a cot protein transgene in potato the coat is expressed all over the plant and will then be exposed to other viruses it does not normally meet.

The frequency of transcapsidation is likely to be very low, and (as there is no exchange of genetic material) effects, such as changes of vector, will be transitory. Nevertheless, some effort has gone into removing the potential problem from more recent constructs. Several strategies have been adopted to reduce the likelihood of transcapsidation. In one approach the CP sequences have been mutated, either to remove sequences that may be involved in transfer (such as the sequences in potyvirus CP required for aphid transfer) or to truncate the sequence to the minimum size required for protection. Another approach, PTGS, fully precludes protein expression; constructs are being generated that do not allow any protein expression, but rely on gene silencing to produce resistance.

Reliance on the RNA-based approaches to viral resistance also has some potential problems—in particular, recombination could occur between different virus genomes. The most likely mechanism is one of template-switching. During replication of the virus genome, the viral RNA polymerase moves along the RNA sequence (the template) synthesising the complementary RNA strand. It has been argued that, in the replication of some multipartite viruses, part of the process involves moving from one RNA fragment to others within the genome. In essence, the viral polymerase pauses during replication then switches to the different template and resumes synthesis. Potentially, this could be a very serious problem if the polymerase was to switch to the RNA of another virus, as this may result in the production of a recombinant virus. In nature, recombination is quite a common phenomenon within virus groups such as luteovirus and tobnavirus, but it has also been shown to occur between totally unrelated viruses. The risk of recombination should be minimised. Several possibilities have been suggested. It is possible to include stop codons in the sequence chosen to ensure that proteins can not be expressed from the sequence. This would also make it unlikely that a viable recombinant virus could form as a result of recombination between the transgenic sequence and any wild-type virus. Other approaches, such as using short sequences that may prime the PTGS, may be considered.

Summary

In this chapter, pathogen-derived resistance has been considered in some detail. Some attempt has been made to explain the effects at the protein level by examining the work done with TMV and other related viruses. PTGS has also been raised as a mechanism of natural virus protection and not just an issue in transgenesis. Some transgenic plants

have been commercialised with varying degrees of success. The transgenic papaya fruit lines have been welcomed in Hawaii as a saviour for the small farmer, whereas the NewLeaf transgenic potato lines currently seem to have been rejected or at least removed from sale in mainland USA.

The future of PDR-type resistance will depend on the public acceptance of these crops as much as whether they provide durable resistance. Other approaches to virus resistance are being considered. Enhancement of the PTGS system is an obvious strategy for virus resistance. Chapter 7 considered the possibility of activating generalised resistance pathways, and this potential route for virus resistance is being explored. It has been known for some time that tobacco strains containing the N gene are resistant to infection by TMV. It is now apparent that the product of the N gene belongs to the toll-interleukin-1 receptor (TIR), nucleotide-binding domain (NB), leucine-rich repeats (LRR) class of R genes (TIR-NBS-LRR).

Further reading

Virus groups and expression patterns

Bustamante, P. I. and Hull, R. (1998) Plant virus gene expression strategies. *Electronic Journal of Biotechnology*—web-link 8.1:

<http://www.ejb.org/content/vol1/issue2/full/3/index.html>

Lomonosoff, G. P. and Johnson, J. E. (1991). The synthesis and structure of comovirus capsids. *Progress in Biophysics and Molecular Biology*, 55, 107–37.

Melcher, U.—web-link 8.2:

<http://www.opbs.okstate.edu/~melcher/mg/MGW3/MG33322.html>

Rybicki, E.—web-link 8.3:

<http://www.uct.ac.za/microbiology/tutorial/RNA%20virus%20replication.htm>

Non-GM approaches to virus resistance

Fenby, N. S., Scott, N. W., Slater, A. and Elliott, M. C. (1995). PCR and non-isotopic labelling techniques for plant virus detection. *Cellular and Molecular Biology*, 41, 639–52.

Fuchs, M., Ferreira, S. and Gonsalves, D. (1997). Management of virus diseases by classical and engineered protection. *Molecular Plant Pathology On-Line*—web-link 8.4: <http://www.bspp.org.uk/mppol/1997/0116fuchs>

PDR-protein and RNA effects

Beachy, R. N. (1997). Mechanisms and applications of pathogen-derived resistance in transgenic plants. *Current Opinion in Biotechnology*, 8, 215–20.

Gonsalves, D. and Slightom, J. L. (1993). Coat protein-mediated protection: analysis of transgenic plants for resistance in a variety of crops. *Seminars in Virology*, 4, 397–405.

Gene silencing

Mlotshwa, S., Voinnet, O., Mette, M. F., Matzke, M., Vaicijeret, H., Ding, S. W. and Pruss, G. and Vance, V. B. (2002). RNA silencing and the mobile silencing signal. *Plant Cell* (Suppl.), S289–S301.

Waterhouse, P. M., Wang, M. B. and Lough, T. (2001). Gene silencing as an adaptive defence against viruses. *Nature*, 411, 834–42.

Risk

Falk, B. W. and Bruening, G. (1994). Will transgenic crops generate new viruses and new diseases. *Science*, 263, 1395–6.

Greene, A. E. and Allison, R. F. (1994). Recombination between viral RNA and transgenic plant transcripts. *Science*, 263, 1423–5.

Tepfer, M. (1993). Viral genes and transgenic plants. *Biotechnology*, 11, 1125–32.

Commercialisation

Fuchs, M. and Gonsalves, D. (1995). Resistance of transgenic hybrid squash ZW-20 expressing the coat protein genes of zucchini yellow mosaic virus and watermelon mosaic virus 2 to infections by both potyviruses. *Biotechnology*, 13, 1466–73.

Tricoli, D. M., Carney, K. J., Russell, P. F., McMaster, J. R., Groff, D. W., Hadden, K. C., Himmel, P. T., Hubbard, J. P., Boeshore, M. L., Reynolds, J. F. and Quemada, H. D. (1995). Field evaluation of transgenic squash containing single or multiple virus coat protein gene constructs for resistance to cucumber mosaic virus, watermelon mosaic virus 2, and zucchini yellow mosaic virus. *Biotechnology*, 13, 1458–65.

9

Strategies for engineering stress tolerance

Introduction

Crop plants are subject to a range of external factors that adversely affect their growth and development. In the previous four chapters, we have considered the effects of various biological factors (weeds, pests and diseases) on crop yields, and biotechnological strategies to resist them. In this chapter, we will look at the effects of environmental conditions such as temperature, water availability and salinity on crop plants. Figure 9.1 shows a number of physical factors that may impose an abiotic stress on plants and adversely affect their growth and development. These include a number that can be grouped together as temperature stresses (heat, chilling and freezing), which in turn belong to a larger subgroup that can be categorised as stresses that result in water-deficit. The figure also emphasises the point that most abiotic and biotic stresses directly or indirectly lead to the production of free radicals and reactive oxygen species (ROS), creating oxidative stress.

The impact of abiotic stresses on crop yield compared to biotic stresses (weed, pest and disease effects) is shown in Table 9.1. One of the first things to notice is the large difference between the average yields of crops and the record yields (an indicator of the maximum yields possible under ideal conditions). It is clear from this data that the major difference between record yield and average yield is accounted for by abiotic stress. Thus, the variation in environmental conditions from one year to the next produces such a variation in yield for wheat that the average yield is only 13% of the maximum. In contrast, as discussed at the beginning of Chapter 5, the control of biotic stresses in industrialised farming is such that they tend to reduce the annual yield by a fairly stable proportion, which is generally less than the most adverse abiotic stresses. Improving the tolerance of crops to abiotic stresses could therefore enable them to maintain growth and development during the normal fluctuations of adverse conditions, and consequently buffer crops against the large swings in yield experienced from one year to the next.

In the longer term, the predicted depletion of the ozone layer and climate changes associated with global warming are likely to add to the burden of environmental stresses on crop plants, and increase the imperative to develop

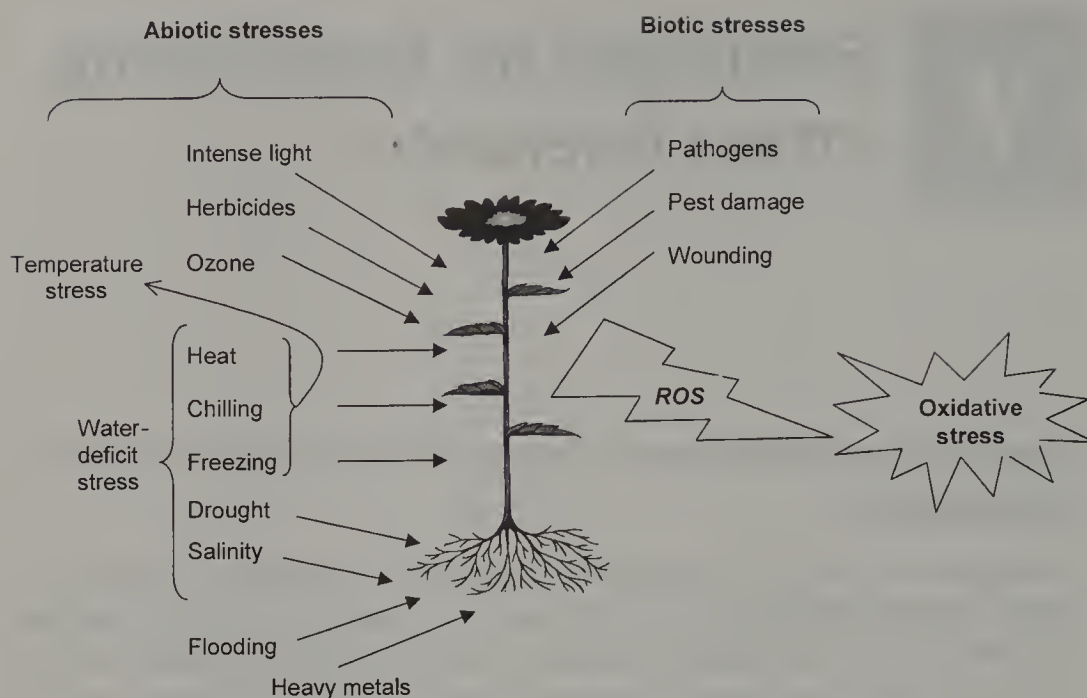


Figure 9.1 The different types of external stresses that affect plant growth and development. The figure shows some of the external stresses impacting upon plant growth and development. The stresses are grouped according to common characteristics. Thus, abiotic and biotic stresses are defined, and the range of abiotic stresses resulting in water deficit are shown. The figure emphasises the point that virtually all stresses result in the production of reactive oxygen species (ROS) and this creates oxidative stress.

stress-tolerant varieties. Furthermore, there is increasing pressure to extend the area of crop cultivation to environments which are not optimal for the growth of major crops (e.g. desert or high-salt conditions). The development of stress-tolerant plants is therefore a major goal of crop biotechnology, and one that is likely to become increasingly important.

The nature of abiotic stress

When discussing the subject of stress tolerance, it is necessary first to try to define stress in relation to plant physiology. Plants are subject to many types of fluctuation in the physical environment. Many of the strategies used by animals to avoid the effects of these fluctuations are not available to plants, because of the sessile nature of their growth habit. Plants therefore depend largely upon internal mechanisms for tolerating variations in the external environment. Not all such fluctuations present a stress to plants, since they are able to cope with normal variation by virtue of their plasticity (the ability to alter metabolic profiles and developmental trajectories in response to external conditions—see Chapter 2). Thus, plants are adapted to function in a fluctu-

Table 9.1 Average and record yields of some major crops

Crop	Record yield (kg/ha)	Average yield (kg/ha)	Average yield (% of record yield)	Average loss (% of record yield) Biotic	Average loss (% of record yield) Abiotic
Wheat	14 500	1 880	13.0	5.0	82.1
Barley	11 400	2 050	18.0	6.7	75.4
Soybean	7 390	1 610	21.8	9.0	69.3
Corn	19 300	4 600	23.8	10.1	65.8
Potato	94 100	28 300	30.1	18.9	54.1
Sugar beet	121 000	42 600	35.2	14.1	50.7

Data from Bray *et al.* (2000).

ating environment, and normal external changes are countered by internal changes without detriment to growth and development. It is only acute or chronic extremes of environmental condition that lead to environmental stress which has the potential to cause physical damage to the plant. Note also that the ‘extremes of environmental condition’ are defined by the normal environment of the particular plant species in question—desert plants do not prosper in the ‘mild’ climate of the temperate regions.

Given the range of abiotic stresses to which plants are exposed, it might be thought that a wide range of different strategies would be required to engineer particular types of stress tolerance. However, many different stresses cause similar types of damage to plants. This chapter will concentrate on the two major types of damage that result from a variety of different stresses. One of these is the damage that results from water deficit caused by a number of different environmental conditions, including drought, salinity, heat and cold (see Figure 9.1). The other type of damage results from the production of reactive oxygen species (ROS), causing chemical damage to the cellular constituents of plants. Most abiotic and biotic stresses cause the production of reactive oxygen species either directly, or indirectly.

The nature of water-deficit stress

Several of the physical changes shown in Figure 9.1 result in a water deficit, i.e. a situation in which the demand exceeds the supply of water. The supply is determined by the availability of water in the soil to the roots. The demand for water is determined mainly by the plant transpiration rate. Transpiration, the loss of water from the aerial parts of the plant via the leaf stomata, is required to dissipate the energy received from solar radiation and the ambient air

temperature. Thus, transpiration cause leaves to cool relative to the ambient temperature when the environmental energy load on the plant is high. The rate of transpiration is also affected by the relative humidity and wind speed.

Two parameters are used to describe the water status of plant cells and tissues: water potential and relative water content. Water potential is measured in units of pressure (M pascal or MPa) and is similar in concept to electrical potential—current flows from a compartment of high potential to one of low potential if a connection is made between them. Thus, water is driven through the plant from the soil to the atmosphere by the difference in water potential between the atmosphere (very low potential) and the soil (relatively high potential when wet). As water transpires from the leaf, leaf water potential is reduced; but if water is available in the soil, water will flow into the leaf to replenish the loss. However, if the soil water potential starts to reduce (because the soil is drying out, for example), a deficit between the supply from the root and demand from the leaf is created. Clearly, a continued water deficit would lead to dehydration and death, but plants can respond to water deficit in a number of different ways. One way is to reduce the water potential in the leaf in order to create the necessary gradient to maintain the flow of water from soil.

Two major factors determine water potential: the concentration of solutes that creates the solute- or osmotic potential; and the physical pressure of the cell or tissue boundaries that forms the pressure- or turgor potential. The osmotic potential is generally lower (more negative) than the water potential, and the turgor potential is the difference between them. Leaf turgor is lost (turgor potential = 0) when the leaf water potential drops to a value equal to the osmotic potential. As leaf turgor falls, the stomata close to reduce transpiration. When turgor is lost, the leaf cells collapse and the leaf wilts. These effects have the result of reducing water loss from the leaf, by closure of the stomata, and because wilting or rolling up of leaves reduces their exposure to sunlight.

However, the reduction in stomatal conductance also causes a reduction in photosynthetic assimilation, thus reducing growth. There may also be an increase in leaf temperature to a level that could cause heat damage to the leaf. The maintenance of turgor and transpiration are therefore important in order to maintain the growth of plants under drought stress. Turgor could be sustained either by keeping the leaf water potential high (by maintaining water uptake from the drying soil) or by reducing the osmotic potential. The latter could be achieved by solute accumulation (remember that an increase in solute concentration creates a more negative osmotic potential, i.e. a steeper gradient for the movement of water)—a process called osmotic adjustment.

Different abiotic stresses create a water deficit

Drought is, by definition, a water-deficit stress, because the environmental conditions either reduce the soil water potential and/or increase the leaf water

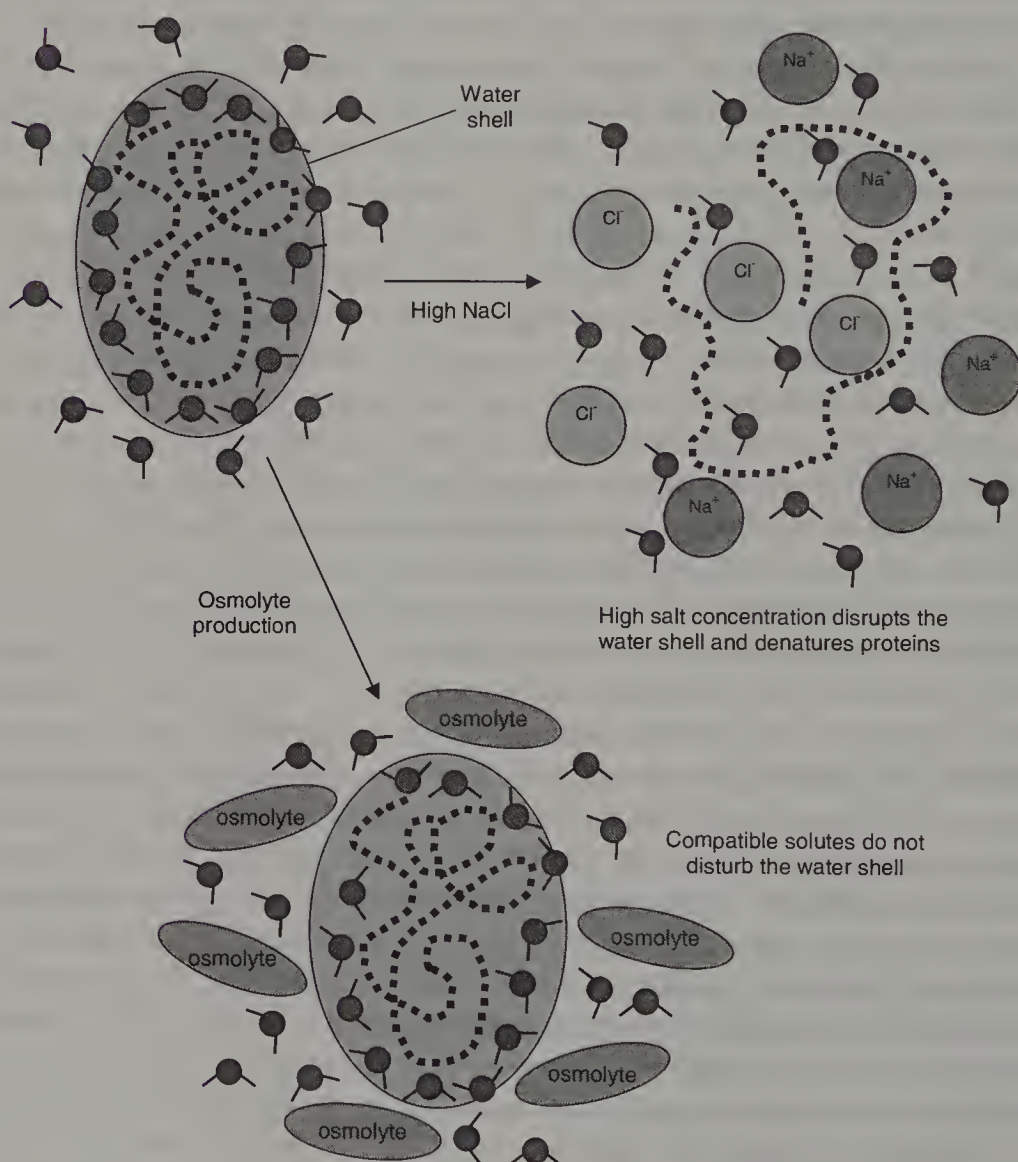
potential due to hot, dry or windy conditions. High-salt conditions can result in water deficit because the soil water potential is decreased (because the osmotic potential of a salt solution is lower than that of water), making it more difficult for roots to extract water from the environment. As stated above, high ambient temperatures cause increased water loss by evaporation, so heat stress creates a water deficit. However, freezing temperatures also cause osmotic stress because the formation of ice crystals in the extracellular space reduces the water potential and results in the efflux of intracellular water. In general, these various causes of water deficit result in the efflux of cellular water, leading to plasmolysis and eventually cell death. Water deficit is a particular problem to plant cells because it inhibits photosynthesis via its effect on the thylakoid membranes. Water deficit is also potentially damaging to all cells because of the increase in the concentration of toxic ions and the loss of the protective hydration 'shell' around vulnerable molecules (Box 9.1).

As described above, cells are able to accommodate some fluctuation in the water potential by osmotic adjustment—increasing the solute concentration and hence reducing the osmotic potential. However, as shown in Box 9.1, an increase in ions such as Na^+ and Cl^- would be counterproductive, given their effects on vulnerable molecules. Instead, plant cells respond to osmotic stress by producing non-toxic compounds called 'compatible solutes' (or 'osmolytes', or 'osmoprotectants') to reduce the osmotic potential. It has been proposed that one reason for the non-toxicity of these compounds is that they can accumulate to high intracellular concentrations without disrupting the hydration shell around proteins and membranes (see Box 9.1).

Figure 9.2 shows a range of compatible solutes produced by plants. These fall into two broad classes: sugars and sugar alcohols; and zwitterionic compounds (i.e. carrying a positive and negative charge). The former class includes sugar alcohols such as mannitol, sorbitol, pinitol and D-ononitol, and oligosaccharides such as trehalose and fructans. (The production of some of these compounds in transgenic plants is described in more detail in Chapter 11). The latter class includes amino acids such as proline, and quaternary ammonium compounds such as glycine betaine. Different plant species tend to produce different compatible solutes. For example, mannitol is produced in large amounts by celery in response to salt stress, whereas spinach accumulates glycine betaine. In fact, spinach cells grown in a high-salt medium accumulate glycine betaine to a concentration of 300 mmol l^{-1} in the cytosol. This maintains the osmotic potential of the cytoplasm relative to the apoplast (the connected extracellular space outside the cell; the symplast is the continuous intracellular compartment within protoplasts, connected by plasmodesmata) whilst keeping the cytoplasmic concentration of Na^+ and Cl^- ions low ($<50 \text{ mmol l}^{-1}$) and storing high concentrations of Na^+ (200 mmol l^{-1}) and Cl^- (150 mmol l^{-1}) ions in the vacuole. It is particularly significant that certain crop plants (e.g. rice and tobacco) lack significant amounts of any of these major classes of osmoprotectants. Basic strategies for engineering resistance to water-deficit stress have therefore focused on the production of

BOX 9.1 The water shell around macromolecules

The precise folding of soluble proteins is determined not only by intramolecular interactions between residues in the molecule, but by its interactions with the surrounding solution. For example, hydrophobic side chains are normally coiled within the interior of the molecule in order to minimise the interaction with the surrounding water, whereas polar and charged groups are often positioned around the external surface of the molecule. The water molecules around the protein form a highly ordered (and therefore low entropy) structure called a 'hydration' or 'water shell', which stabilises the protein and serves to buffer it from interactions with polar solutes.



The water shell around a protein is disrupted when the concentration of ions is increased, and the protein is denatured. An equivalent concentration of a compatible solute does not disrupt the water shell and the protein is not denatured.

BOX 9.1 Continued

Increases in the ionic concentration of the medium surrounding the protein will disrupt the structure of the protein, because Na^+ and Cl^- ions can effectively penetrate the hydration shell, and interfere with the non-covalent bonding that contributes to the stability of the protein.

On the other hand, compatible solutes are excluded from the protein water shell, so the protein does not come into contact with the solute, and is not denatured.

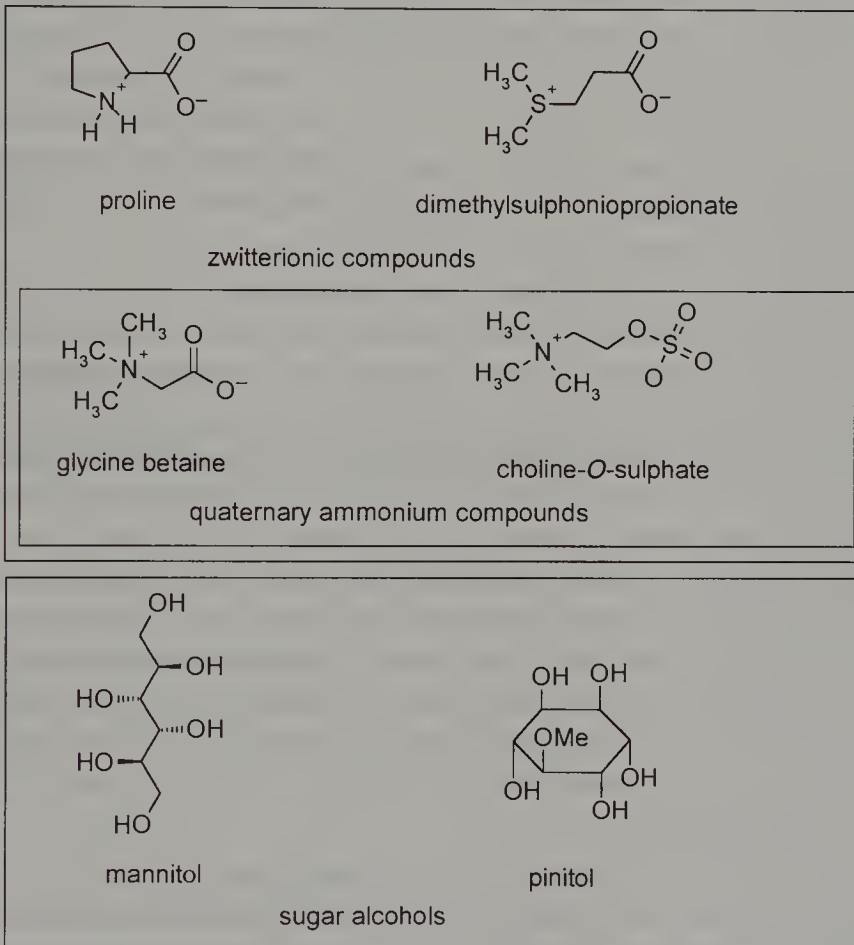


Figure 9.2 Compatible solutes. The chemical structures of a number of compatible solutes are shown, grouped together as zwitterionic or sugar-based compounds.

osmoprotectants as a mechanism for overcoming the osmotic stress generated by water deficit. These strategies have generally required the determination of the biosynthetic pathways for various osmoprotectants, isolation of the relevant genes and appropriate engineering of constructs to target gene expression and protein destination.

CASE STUDY 1 Glycine betaine production

Glycine betaine is a quaternary ammonium compound found in at least 10 flowering plant families and also in marine algae. The charged groups mean it is extremely soluble in water, and is electrically neutral over a wide range of physiological pH values. The three methyl groups that comprise the quaternary ammonium group enable it to interact with hydrophobic as well as hydrophilic molecules. Thus, glycine betaine appears to stabilise proteins and membranes via these interactions, in addition to acting as a cellular osmolyte.

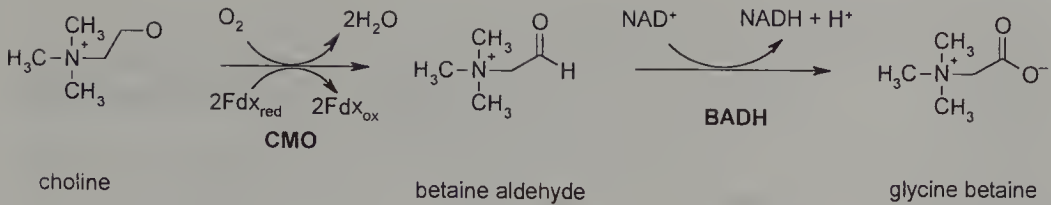
The synthesis of glycine betaine in plants has been studied mainly in species of *Chenopodaceae* (spinach, beet, etc.). In these species, glycine betaine is formed by a two-stage oxidation of choline, with the first step catalysed by choline monooxygenase (CMO) (Figure 9.3). CMO encodes an unusual ferredoxin-dependent enzyme with an iron–sulphur cluster similar to that found in the Rieske protein of the cytochrome *b₆/f* complex. This enzyme is localised in the chloroplast stroma, where its activity is strongly light-dependent due to its requirement for reduced ferredoxin generated by photosynthetic electron transport. The second step in glycine betaine synthesis is catalysed by betaine aldehyde dehydrogenase (BADH), which is also stromal in *Chenopodaceae*. CMO and BADH are both induced by osmotic stress and have both been cloned from several plants.

The biosynthetic pathway of glycine betaine is different in *Escherichia coli* compared to plants, in that the production of betaine aldehyde from choline is catalysed by choline dehydrogenase (CDH), but the subsequent step still involves a betaine aldehyde dehydrogenase enzyme. In contrast, the synthesis of glycine betaine in *Arthrobacter globiformis* involves a single enzyme, choline oxidase (encoded by the *codA* gene), which catalyses the reaction directly from choline to glycine betaine (Figure 9.3), producing hydrogen peroxide.

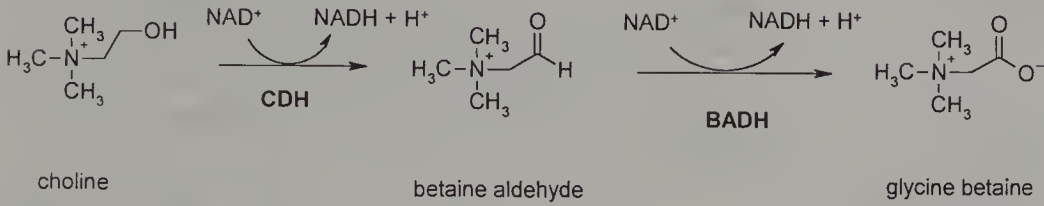
Attempts have been made to manipulate glycine betaine production by transformation with the relevant plant, *E. coli* or *A. globiformis* genes (Table 9.2). Transformation of plants with the betaine aldehyde dehydrogenase genes from plants or *E. coli* permitted the accumulation of glycine betaine when plants were supplied with betaine aldehyde, and conferred resistance to this toxic compound (see Chapter 4). However, plants did not accumulate glycine betaine in the absence of betaine aldehyde, presumably because the conversion of choline to betaine aldehyde does not occur in plants that do not normally accumulate glycine betaine.

Transformation of tobacco with the plant CMO gene targeted to the chloroplast resulted in very low levels of glycine betaine accumulation (Table 9.2). It has been suggested that this is a result of the very low pool of choline present in tobacco cells, or it could be attributable to the lack of betaine aldehyde dehydrogenase activity in the tobacco chloroplasts. Tobacco plants transformed with the *E. coli* choline dehydrogenase gene, *betA*, also failed to accumulate glycine betaine, even though this enzyme can also catalyse the conversion of betaine aldehyde to betaine. However, in this case, the gene was not targeted to the chloroplast, which might explain this failure. But there was an observed increase in the salt tolerance of these transgenic plants, that can not be accounted for by the accumulation of glycine betaine to high concentrations.

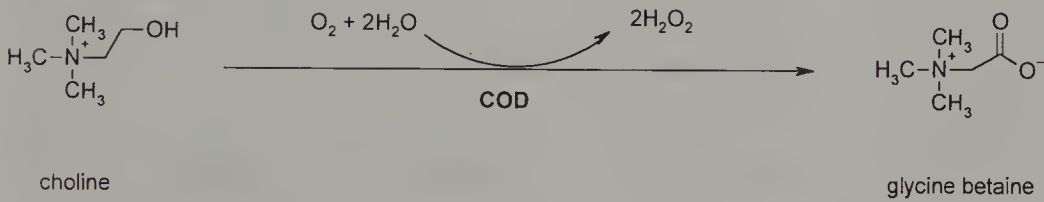
Higher levels of glycine betaine in transgenic plants have been obtained using the choline oxidase gene from certain *Arthrobacter* species particularly when transformed into plant



1. Plants



2. *E. coli*



3. *Arthrobacter globiformis*

Figure 9.3 Pathways to glycine betaine. The different pathways from choline to glycine betaine in plants, *E. coli* and *Arthrobacter globiformis* are compared.

species (e.g. *Arabidopsis* and rice) that have significantly higher choline pools than does tobacco. In a number of cases where glycine betaine accumulation has been achieved, the tolerance to various water-deficit stresses has been determined, including salt, chilling, freezing heat and drought (see Table 9.2). It is apparent that production of glycine betaine in transgenic plants can improve tolerance to a range of osmotic stresses under the test conditions applied, validating the initial concept that osmoprotectants can enhance tolerance to a variety of stresses that cause water deficit. However, it is not entirely clear what the protective effect of the glycine betaine is. The levels of glycine betaine produced in most of these transgenic plants is relatively low (about 10% of the levels that accumulate in plants such as spinach) and is not sufficient to account for the improved osmotic stress tolerance by osmotic adjustment. It may be that the glycine betaine exerts a protective effect on vulnerable macromolecular structures at a lower concentration than its effect as an osmolyte.

In the same way that production of glycine betaine increases tolerance to a range of stresses that cause water deficit, other osmoprotectants have been produced in transgenic plants to test the feasibility of this approach. Table 9.3 shows the range of genes and products that have been used to test this strategy in a range of crop plants.

Table 9.2 Glycine betaine production in transgenic plants

Transgene	Host plant	Accumulation of glycine betaine	Stress tolerance tested
Barley <i>badh</i> (betaine aldehyde dehydrogenase)	Tobacco peroxisome	Not tested	Not tested
Spinach <i>badh</i>	Tobacco chloroplast	20 $\mu\text{mol g}^{-1}$ FW (in 5 mmol l^{-1} betaine aldehyde)	Not tested
Spinach <i>cmo</i> (choline monooxygenase)	Tobacco chloroplast	<0.05 $\mu\text{mol g}^{-1}$ FW	Not tested
<i>E. coli betB</i> (betaine aldehyde dehydrogenase)	Tobacco chloroplast or cytosol	Not tested	Not tested
<i>E. coli betA</i> (choline dehydrogenase)	Tobacco cytosol (membranes)	Not detected	Salt
<i>betA/betB</i>	Tobacco	0.035 $\mu\text{mol g}^{-1}$ FW	Chilling Salt
<i>betA</i>	Rice	5.0 $\mu\text{mol g}^{-1}$ FW	Drought Salt
<i>A. globiformis codA</i> (choline oxidase)	<i>Arabidopsis</i> chloroplast	1.2 $\mu\text{mol g}^{-1}$ FW	Salt Chilling Freezing Heat Strong light
<i>codA</i>	Rice	5.3 $\mu\text{mol g}^{-1}$ FW	Salt Chilling
<i>A. pascens cox</i> (choline oxidase)	<i>Arabidopsis</i>	19 $\mu\text{mol g}^{-1}$ DW	Freezing Salt
<i>cox</i>	<i>Brassica napus</i>	13 $\mu\text{mol g}^{-1}$ DW	Drought Salt
<i>cox</i>	Tobacco	13 $\mu\text{mol g}^{-1}$ DW	Salt

FW, fresh weight DW, dry weight.

What is clear from this table is the existence of a range of effective transgenes for the production of osmoprotectants in plants, and that these generally have a measurable effect on the tolerance of the plant to various stresses that cause water deficit. In some cases, tolerance to more than one stress has been demonstrated for the same transgenic plant, confirming the view that generic strategies to combat water-deficit stress could provide tolerance against a range of different stresses. However, several aspects of osmolyte

accumulation still require further investigation and explanation. One point to note from Tables 9.2 and 9.3 is that the level of accumulation of most of the compatible solutes, apart from proline and D-ononitol, is considerably lower than could account for their protective effects by osmotic adjustment. The reasons for stress tolerance may be different in each case, with some compounds like glycine betaine acting primarily to stabilise macromolecular structures, whilst other compatible solutes such as mannitol, sorbitol and proline are effective scavengers of reactive oxygen species (see below). Another point to make is that the tests of stress tolerance of these transgenic plants are often performed under highly controlled laboratory conditions that present an acute osmotic shock rather than the chronic water-deficit stresses more generally experienced in the field. It should be noted that few of these transgenic plants have been tested for stress tolerance in the field. Another consideration for developing these strategies further is that the accumulation of trehalose and sorbitol was found to inhibit plant growth.

A final note of caution relates to the assumption that increasing osmolyte accumulation in transgenic plants to levels that can maintain cell and tissue turgor pressure under water-deficit conditions will improve crop yield under drought conditions. Under conditions of reduced soil water, the maintenance of leaf turgor and hence transpiration in order to sustain growth will reduce soil water further and hasten the dehydration of the leaves. What needs to be understood is that under conditions where the water deficit is severe enough to threaten plant survival, crop yields are so low that even large increases in fractional yield are of little practical benefit to the grower.

Targeted approaches towards the manipulation of tolerance to specific water-deficit stresses

The production of osmoprotectants is not the only possible strategy to engineer resistance to the water-deficit stresses shown in Figure 9.1. Individual stresses may be susceptible to specific mechanisms to improve tolerance. Some of these will be described in the following section.

Alternative approaches to salt stress

Salt tolerance is increasingly becoming a major target for crop improvement as substantial areas of irrigated land are damaged by the accumulation of salt. Furthermore, the pressure for land has made it necessary to consider the possibility of growing crops in more saline conditions, with poorer quality water. As described above, saline conditions lead to osmotic stress by preventing water uptake by the roots and water efflux from the cells. However, the accumulation of Na^+ and Cl^- ions in the cytoplasm may also have direct toxic

Table 9.3 Transgenic plants engineered to synthesise osmoprotectants other than glycine betaine

Osmoprotectant	Transgenes	Crop plants	Accumulation	Stress tolerance
Proline	Mothbean <i>P5CS</i>	Tobacco		Salt
	(Pyrroline carboxylate synthetase)	Rice		Drought, salt, oxidative
	<i>P5CS_{F127A}</i>	Soybean		Osmotic, heat
	(feedback inhibition insensitive)	Tobacco	4 mg g ⁻¹ FW	Salt
Mannitol	Anti-ProDH			
	(proline dehydrogenase)	<i>Arabidopsis</i>	0.6 mg g ⁻¹ FW	Salt, freezing
	<i>E. coli mt1D</i>			
	(mannitol-1-phosphate dehydrogenase)	<i>Arabidopsis</i>	10 µg g ⁻¹ FW	Salt
Sorbitol		Tobacco	6 µmol g ⁻¹ FW	Salt
	Apple <i>s6pdlh</i>			
	(sorbitol-6-phosphate dehydrogenase)	Tobacco		Oxidative stress
		Persimmon	61.5 µmol g ⁻¹ FW	Salt

Trehalose	Yeast <i>tps1</i> (trehalose-6-phosphate synthase, T-6-PS) <i>E. coli otsA + otsB</i> (T-6-PS and T-6-P phosphatase)	Tobacco Potato	3.2 $\mu\text{g g}^{-1}$ DW	Drought Drought
		Tobacco	90 $\mu\text{g g}^{-1}$ FW	Drought
D-Ononitol	Ice plant <i>imtl</i> (Myo-inositol o- methyltransferase)	Tobacco	35 $\mu\text{mol g}^{-1}$ FW	Drought, salt
Fructans	<i>B. subtilis sacB</i>	Tobacco Sugar beet	0.35 mg g^{-1} FW 5 mg g^{-1} DW	Drought Drought
Glutamine	GS2 (chloroplastic glutamine synthetase)	Rice		Salt, chilling
Osmotin	<i>Osm1-Osm4</i> (protein accumulation)	Tobacco		Drought, salt

FW, fresh weight; DW, dry weight.

effects by inhibiting protein synthesis, photosynthesis and susceptible enzymes. Thus, strategies for engineering water stress tolerance via the production of compatible solutes may provide protection against the osmotic effects of saline conditions, but not against ion toxicity. Additional approaches to minimise the toxic effects of specific ions may also be required.

CASE STUDY 2 **Vacuolar Na/H antiport in transgenic plants improves salt tolerance**

One approach to tackling this problem is to look at the response of plants to salt stress. Some halophytes actually excrete salt via specialised glands on their leaf surfaces. However, it is more common for plants to avoid sodium ion accumulation in the cytoplasm by transporting them into the vacuole. Thus, one strategy would be to improve the transport of ions out of the cytoplasm and into the vacuole, particularly using older leaves as a sink and avoiding salt accumulation in sensitive tissues such as meristems.

In order to put this into practice, it is necessary to consider the mechanism of ion transport into the vacuole. Since this transport is working against a concentration gradient, it requires the input of energy. This is achieved by coupling the transport protein to a proton pump, transporting H^+ ions in the opposite direction (hence it is called an antiport protein). The vacuolar Na/H antiport protein AtNHX1 of *Arabidopsis* has been extensively studied and is known to be coupled to proton pumps such as AVP1, a vacuolar H^+ -translocating pyrophosphatase. An analogy that has been used is to compare AtNHX1 with a revolving door, and AVP1 as providing the energy for the door to spin. To increase the traffic through the membrane, one could therefore either increase the number of doors, or provide more energy for the existing doors to spin faster.

The first approach has been successfully used to engineer salt tolerance in tomato and canola plants by transformation with the *Arabidopsis* AtNHX1 antiport protein gene. These normally susceptible plants were able to grow in a high-salt environment to fertile maturity. Tomato plants grown in 200 mmol/l NaCl accumulated salt in the leaves to a level 20-fold higher than normal, but, significantly, the fruits had a low sodium and chloride content. The transgenic canola plants accumulated up to 6% sodium, but the oil quality and yield of the seeds were normal.

The other approach has been to overexpress the encoding gene AVP1, initially in *Arabidopsis*, in order to increase the proton-pumping potential of the vacuole, and hence its ability to transport sodium. This has improved not only the salt tolerance of these experimental plants, but also the drought tolerance, since the altered ion balance has enabled the plants to retain more water.

Alternative approaches to cold stress

Different plants vary enormously in their ability to withstand cold and freezing temperatures. Most tropical plants have virtually no capacity to survive freezing conditions. On the other hand, many temperate plants can survive

a range of freezing temperatures from -5 to -30°C depending upon the species. Plants from colder regions routinely withstand temperatures even lower than this. It is known that plants are better able to withstand cold or freezing stress if they first undergo a period of cold acclimation, at a low but non-freezing temperature. For example, wheat plants grown at normal warm temperatures are killed by freezing at -5°C , but after a period of cold acclimation when the plant grows at temperatures below 10°C , they can survive freezing temperatures down to -20°C .

Plants differ in their ability to withstand cold or freezing conditions and cold tolerance is one of the traits that plant breeders have selected for over many centuries. However, there has been little improvement in the cold tolerance of major crop species over the past two decades by conventional breeding, prompting the search for molecular solutions to this problem.

CASE STUDY 3 **The COR regulon**

One approach has been to study the mechanisms of freezing resistance that exist in some plant species. During the period of acclimation, plants produce a number of cold-induced proteins that are assumed to play a role in the subsequent cold-resistance. About 50 cold-induced proteins have been identified in different plant species. These fall into a small number of groups, but they all share the property of being extremely hydrophilic. Many of them also have relatively simple amino acid compositions, with repeated motifs. Some of these groups had previously been identified as 'late embryogenesis abundant' (LEA) proteins. Other groups of proteins are encoded by a class of genes designated as *COR* (cold-responsive) genes according to their patterns of expression. The precise function of these cold-induced genes is not yet known, but it has been speculated they might contribute directly to freezing tolerance by mitigating the potentially damaging effects of dehydration associated with freezing. Overexpression or ectopic expression of these cold-induced proteins could therefore be a possible route to the specific engineering of cold or freezing stress tolerance.

There are some examples of the expression of cold-induced proteins in transgenic plants. For example, constitutive expression of the small, hydrophilic, chloroplast-targeted *COR* protein *COR15a* in *Arabidopsis* improved the freezing tolerance of chloroplasts frozen *in situ* and of protoplasts frozen *in vitro*. However, *COR15a* expression has no discernible effect on the survival of frozen plants. One explanation for this observation is that the cold-induced proteins may be targeted to different vulnerable cell components, and that they are all required to provide complete protection to the cell. By implication, many *COR* genes would need to be transformed into a transgenic crop in order to obtain an appreciable improvement in cold tolerance. As discussed in Chapter 4, insertion of several genes into a transgenic crop on the same vector is technically difficult, and the process of 'pyramiding' or 'stacking' genes by crossing transgenic lines (see, for example, Chapters 6 and 11) is time-consuming. A different strategy is therefore needed.

One solution to the problem of engineering a multigene trait has emerged following the recognition that several different cold tolerance-related genes contain a similar regulatory

element in their promoters—the CRT/DRE element. Furthermore, it has been found that the transcription factor CBF1 binds to the CRT/DRE element and activates expression of this group of genes, which comprise the COR regulon (see Box 9.2). Therefore, the strategy is to overexpress the *CBF1* gene, leading to the induction of this entire group of COR cold-tolerance genes. Transgenic *Arabidopsis* plants carrying a 35S promoter::*CBF1* gene construct have been produced. These plants express a number of COR genes without cold acclimation and have been shown to be freezing-tolerant without prior cold acclimation. As a control, transgenic plants overexpressing a single COR protein, COR15a, were found to be less freezing-tolerant than the *CBF1* plants. The interrelated nature of different stress responses was demonstrated in a similar experiment. The expression of a CBF1 homologue, DREB1A (dehydration responsive-element binding protein) under the control of a stress-induced promoter in transgenic *Arabidopsis* resulted in plants that had improved drought, salt and freezing tolerance.

Tolerance to heat stress

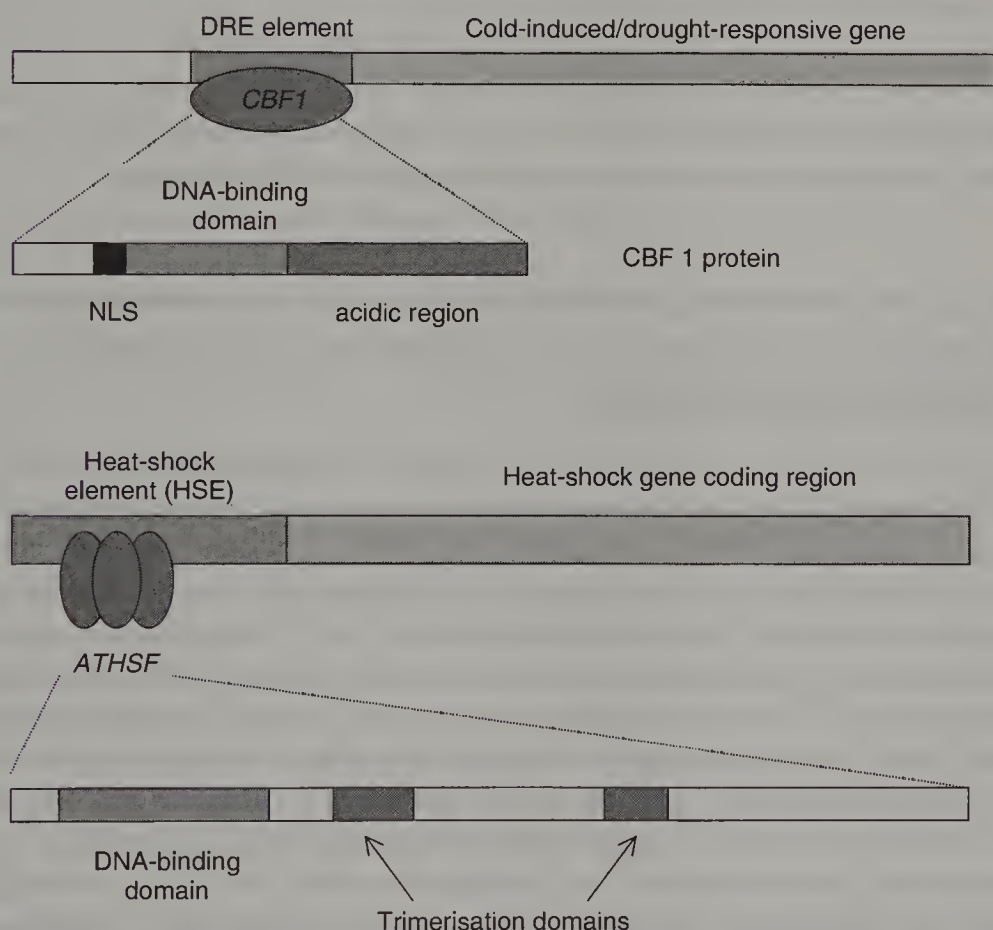
For many years it has been known that heat stress applied to a wide range of organisms induces a specific set of ‘heat shock’ proteins (HSPs); they fall into five classes, four of which are highly conserved in prokaryotes and eukaryotes. These four are categorised according to size as the HSP100, HSP90, HSP70 and HSP60 classes whose members appear to function as molecular chaperones. Some of them are expressed constitutively and are involved in normal protein synthesis and folding. Those induced by heat appear to be involved in countering the effects of heat stress by protecting or refolding denatured proteins. Their expression is induced by heat treatment and, in some cases, can be correlated with the acquisition of thermotolerance. The fifth group of several classes of small heat-shock proteins are particularly abundant in plants, but their function is not yet clear.

In a way analogous to the cold/freezing tolerance strategies, individual heat-shock proteins have been transformed into plants in order to enhance heat tolerance (Table 9.4). However, it is also known that the rapid heat-shock response is co-ordinated by a heat-shock transcription factor (HSF). This protein is expressed constitutively, but in normal conditions exists as a monomer bound to one of the HSP70 heat-shock proteins. Upon heat stress, the HSP70 dissociates and the HSF assembles into a trimer (Box 9.2) which binds to a heat-shock element (HSE) common to the promoters of heat-shock protein genes. The HSE is made up of 5-bp repeats in alternating orientations with the consensus *NGAAN*; five to seven of these repeats occur in the promoter close to the TATA box.

When the *AtHSF1* gene was overexpressed in *Arabidopsis*, the transcription factor was not active, and there was no effect on thermotolerance. However, fusion of *AtHSF1* to the N- or C- terminus of the *gusA* reporter gene produced a fusion protein that was able to trimerise in the absence of heat. Transformation of this fusion protein into *Arabidopsis* produced transgenic

BOX 9.2 COR and heat-shock regulons

A number of the cold-induced, *COR*, genes have been characterised, and the sequences of their promoters compared. One of the features of several different *COR* genes is that their promoters share a common regulatory element termed the 'C-repeat' (CRT) or 'low temperature-response element' (LTRE) which is 5 nucleotides long and has a consensus sequence of CCGAC. This element had already been linked to drought resistance and termed the 'dehydration responsive element' (DRE). The CRT/LTRE/DRE is bound by a transcription factor termed 'CBF1'. The structure of the CBF1 transcription factor is shown, indicating the nuclear localisation sequence, DNA-binding domain and an acidic region that may be involved in interactions with other proteins. *CBF1* expression is induced by cold acclimation, and leads to the expression of the *COR* genes. This group of genes, sharing a common regulatory mechanism, has been termed the 'COR regulon'.



A number of cold-induced/drought responsive genes contain the DRE sequence element in their promoters, which is bound by transcription factors of the CBF family, activating transcription. Overexpression of a single CBF gene therefore induces the expression of several cold-induced/drought responsive genes. Heat shock-induced genes also contain a heat shock element (HSE) in their promoters. Heat shock factor (HSF) proteins bind the HSE as a trimer. Overexpression of HSF encoding gene induces the expression of several heat shock proteins.

BOX 9.2 Continued

CBF1 is a member of a small gene family; *CBF1–3*. *CBF2* and *CBF3* are also transcription factors, and constitutive expression improves freezing tolerance. Expression of all three *CBF* genes is induced rapidly by low temperatures (but not by dehydration, ABA or high salt concentrations). In addition, *CBF3* overexpression results in several biochemical changes associated with cold acclimation, such as elevated levels of compatible osmolytes, proline and soluble sugars.

Although low temperature-induced gene expression, mediated by the CRT element, appears to be well conserved in plants, not all cold-induced genes (including the *CBF* genes) have the CCGAC element in their promoters. Other pathways of low temperature-induced gene expression, not mediated via CRT/*CBF*, appear then to be present in plants. Another sequence element, CCGAAA, has been identified as conferring low temperature inducibility in some genes.

The heat-shock genes also comprise a regulon. They all contain the heat-shock element (HSE) consensus sequence in their promoter regions. The transcription factor HSF binds to the HSE element as a trimer.

plants that expressed heat-shock proteins constitutively and demonstrated enhanced thermotolerance without requiring prior heat treatment.

Secondary effects of abiotic stress—the production of reactive oxygen species

Many, if not all, forms of environmental stress (abiotic and biotic) result in oxidative stress. This occurs most directly as a result of ozone pollution or ionising radiation. However, oxidative stress is a secondary effect of many type of stress, from pathogen attack to water-deficit stress. The oxidative stress arises from the production of free radicals and the subsequent ensuing cascade of reactions. Much of the oxidative stress-induced damage to cells occurs as a result of the formation of reactive oxygen species (ROS) (see Box 9.3). Figure 9.4 shows the cascade of successive reactions that form superoxide, hydrogen peroxide and hydroxyl-radical species. The ROS cause damage as a result of their reactions with cellular macromolecules. Damage to cellular membranes may result from a low concentration of hydroxyl radicals triggering a chain reaction of lipid peroxidation. Oxidative damage to proteins may involve specific amino acid modifications, fragmentation of the peptide chain, aggregation of cross-linked reaction products, altered electrical charge and increased susceptibility to proteolysis. In DNA, both the sugar and bases are susceptible to oxidation by reactive oxygen species, causing base degradation, single-strand breakage, and cross-linking to protein. These lesions in DNA can result in deletions, mutations and other damaging genetic effects.

Plants contain a number of enzymes that catalyse the cascade of reactive

Table 9.4 Transgenes used to manipulate heat tolerance

Gene	Protein	Transgenic plant
<i>AtHSF1</i>	Heat-shock transcription factor HSF1::GUS fusion	<i>Arabidopsis</i>
<i>Hsp101</i>	HSP100 class heat-shock protein	<i>Arabidopsis</i>
<i>Hsp70</i>	HSP70 class heat-shock protein	<i>Arabidopsis</i>
<i>Hsp17.7</i>	SmHSP small heat-shock protein family	Carrot
<i>TLHS1</i>	Class I smHSP	Tobacco

oxygen species and convert them to less reactive products (Figure 9.4). One key point to notice is the key role of certain enzymes in this cascade, notably superoxide dismutase, catalase and peroxidases in neutralising these reactive species. Also shown in the figure are a number of ‘antioxidant’ compounds that react with the activated oxygen compounds to produce harmless, regenerable products. These include three different vitamin compounds: β -carotene (provitamin A), ascorbic acid (vitamin C) and α -tocopherol (vitamin E). Other important antioxidants include glutathione and zeaxanthin (see Box 9.4). Central to this antioxidant system is the ascorbate–glutathione cycle (Figure 9.4). The coupling of ascorbate and glutathione redox-cycling has been most extensively characterised in the chloroplast. Chloroplasts produce superoxide and hydrogen peroxide in the light, particularly from photosystem I. The superoxide generated is itself converted into hydrogen peroxide by either spontaneous dismutation or by the activity of superoxide dismutase. The hydrogen peroxide produced is scavenged by ascorbate and the enzyme ascorbate peroxidase. The monodehydroascorbate produced from this reaction (see Box 9.4) may be regenerated in two ways. One is by reduction of NAD(P)H catalysed by monodehydroascorbate reductase. The other occurs via the dismutation of two monodehydroascorbate molecules to form ascorbate and dehydroascorbate, followed by the reduction of the dehydroascorbate by glutathione, catalysed by dehydroascorbate reductase.

From the outline of plant antioxidants and oxygen-scavenging activities given above, two general strategies for engineering tolerance to oxidative stresses are apparent—either to increase the level of enzymes that remove ROS, or to increase the level of antioxidant compounds that react with ROS.

Strategy 1: Expression of enzymes involved in scavenging ROS

Superoxide dismutases (SODs) catalyse the reaction shown in Figure 9.4, in which the superoxide radical combines with two hydrogen ions to form

BOX 9.3 Reactive oxygen species

In order to understand the significance of oxidative stress and the production of reactive oxygen species, it is necessary to review some of the basic chemistry of oxygen. Atmospheric oxygen in its ground-state has two unpaired electrons. These two unpaired electrons have parallel spins and, in consequence, oxygen is usually non-reactive to most organic molecules, which have paired electrons with opposite spins. (Oxygen cannot react with a divalent reductant unless it has two unpaired electrons with a parallel spin opposite to the oxygen. This spin restriction means that the most common mechanisms of oxygen reduction in biochemical reactions are those involving transfer of only a single electron). This means that oxygen has to be 'activated' before it can participate in reactions with organic molecules. The activation of oxygen may occur by two different mechanisms:

1. Absorption of sufficient energy to reverse the spin on one of the unpaired electrons to form the singlet state, in which the two electrons have opposite spins (see Figure 9.4). Singlet oxygen can participate in reactions involving the simultaneous transfer of two electrons, and since paired electrons are common in organic molecules, singlet oxygen is much more reactive towards organic molecules than its triplet counterpart. In photosynthetic plants, singlet oxygen is often formed by the photosystems and plays a role in metabolic reactions such as the oxidation of xenobiotics and the polymerisation of lignin. However, it may also be formed by the dysfunctioning of enzymes or electron transport systems, as a result of perturbations in metabolism caused by chemical or environmental stress.
2. The stepwise reduction of oxygen to form reactive oxygen species—superoxide (O_2^-), hydrogen peroxide (H_2O_2) and the hydroxyl radical ($OH^\cdot$).

Superoxide is a powerful oxidant or a reductant. It can oxidise sulphur, ascorbic acid or NADPH, but it can also reduce cytochrome *c* and metal ions. The reaction leading to the formation of hydrogen peroxide and oxygen is catalysed by the enzyme superoxide dismutase, but it can also occur spontaneously (see Figure 9.4).

Hydrogen peroxide is a substrate in a number of oxidation reactions catalysed by peroxidases involving the synthesis of complex organic molecules. It readily permeates membranes and is therefore not compartmentalised in plant cells. The reactivity of hydrogen peroxide is the result of its reduction by metal ions to form the highly reactive hydroxyl radical.

The hydroxyl radical oxidises organic molecules by two different reaction mechanisms: addition of the hydroxyl radical to form a hydroxylated compound; or the formation of an organic radical plus water. The organic radical can react with triplet oxygen to form a peroxy radical, which can produce another organic radical. This can proceed to generate a chain reaction in which further organic radicals are produced.

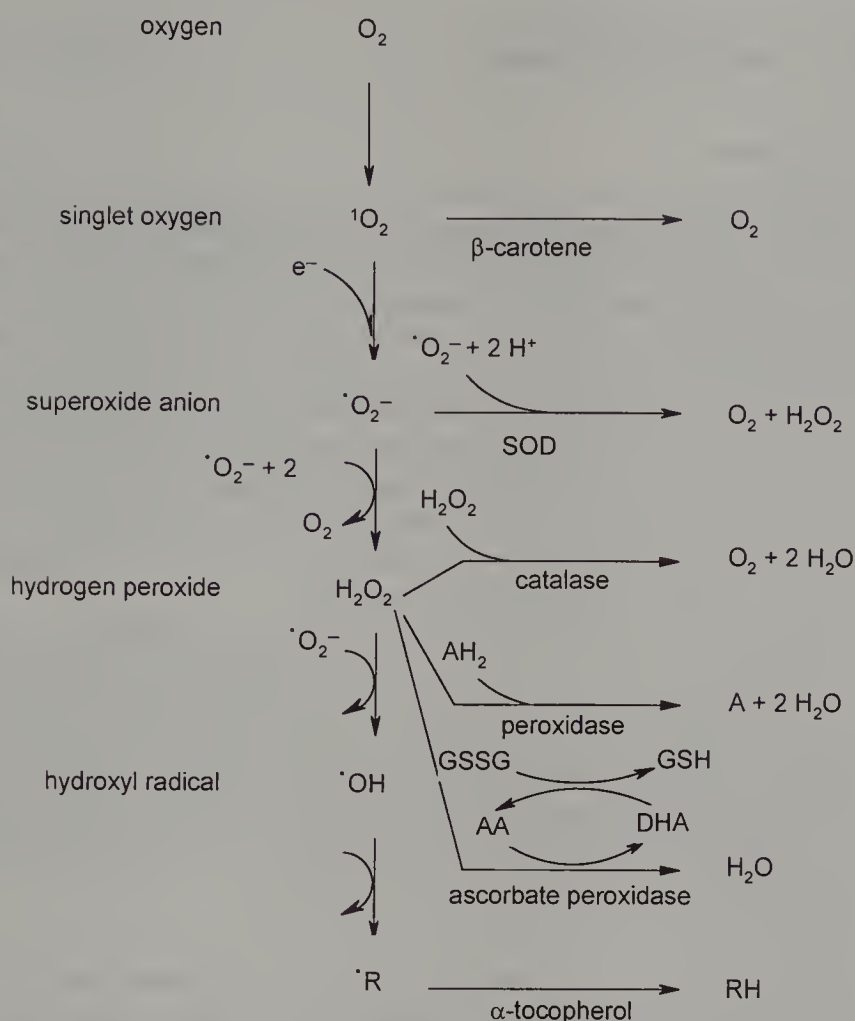


Figure 9.4 Formation of ROS and enzymes that remove ROS. The formation of reactive oxygen species and the cascade of subsequent reactions that result are shown. The roles of the antioxidants and antioxidant enzymes are shown. AA, ascorbic acid; DHA, dehydroascorbate; GSH, glutathione; GSSG, oxidised glutathione; SOD, superoxide dismutase.

oxygen and hydrogen peroxide. The activity of this enzyme therefore determines the concentration of the two substrates (superoxide and hydrogen peroxide) that react to form the hydroxyl radical. There is a broad family of SOD enzymes, requiring different metal ions for activity, and predominating in different cellular compartments. Thus, Mn^{2+} , Fe^{2+} and Cu^{2+}/Zn^{2+} forms are found, specific to different organelles (Box 9.5).

Transgenic tobacco plants containing each class of enzyme have been produced and tested for their ability to withstand oxidative stress, as shown in Table 9.5. In each case, there has been some measurable increase in the ability to withstand oxidative stresses such as exposure to ozone. The Mn -SOD has also been transformed into alfalfa and targeted to the chloroplast or mitochondria. In this case, the results of a 3-year field trial indicated significant improvements in the yield and survival of the transgenic plants. However, it is not clear whether the protective effect is due to the removal of superoxide by

BOX 9.4 Antioxidants in plants**L-Ascorbic acid**

Ascorbate (vitamin C) has an important role in plant metabolism, growth and development. Amongst other functions, it acts a reductant for many free radicals to minimise the damage caused by oxidative stress. Ascorbate can directly scavenge reactive oxygen species with and without enzyme catalysts, and can indirectly scavenge them by recycling α -tocopherol (see below) to the reduced form. Indeed, it reacts with ROS more readily than any other aqueous component. For example, ascorbate reacts with superoxide to form hydrogen peroxide and dehydroascorbate. It also reacts with hydrogen peroxide to form monodehydroascorbate and water, in a reaction catalysed by ascorbate peroxidase. It also regenerates membrane-bound antioxidants such as α -tocopherol that scavenge peroxy radicals and singlet oxygen. The tocopheroxyl radical is converted to α -tocopherol via the oxidation of ascorbate to monodehydroascorbate.

Note that there are two different products of ascorbate oxidation—monodehydroascorbate and dehydroascorbate—which represent one- and two-electron transfers, respectively. Two molecules of monodehydroascorbate may spontaneously dismutate to form ascorbate and dehydroascorbate, or can be reduced to ascorbate by NAD(P)H monodehydroascorbate reductase. Dehydroascorbate is unstable at pH > 6, forming tartrate and oxalate. To prevent this, dehydroascorbate is rapidly reduced to ascorbate by dehydroascorbate reductase using reducing equivalents from glutathione (GSH) (see below).

Glutathione

Glutathione (GSH) is a tripeptide (Glu–Cys–Gly) which was met in Chapter 5 with reference to herbicide detoxification. The antioxidant function of GSH involves the sulphhydryl group of the cysteine residue, which forms a thiyl radical on oxidation, and then reacts with a second oxidised glutathione to form a disulphide bond (GSSG). The redox potential of GSH enables it to reduce dehydroascorbate to ascorbate or to reduce the disulphide bonds of proteins.

GSH functions directly as a free-radical scavenger by reacting chemically with singlet oxygen, superoxide and hydroxyl radicals. It also stabilises membranes by removing acyl peroxides formed by lipid peroxidation reactions. It also functions to regenerate reduced ascorbic acid from dehydroascorbate.

The reduction of GSSG to GSH is catalysed by the enzyme glutathione reductase (GR), which exists in plants in multiple forms associated with different subcellular compartments. Glutathione peroxidase provides an alternative means of detoxifying activated oxygen, by using GSH to reduce hydrogen peroxide, producing GSSG.

 α -Tocopherol

α -Tocopherol (vitamin E), is a membrane-bound scavenger of oxygen free radicals, lipid peroxy radicals and singlet oxygen. It also serves as a membrane stabilising agent. The reaction with peroxy radicals formed in the lipid bilayer appears to be particularly important. The tocopheroxyl radical produced is stable and therefore does not propagate further free radicals. α -Tocopherol can be regenerated by ascorbate, as described above.

BOX 9.4 Continued**Carotenoids**

Carotenoids are C_{40} isoprenoids and tetraterpenes, and include the carotenes and xanthophylls (see Chapter 10). They are located in the plastids of both photosynthetic and non-photosynthetic plant tissues. In chloroplasts, they protect the photosystems by detoxifying ROS and triplet chlorophyll produced as a result of excitation of the photosynthetic complexes by light. Thus, they can perform a similar role to α -tocopherol in scavenging the products of lipid peroxidation. They can react with singlet oxygen and dissipate the energy as heat. They also dissipate excess excitation energy through the xanthophyll cycle, and they can quench triplet or excited chlorophyll molecules to prevent the formation of singlet oxygen.

BOX 9.5 The SOD enzyme family

Superoxide dismutase (SOD) catalyses the dismutation of superoxide to hydrogen peroxide and oxygen. SOD is present in most (if not all) subcellular compartments of the plant cell that generate activated oxygen, and is assumed to play a central role in the defence against oxidative stress. There are three distinct types of SOD classified on the basis of the metal cofactor: the copper/zinc (Cu/Zn-SOD), the manganese (Mn-SOD) and the iron (Fe-SOD) isozymes. The Mn-SOD is found in the mitochondria; whilst Cu/Zn-SOD isozymes are found in the cytosol, or in the chloroplasts. The Fe-SOD isozymes, though not always detected in plants, are usually found in the chloroplast. All of the plant SOD genes are in the nuclear genome and are targeted to their respective subcellular compartments by an amino-terminal targeting sequence. However, they are not regulated co-ordinately, but independently according to the degree of oxidative stress experienced in the respective subcellular compartments.

the enzyme, or whether the hydrogen peroxide product enhances stress tolerance by inducing other stress-related genes. More recently, the introduction of Cu/Zn-SOD, Mn-SOD and Fe-SOD genes together into alfalfa has been shown to increase the winter hardiness of the crop.

Strategy 2: Production of antioxidants

Table 9.5 also shows the alternative strategy for oxidative stress tolerance in action. Thus, three of the enzymes shown in Figure 9.4—ascorbate peroxidase, glutathione peroxidase and glutathione reductase—have been transformed into *Arabidopsis* and tobacco plants and shown to have some effect on tolerance to various abiotic stresses such as heat, cold and salinity. Glutathione reductase also provided resistance to the oxidative stress resulting

Table 9.5 Transgenes used to engineer tolerance to oxidative stress

Gene	Host	Stress tolerance
Mitochondrial Mn-SOD Tobacco	Alfalfa chloroplast	2× increase in SOD Increased field drought tolerance Increased freezing tolerance
Mitochondrial Mn-SOD Tobacco	Tobacco chloroplast	2–4× increase in SOD Increased ozone tolerance
Mitochondrial Mn-SOD Tobacco	Tobacco mitochondria	8× increase in SOD No effect on ozone tolerance
Mn SOD	Canola	Increased aluminium tolerance
Chloroplast Cu/Zn-SOD Pea	Tobacco chloroplast	3–15× increase in SOD Increased tolerance to high light and chilling
Cytosolic Cu/Zn-SOD	Tobacco cytosol	1.5–6× increase in SOD Reduced damage from acute ozone exposure
Fe-SOD <i>Arabidopsis</i>	Tobacco	Protected plants from ozone damage
<i>Apx3</i> (ascorbate peroxidase)	Tobacco	Increased protection against oxidative stress
<i>Apx1</i> (ascorbate peroxidase)	<i>Arabidopsis</i>	Heat tolerance
GST/GPX (glutathione <i>S</i> -transferase with glutathione peroxidase)	Tobacco	Increased stress tolerance
<i>Nt107</i> (glutathione <i>S</i> -transferase)	Tobacco	Sustained growth under cold and salinity stress
<i>ParB</i> (glutathione <i>S</i> -transferase)	<i>Arabidopsis</i>	Protects against aluminium toxicity and oxidative stress
<i>NtPox</i> (glutathione peroxidase)	<i>Arabidopsis</i>	Protects against aluminium toxicity and oxidative stress
Glutathione reductase <i>E. coli</i>	Tobacco chloroplast	3–6× increase in foliar GR Increased tolerance to SO ₂ and paraquat
Glutathione reductase <i>E. coli</i>	Tobacco cytosol	1–35× increase in GR Increased tolerance to paraquat
Glutathione synthetase <i>E. coli</i>	Poplar cytosol	100× increase in GS GSH not increased No effect on paraquat tolerance
<i>MsFer</i> Alfalfa ferritin	Tobacco	Increased tolerance to oxidative damage caused by excess iron

SOD, superoxide dismutase; GST, glutathione *S*-transferase; GPX, glutathione peroxidase; GR, glutathione reductase; GS, glutathione synthetase; GSH, glutathione.

from paraquat treatment. On the other hand, simply increasing the cellular concentration of glutathione by expressing glutathione synthase had no effect on stress tolerance.

Summary

This chapter has highlighted the importance of abiotic stress in determining the large annual fluctuations in crop yield. Although crops experience a number of different abiotic stresses, several of these cause two major problems in common—water-deficit and oxidative stress. Two general strategies for engineering tolerance to abiotic stresses in plants are therefore possible. Some measure of tolerance to water-deficit stress can be provided by the synthesis of compatible solutes. On the other hand, expressing enzymes involved in protection against reactive oxygen species can combat oxidative stress.

In some cases, single gene mechanisms for tolerating specific stresses can be deployed (e.g. salt stress and cold stress). However, the overriding theme from this chapter is that abiotic stresses induce complex reactions from plants, and that optimal protection may well involve several genes.

Further reading

Plant stress—general

Bray, E. A., Bailey-Serres, J. and Weretilnyk, E. (2000). Responses to abiotic stresses. In *Biochemistry and molecular biology of plants* (ed. B. B. Buchanan, W. Gruissem and R. L. Jones), Chapter 22, pp. 1158–1203. American Society of Plant Physiologists, Rockville, Maryland, USA

Plant stress—web-link 9.1: <http://www.plantstress.com>

Stress tolerance—general

Holmberg, N. and Bulow, L. (1998). Improving stress tolerance in plants by gene transfer. *Trends in Plant Science*, 3, 61–6.

Zhang, J., Kueva N. Y., Wang, Z., Wu, R., Ho, T.-H. and Nguyen, H. T. (2000). Genetic engineering for abiotic stress resistance in crop plants. *In vitro Cellular and Developmental Biology—Plant*, 36, 108–14.

Osmoprotectants and stress tolerance

Chen, T. H. H. and Murata, N. (2002). Enhancement of tolerance of abiotic stress by metabolic engineering of betaines and other compatible solutes. *Current Opinion in Plant Biology*, 5, 250–7.

Nuccio, M. L., Rhodes, D., McNeil, S. D. and Hanson, A. D. (1999). Metabolic engineering of plants for osmotic stress resistance. *Current Opinion in Plant Biology*, 2, 128–34.

Sakamoto, A. and Murata, N. (2000). Genetic engineering of glycine betaine synthesis in plants: current status and implications for enhancement of stress tolerance. *Journal of Experimental Botany*, 51, 81–8.

Sakamoto, A. and Murata, N. (2002). The role of glycine betaine in the protection of plants from stress: clues from transgenic plants. *Plant, Cell and Environment*, 25, 163–71.

Serraj, R. and Sinclair, T. R. (2002). Osmolyte accumulation: can it really help increase crop yield under drought conditions? *Plant, Cell and Environment*, 25, 333–41.

Salt tolerance

Apse, M. P., Aharon, G. S., Snedden, W. A. and Blumwald, E. (1999). Salt tolerance conferred by overexpression of a vacuolar Na⁺/H⁺ antiport in Arabidopsis. *Science*, 285, 1256–8.

Frommer, W. B., Ludewig, U. and Rentsch, D. (1999). Enhanced: taking transgenic plants with a pinch of salt. *Science*, 285, 1222.

Cold tolerance

Jaglo-Ottosen, K. R., Gilmour, S. J., Zarka, D. G., Schabenberger, O. and Thomashow, M. F. (1998). Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science*, 280, 104–6.

Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1999). Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nature Biotechnology*, 17, 287–91.

Thomashow, M. F. (1998). Role of cold-responsive genes in plant freezing tolerance. *Plant Physiology*, 118, 1–7.

Heat stress

Lee, J. H., Hubel, A. and Schoffl, F. (1995). Derepression of the activity of genetically engineered heat shock factor causes constitutive synthesis of heat-shock protein and increased thermotolerance in transgenic Arabidopsis. *Plant Journal*, 8, 603–12.

Oxidative stress

McKersie, B. Tutorial on oxidative stress—web-link 9.2:
<http://www.agronomy.psu.edu/Courses/AGRO518/oxygen.htm>

The improvement of crop yield and quality

Introduction

The previous five chapters have considered a variety of ways in which the productivity of crop plants can be improved by enhancing their ability to resist or tolerate biotic and abiotic stresses. These strategies can all contribute to an improvement in crop yield by allowing the plants to better withstand external factors that reduce the amount and quality of harvestable plant material. However, the performance of crop plants is also determined by endogenous factors that affect the yield and quality of the harvestable material produced. This chapter will look at a number of examples of how crop productivity can be genetically enhanced by increasing the amount, or improving the quality, of material produced by the crop.

The yield of a crop is ultimately determined by the amount of solar radiation intercepted by the crop canopy, the photosynthetic efficiency (i.e. the efficiency of conversion of radiant to chemical potential energy) and the harvest index (the fraction of dry matter allocated to the harvested part of the crop). Thus, manipulation of crop yield requires some understanding of the fundamental processes that determine photosynthetic efficiency and dry-matter partitioning. On the other hand, the quality of a crop is determined by a wide range of desirable characteristics such as nutritional quality, flavour, processing quality and shelf-life. The nature of yield and quality are therefore much more varied and wide-ranging than the other traits described to this point, and the structure of this chapter is therefore somewhat different to reflect this. Rather than try to cover the entire range of targets for different crops, a small number of yield and quality traits have been selected to exemplify the types of approaches that have been adopted in this area. The chapter therefore starts with an extended discussion of one system—tomato ripening—because it is possible to draw several general lessons from the techniques and concepts that were deployed in this project.

The genetic manipulation of fruit ripening

The predominance of herbicide and pest resistance amongst the commercially developed GM traits has been discussed previously (Chapters 5 and 6). However, amongst the first GM products to reach the market, around 1994, were Calgene's FlavrSavr fresh tomatoes, and processed tomato products containing delayed-softening tomato fruit developed in the UK by Zeneca and the University of Nottingham. This may seem rather puzzling, because the tomato crop is relatively small compared to the major GM crops of Northern America (maize, soybean, oilseed rape and cotton). Furthermore, it is not immediately clear why a large biotechnology company like Zeneca would be involved, when they had no direct interest in tomato seeds or processed products when the project was started (when the company was still ICI). Part of the answer lies in the early recognition by ICI and Calgene of the value of using tomato ripening as a model system on which to develop expertise and test the

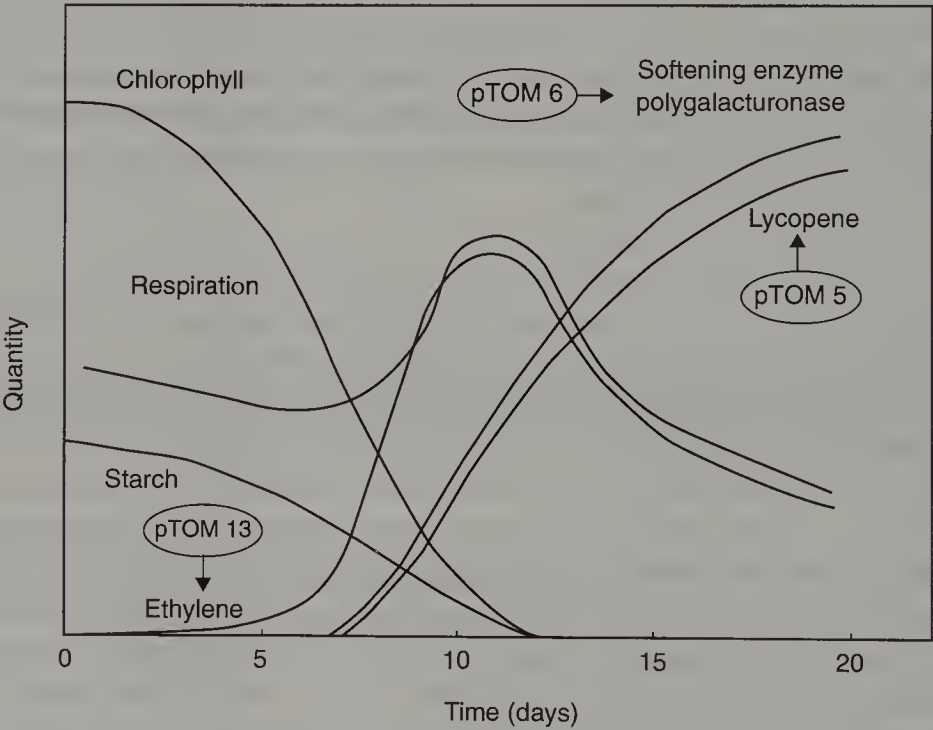


Figure 10.1 Biochemical changes during tomato ripening. The graph shows the changes that occur as tomato fruits progress from the mature green stage to fully ripe fruit. The onset of ripening is characterised by a burst of respiration, the 'respiratory climacteric', which is accompanied by a rapid increase in ethylene synthesis. The green coloration of the fruit is lost as chlorophyll is degraded, and the sweetness of the fruit is enhanced as starch is converted to sugars. The red pigment, lycopene, is synthesised, and the fruit softens as a result of the appearance of cell wall degrading enzymes such as polygalacturonase (PG). The graph also indicates the role of three ripening-related clones in the ripening process. The clone pTOM6 encodes PG, pTOM5 codes for phytoene synthase, and pTOM13 encodes the ethylene forming enzyme, ACC oxidase. (Redrawn with permission from Grierson, D. and Covey, S. (1988).)

potential of plant biotechnology to modify crop quality. In order to understand why tomato ripening was seen as such a good model system, it is necessary to look at some of the basic biology of fruit ripening.

Fruit ripening is an active process that, in climacteric fruit such as tomatoes, is characterised by a burst of respiration (the respiratory climacteric), ethylene production, softening and changes to colour and flavour. Figure 10.1 shows the transient peaks in respiration and ethylene production at the start of tomato ripening, which are accompanied by softening, a change in colour from green to red and enhanced flavour. The peak of ethylene production is significant, because ethylene is known to be the phytohormone that triggers ripening in climacteric fruit. The colour change results from the degradation of chlorophyll and the production of the red pigment, lycopene. Flavour changes occur as starch is broken down and sugars accumulate. A large number of secondary products that improve the smell and taste of the fruit are also produced. The softening of the fruit is largely the result of the cell wall degrading activity of the enzymes polygalacturonase (PG) and pectin methylesterase (PME). The PG enzyme is synthesised *de novo* during ripening and acts to break down the polygalacturonic acid chains that form the pectin ‘glue’ of the middle lamella, which ‘sticks’ neighbouring cells together. The realisation that the ripening process involved the activation of specific ripening-related genes, such as the one encoding polygalacturonase, enabled these genes to be cloned, as shown in Box 10.1.

BOX 10.1 cDNA cloning of ripening-related genes

The three ripening genes that will be discussed in this chapter were isolated from the same cDNA library made from ripe tomato fruit. Ripening-related clones were identified by differential screening of the cDNA library with cDNA sequences prepared from mRNA populations from mature green fruit and ripe fruit. Clones that hybridised to the ripe cDNA but not to the mature green cDNA were classed as ripening-related. Over 150 clones were isolated, grouped and characterised. Three groups of clones, represented by pTOM5, pTOM6 and pTOM13, were confirmed as ripening-related and were also shown to be induced by ethylene. Their identities and roles in the ripening process were subsequently determined, as shown in the table below.

Clone	Gene product	Function	Role in ripening
pTOM5	Phytoene synthase	Lycopene synthesis	Red coloration
pTOM6	Polygalacturonase	Cell wall degradation	Fruit softening
pTOM13	ACC oxidase	Ethylene formation	Ripening trigger

CASE STUDY 1 The genetic manipulation of fruit softening

Sequencing of the TOM 6 clone showed that the sequence matched the N-terminal sequence of the polygalacturonase (PG) protein. In other words, one of the genes implicated in a particular part of the ripening process—softening—had been cloned. Northern blot analysis confirmed that expression of the PG gene was ripening-specific, and was induced by ethylene. The availability of the PG gene led to two types of question: (1) was it possible to use the cloned gene to investigate the precise role of PG in ripening; and (2) was it possible to modify the expression of endogenous PG in order to manipulate the ripening process. This led to one of the first examples of using antisense techniques to alter plant gene expression (see Box 10.2).

The initial antisense PG construct comprised a partial cDNA fragment of pTOM6 inserted in reverse orientation downstream of the 35S promoter (Figure 10.2). Note that expression of the antisense gene will therefore be 'constitutive', whilst expression of the endogenous gene is, of course, ripening-specific.

The effect of transforming this construct into tomato plants was quite instructive. Table 10.1 summarises the results of Northern blots of RNA from mature green and ripe fruit from both transgenic and non-transgenic tomatoes, hybridised with probes to sense and antisense mRNA. There are a number of things to note from this blot. PG mRNA is not present in mature green wild-type fruit, but is induced during ripening. In ripe transgenic fruit, the level of PG mRNA is much reduced. The antisense RNA probe reveals that the antisense RNA accumulates to high levels in mature green fruit, but its level is also reduced in ripe fruit. The conclusion is that the interaction of sense and antisense mRNA leads to a reduction in both transcripts (Box 10.2).

In these transgenic fruit, it could be shown that polygalacturonase activity was much reduced during ripening. Furthermore, this caused a specific delay in softening, but not

BOX 10.2 Antisense RNA

The basic antisense concept is relatively simple to describe—explaining why it works has proved to be more difficult until relatively recently. Essentially, it involves creating a construct in which the gene sequence is transcribed in the reverse orientation, using the opposite strand as the template. Therefore the resulting antisense transcript has a sequence complementary to the normal (sense) mRNA. The original concept assumed that some interaction at the transcriptional, post-transcriptional or translational level would reduce the expression of the endogenous mRNA. This turned out to be the case, and subsequent evidence from different antisense constructs indicated that the levels of both sense and antisense RNA were reduced, and that the antisense construct did not need to be full length (see Table 10.1). Current thinking about the mode of action is that double-stranded RNA hybrids formed between the antisense RNA and endogenous mRNA are recognised by plant cell-defence mechanisms and degraded. Detailed knowledge of these mechanisms has led to the recent elaboration of the RNAi strategy (see Chapter 3, Box 3.4).

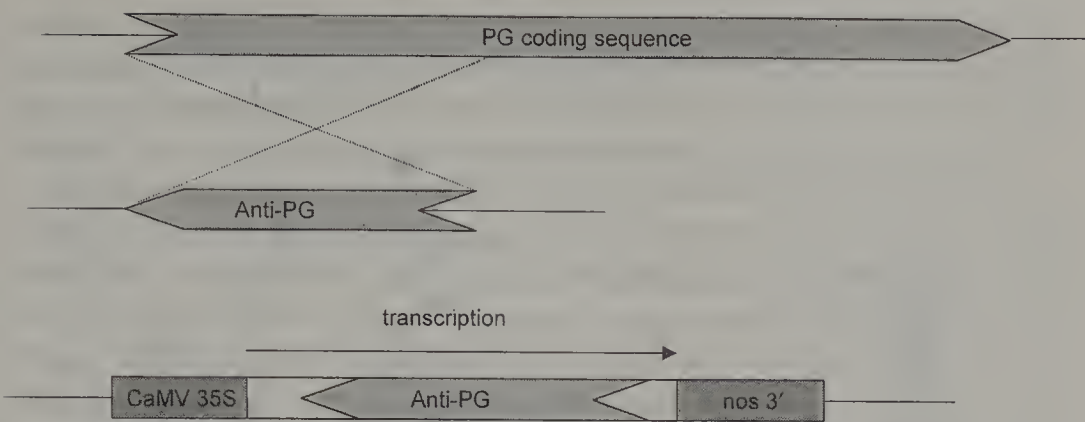


Figure 10.2 Construction of an antisense PG vector. The diagram shows the antisense concept in operation. Basically, all or part of a gene sequence is inserted into a vector in the ‘wrong’ orientation, so that the opposite strand from normal is transcribed to produce an ‘antisense’ transcript. This will be complementary to the endogenous mRNA, allowing the formation of double-stranded RNA.

Table 10.1 The expression of pTOM6 in transgenic tomato plants. The results of Northern blot hybridisation of samples from transgenic tomato fruit with sense and antisense (AS) pTOM6 probes are shown.

Fruit	Non-transgenic Mature green		Non-transgenic Ripe		Transgenic Mature green		Transgenic Ripe	
Probe	Sense	AS	Sense	AS	Sense	AS	Sense	AS
Signal	ND	ND	Intense band	ND	ND	Intense bands	Faint band	Faint bands
Size of band			PG mRNA			<PG mRNA	PG mRNA	<PG mRNA

AS = antisense; ND= Not detected.

in other ripening events. Thus, Figure 10.3 shows that the red lycopene pigment still accumulated during the ripening of these transgenic fruit, whilst production of the softening enzyme was inhibited.

It was possible to modulate the antisense effect by analysing heterozygote (1 antisense gene) and homozygote (2 antisense genes) plants. Two copies of the antisense gene were found to reduce the levels of PG to virtually undetectable levels during ripening. Thus, tomato fruit in which the activity of one specific enzyme had been effectively knocked out had been generated. It subsequently proved possible to reduce PG activity by co-suppression (see also Chapters 3 and 8) using sense PG constructs (Box 10.3).

The effect of reducing polygalacturonase activity in low-PG transgenic fruit was to inhibit pectin degradation, as indicated by the large size of polyuronides remaining in the cell wall. The resulting changes to food processing properties were determined by Zeneca and are shown in Figure 10.4. The Bostwick viscosity is influenced largely by the insoluble cell wall polymers and is a measure of the potential paste yield of tomato fruit. The Bostwick

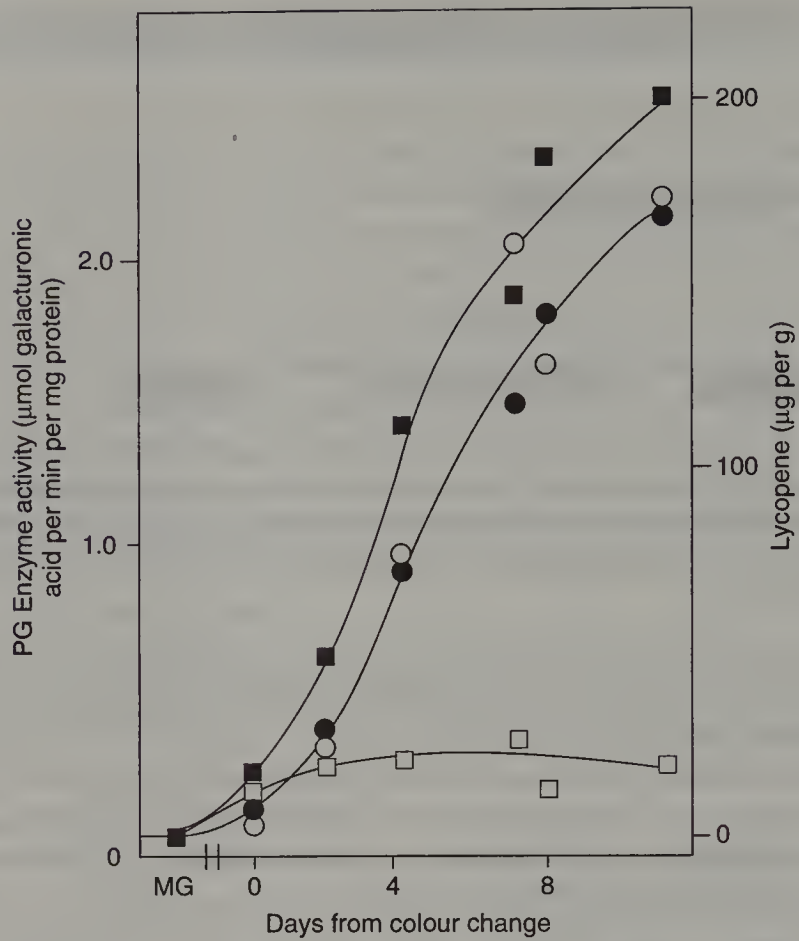


Figure 10.3 Polygalacturonase (PG) activity and lycopene content during the ripening of antisense PG fruit. The effect of expression of the antisense PG construct on PG enzyme activity (■) and lycopene content (□) is compared with PG activity (●) and lycopene (○) in untransformed control fruit. The antisense PG markedly reduces PG activity, but does not affect the accumulation of lycopene. (Redrawn with permission from Smith, C. J. S., *et al.* (1988).)

BOX 10.3 Co-suppression of PG gene expression by sense PG constructs

In addition to pioneering the use of antisense techniques, the tomato ripening system was also one of the first to reveal the phenomenon of co-suppression. This was observed in fruit in which sense rather than antisense PG constructs were expressed. In several cases, the overall level of polygalacturonase (PG) activity was less than in non-transgenic controls. This implied some mechanism whereby the plant cell was able to detect the presence of extra gene copies and/or increased levels of specific gene expression of the endogenous gene, and to suppress both endogenous and transgene expression. The term ‘co-suppression’ is used specifically to describe this mutual suppression of endogenous gene and transgene expression. It is one example of a more general phenomenon called ‘transgene silencing’, in which the expression of a transgene, even though intact and stably integrated, is suppressed (see Chapter 3).

viscosity in low-PG fruit (PG activity reduced by 99%) was enhanced by >80% as a result of the increase in size of the cell wall pectins, indicating a significant increase in the yield of paste from these tomatoes. This characteristic enabled Zeneca to exploit these transgenic tomatoes for the production of paste for processed foods such as sauces and pizzas.

Calgene in the USA exploited the delayed softening characteristics of low-PG fruit in quite a different way. Their low-PG varieties were found to remain firmer on the vine and during storage than unmodified fruit. Fresh tomatoes are generally harvested before they are fully ripe, to ensure they are firm enough to withstand handling during harvest, processing and storage. Ripening may then be induced by ethylene prior to sale. Whilst this ensures the fruit are sold intact, at a uniform size, colour and stage of ripeness, it is a common perception of consumers that tomatoes processed in this way have lost much of the flavour and aroma of those eaten straight from the vine. In contrast, the Calgene 'FlavrSavr' tomatoes were allowed to ripen and develop their full flavour on the vine, but the ripe fruit were still firm enough to withstand damage from handling and also from postharvest fungal infections.

Figure 10.4 also shows the effect of reducing pectin methylesterase activity in tomato fruit. This enzyme is present in unripe as well as ripe fruit, and removes methyl groups from cell wall pectin. Low-PME fruit were found to have an enhanced serum viscosity, reflecting an increase in the proportion of soluble pectin. Serum viscosity contributes to the glossy appearance of tomato paste. When the low-PG and low-PME traits were combined (either by crossing parent tomatoes with a single trait, or by re-transformation with both genes in tandem), the fruit had the quality characteristics of both single gene lines, plus a significant increase in soluble solids as indicated by a high Brix value. The soluble solids are measured on the Brix scale—a value obtained using refractometry. The soluble solids are an important determinant of tomato processing characteristics, with a high Brix value indicating good processing quality.

Despite their initial success, neither the Zeneca nor the Calgene low-PG tomato products can now be found, though for quite different reasons. The Zeneca tomato pastes were clearly labelled as GM products yet sold well when they were first introduced. However, they were cleared from the shelves of UK supermarkets following the anti-GM backlash catalysed by the Pusztai affair (see Chapter 6), despite the fact there has been no evidence of any risk to human health from this product. The FlavrSavr tomato met a different fate, in that the trait was not originally introduced into a suitable background and the tomato did not perform well in commercial production.

The genetic modification of fruit softening demonstrated the general validity of the concept that a specific trait could be manipulated by downregulating the expression of an endogenous gene using antisense and co-suppression techniques. The specific example of reducing cell wall degradation by targeting endogenous PG and PME has been applied to a number of other fruit, such as mango, peach and pear. Furthermore, it is now apparent that pectin depolymerisation is not the only determinant of texture, and the roles of other enzymes such as pectate lyases, cellulases and xyloglucan hydrolases are being

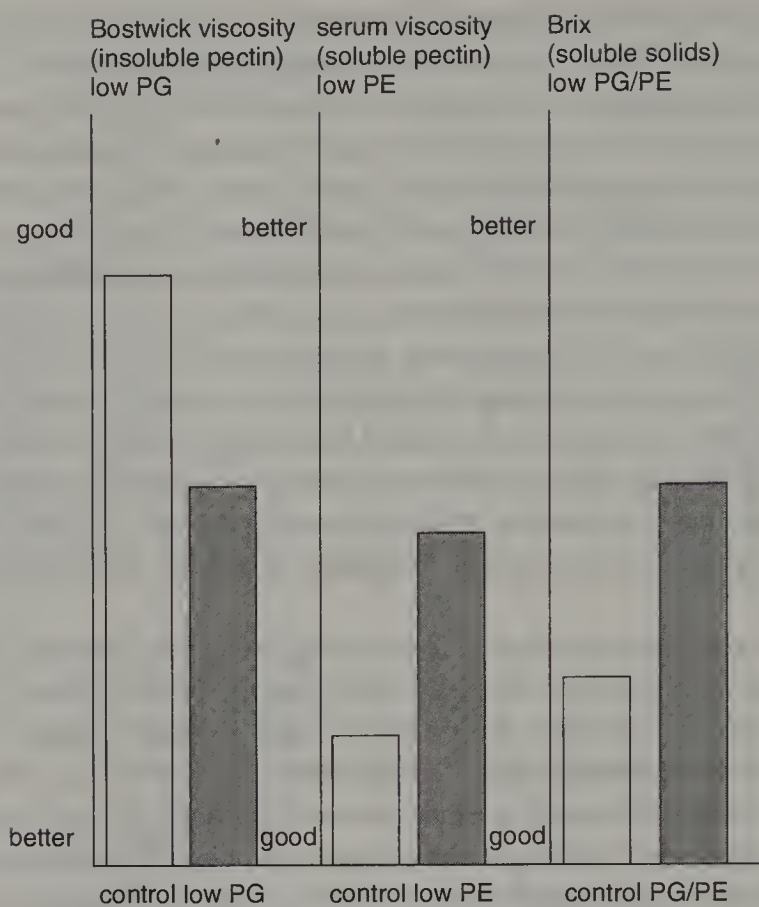


Figure 10.4 Relative processing performance of low-PG and low-PE fruit. Three different assays of processing quality were used to assess the effects of reducing two cell wall degrading enzymes—PG and pectin methylesterase (PE). The low-PG fruit show an improved Bostwick viscosity (determined by insoluble pectin), whilst low-PE fruit have a better serum viscosity (related to soluble pectin). The total soluble solids (Brix) is improved in fruit with combined low-PG and low-PE traits. (Redrawn with permission from Figure 6, p. 248 in Grierson, D. and Schuch, W. (1993).)

investigated. Plant cell wall structure and metabolism are still poorly understood, but research in these areas holds out the prospect of being able to manipulate many aspects of plant growth and development.

CASE STUDY 2 **The genetic modification of ethylene biosynthesis**

The use of pTOM6 indicated that, although the ripening process is co-ordinated, specific elements of the process could be manipulated without affecting others. Thus, antisense PG fruit still change colour and accumulate sugars and flavour compounds. However, the cloning of several ripening-related sequences led to the proposal that some of these might be involved in the transient burst of ethylene production, and that the study of these clones could provide some insight into the regulation of the entire ripening process. Investigation of the expression of ripening-related clones in wounded leaves (which also produce ethylene) showed that the expression of pTOM13 was related to ethylene production rather than to ripening *per se*. Sequence analysis led to the preliminary identification of pTOM13 as the

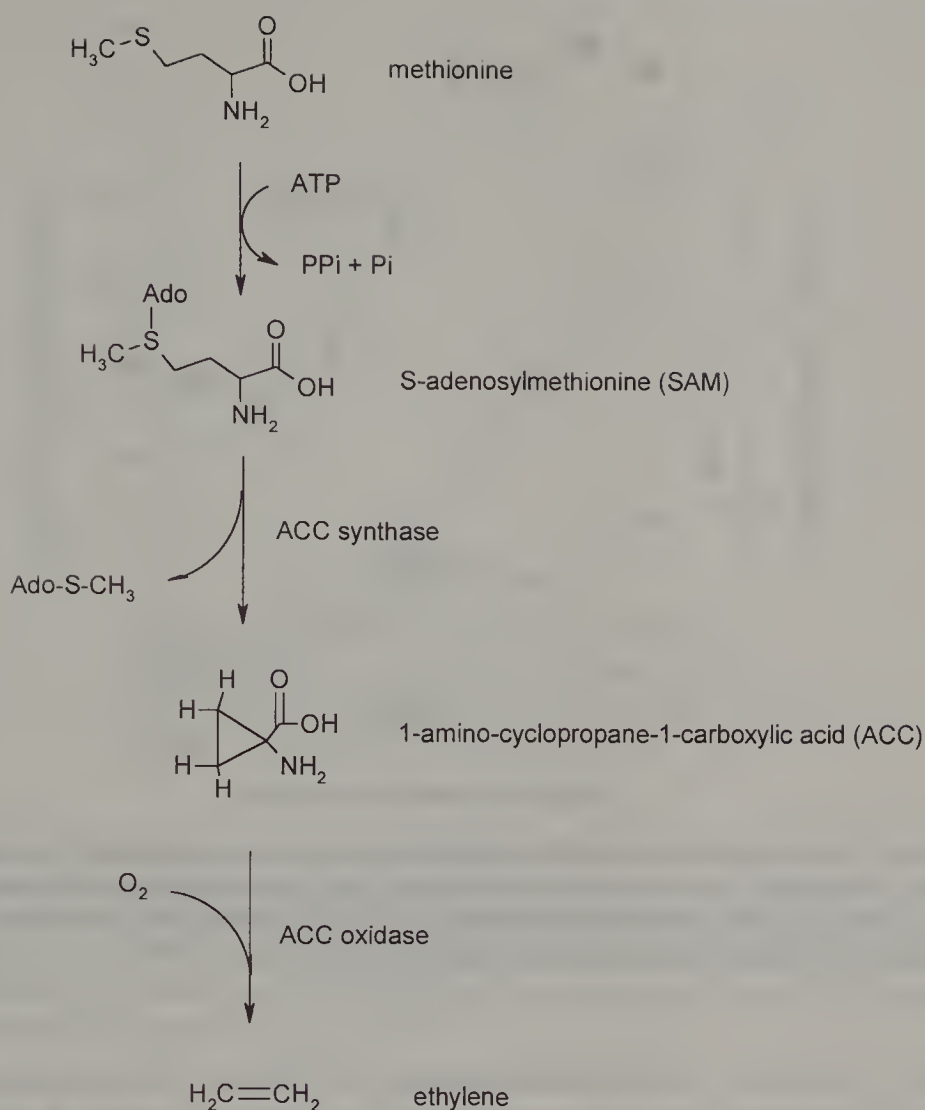


Figure 10.5 Ethylene biosynthesis and its regulation. The pathway to ethylene from methionine is shown. Only two committed steps from *S*-adenosylmethionine are involved. The first is catalysed by ACC synthase, generating the unusual three-sided ring compound, 1-amino-cyclopropane-1-carboxylic acid (ACC). This is converted to ethylene by ACC oxidase, producing CO_2 , HCN and H_2O in the process. Ascorbate is a co-substrate and is oxidised to dehydroascorbate (see Chapter 9).

second ethylene-forming enzyme (EFE) in the two-step pathway from *S*-adenosylmethionine to ethylene (Figure 10.5). The protein sequence, plus the availability of the clone for protein expression studies, also allowed the classification of the previously uncharacterised ethylene forming enzyme as ACC oxidase.

The antisense strategy was initially used to confirm the role of pTOM13 in ethylene synthesis. Transgenic tomato plants carrying the antisense pTOM13 construct produced much less ethylene, either in wounded leaves, or during ripening (Figure 10.6) than did control plants.

The effect of the much-reduced levels of ethylene produced during ripening was to delay the entire ripening process. Thus, lycopene production (Figure 10.7) and softening (Figure 10.8) occur much more slowly than in normal plants. These ripening phenomena can be

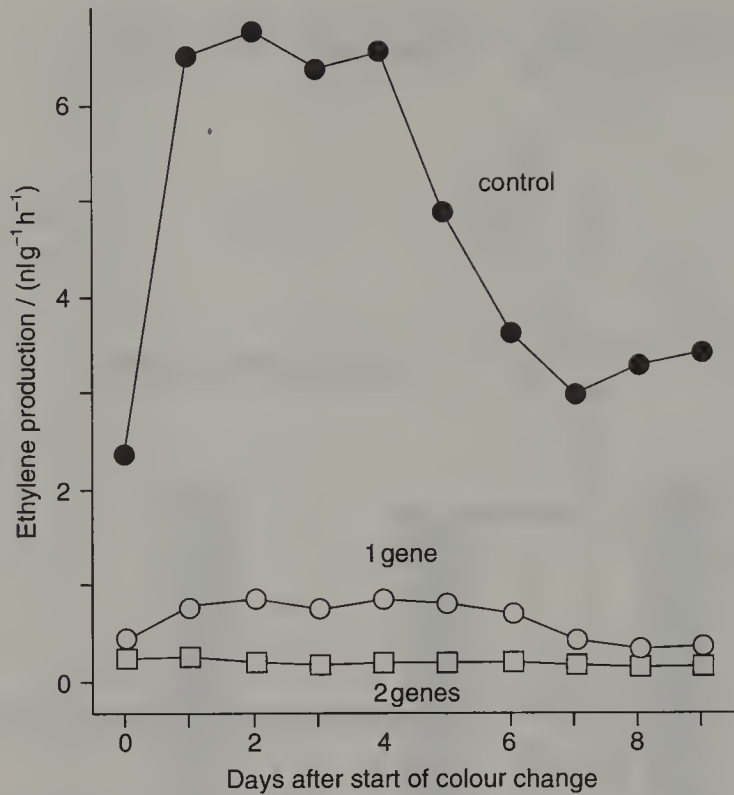


Figure 10.6 Ethylene production in tomato fruit transformed with an antisense ACC oxidase construct. The climacteric burst of ethylene production seen in control fruit is greatly decreased in fruit carrying one antisense gene, and effectively abolished in homozygotes carrying two antisense genes. (Adapted with permission from Hamilton, A. J., *et al.* (1990).)

induced in the antisense plants by applying exogenous ethylene. The increased storage capacity of these fruit has tremendous potential for reducing spoilage of the fruit during storage and transport.

“Antisense ethylene” technology is applicable to all climacteric fruits, and to other systems triggered by ethylene, such as the spoilage of vegetables and the senescence of picked flowers. For example, it has been known for some time that the vase-life of cut flowers can be extended by adding silver salts to the water, which blocks the response to ethylene. The introduction of anti-sense ACC synthase or ACC oxidase would have a similar effect of extending vase-life by suppressing ethylene synthesis. More importantly, the postharvest spoilage of fruit and vegetables is often very rapid in tropical countries, and this can be a major barrier to the efficient distribution of food, particularly where the transport and food storage infrastructure is not well developed. The ability to delay ripening, senescence and spoilage could make a significant contribution to the problems of food distribution in developing countries.

In addition, the successful manipulation of one plant growth regulator indicates the potential for engineering plant development by manipulating other plant growth regulators. The early availability of auxin and cytokinin biosyn-

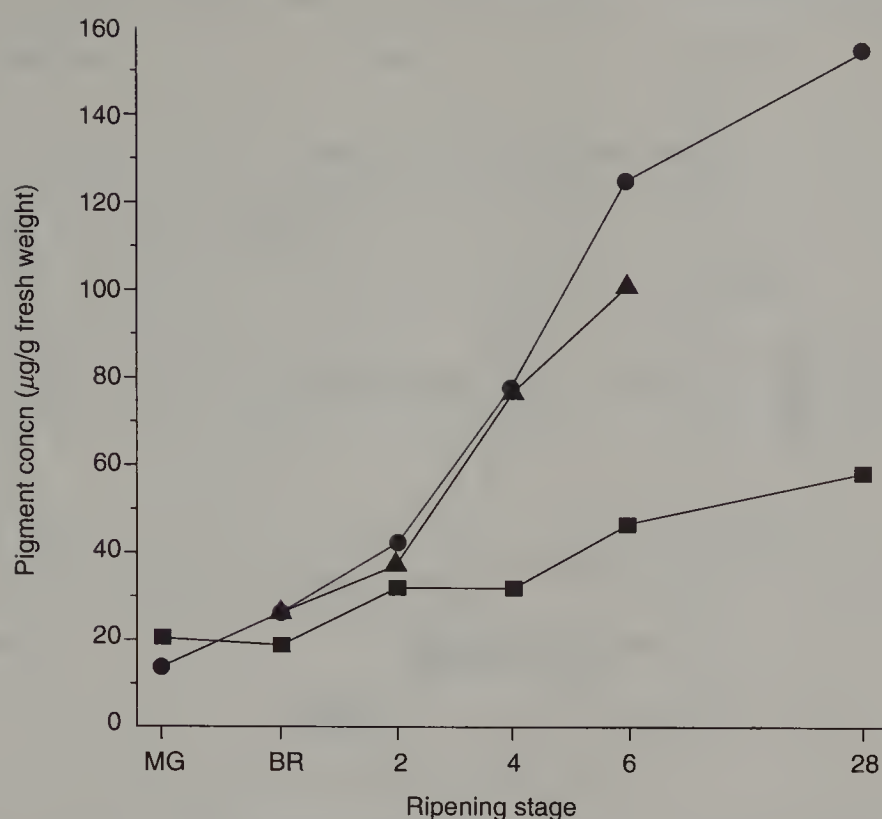


Figure 10.7 Pigments in antisense ACC oxidase fruit during ripening. Tomato fruits were removed from the plant at the mature green stage and the total carotenoid pigments were measured at the time intervals indicated. The accumulation of carotenoids in the antisense ACC oxidase fruit (square) was much reduced compared to the control (circle), but could be restored by incubating the fruit in a container containing exogenous ethylene (triangle). (Redrawn with permission from Bouzayen, M., *et al.* (1992).)

thetic genes from the *Agrobacterium tumefaciens* Ti plasmid (see Chapter 3) has meant that both phytohormones have been synthesised in a range of plants under the control of various promoters. The more recent cloning of genes involved in gibberellin and abscisic acid synthesis/response has also permitted the genetic manipulation of important developmental processes (see below).

CASE STUDY 3 Modification of colour

The third, ripening-related clone, pTOM5, highlighted at the beginning of this chapter (Box 10.1) encodes phytoene synthase, a key enzyme in the biosynthesis of the red pigment, lycopene (Figure 10.9). The position of this step in the isoprenoid pathways of plants is shown. Many classes of compound are synthesised by building up molecules from the 5-carbon isoprenoid unit. In particular, this route forms the basis of the biosynthetic pathways of three classical plant hormones (Chapter 2). Cytokinins such as zeatin are produced by the addition of a single isoprenoid unit to adenosine. The biosynthesis of gibberellins (GA) starts from the 20-carbon unit geranyl geranyl diphosphate (GGDP). Abscisic acid (ABA) is formed from β -carotene, as are a wide range of carotenoids and xanthophylls.

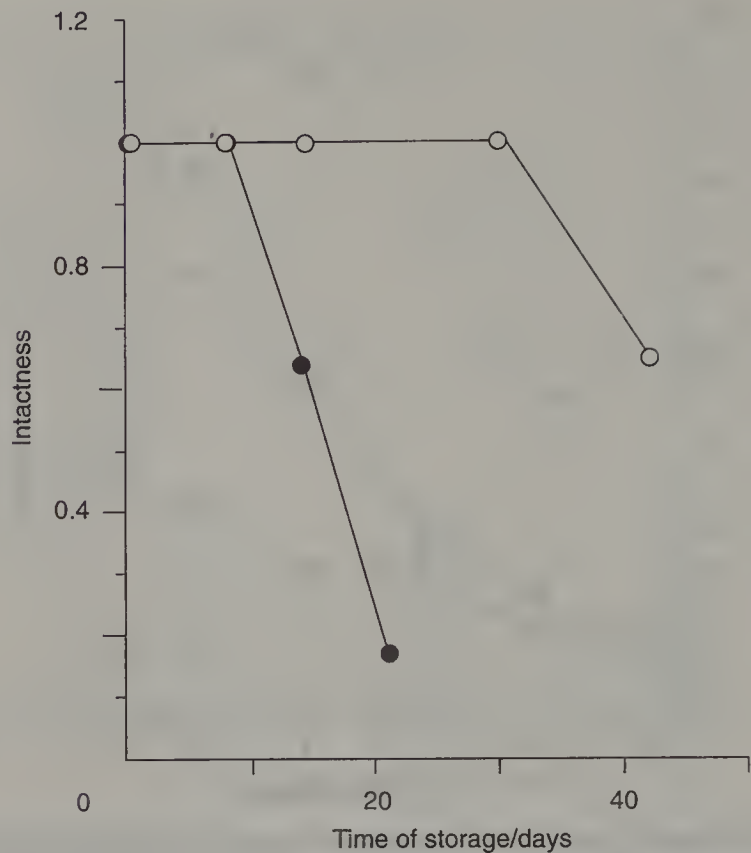


Figure 10.8 Increased storage of low ethylene fruit. The reduction of ethylene formation in antisense ACC oxidase fruit (○) slows down the ripening process*. One qualitative measure of this effect is on the intactness of the fruit—in other words, it takes a lot longer for the antisense fruit to become a soft pulp. (Redrawn with permission from Grierson, D. and Schuch, W. (1993).)

As might be expected, antisense pTOM5 tomato plants produce yellow fruit during ripening. On the other hand, plants overexpressing pTOM5 under the control of the 35S promoter were found not only to produce lycopene ectopically, but also to have a shorter height. This latter finding was unexpected, but can be explained with reference to Figure 10.9. Analysis of gibberellin synthesis in these plants showed the ectopic phytoene synthase activity was channelling 20-C units into the lycopene pathway, leading to a reduction in GA synthesis. Gibberellins play a role in stem elongation, so the reduction in GA synthesis caused the dwarf phenotype.

The manipulation of lycopene synthesis, and its effect on gibberellin synthesis, leads on to three different examples of the manipulation of yield and quality. The first general area to consider is the modification of colour. There are many examples of this, particularly amongst the ornamental flowers (see Box 10.4).

Another area that relates to the genetic manipulation of tomato colour is the effect that overexpression of phytoene synthase had on the gibberellin

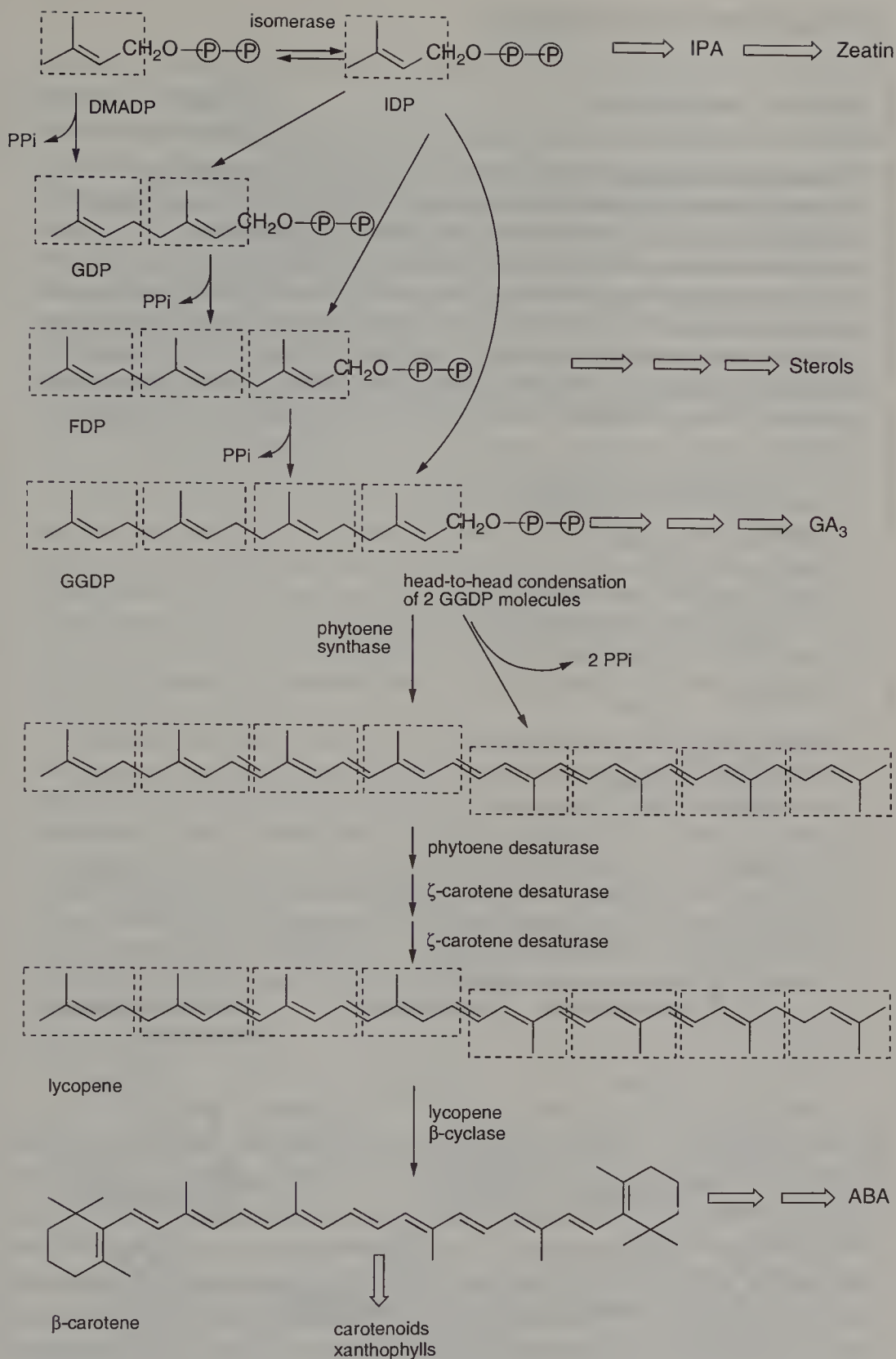
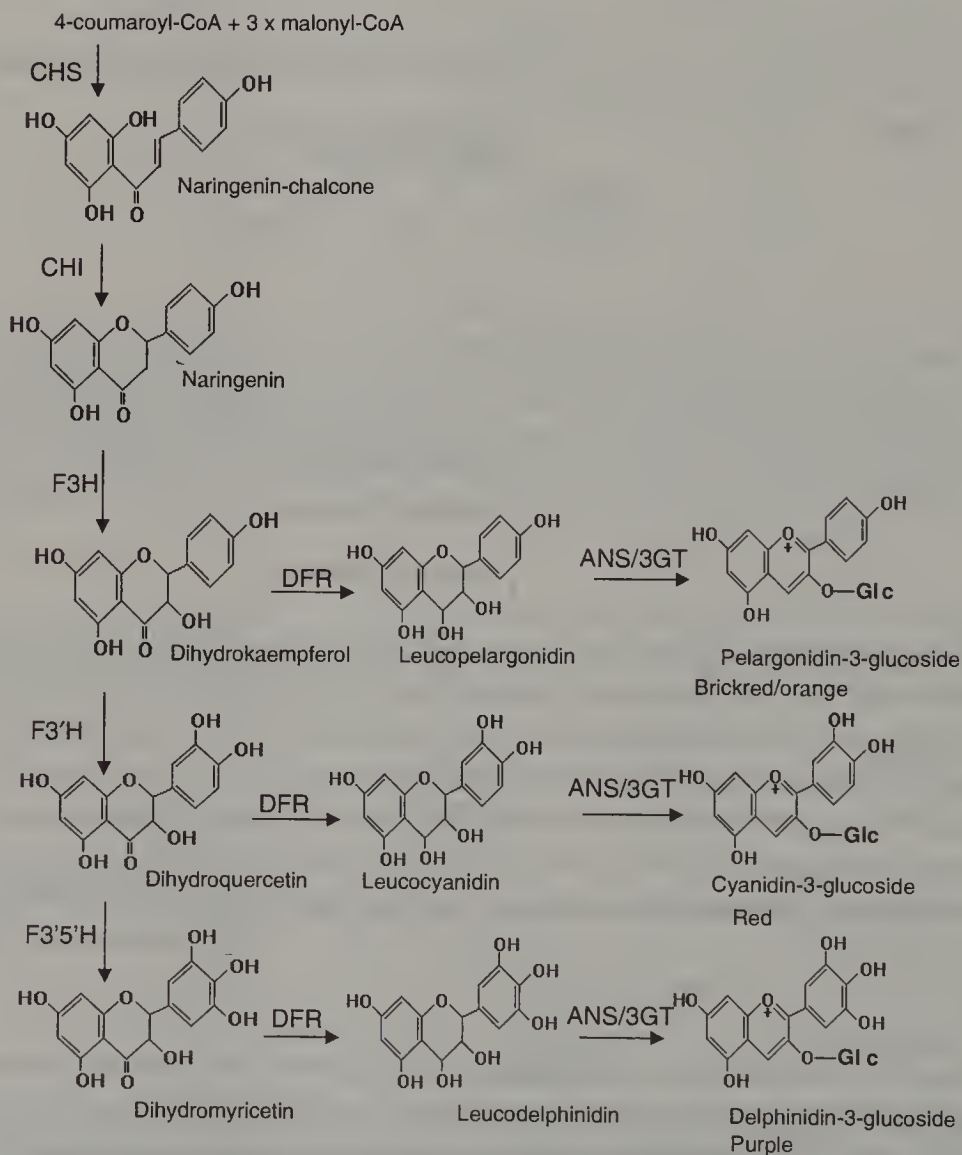


Figure 10.9 The carotenoid biosynthesis pathway. As shown, the synthesis of lycopene and β-carotene originates from the isoprenoid pathway. The sequential addition of 5-carbon isoprenoid units (isopentenyl diphosphate (IDP) and dimethylallyl diphosphate (DMADP)) to form 10-carbon (geranyl diphosphate (GDP)), 15-carbon (farnesyl diphosphate (FDP)) and 20-carbon (geranylgeranyl diphosphate (GGDP)) units is an important starting point for the synthesis of a range of compounds that include steroids, carotenoids, xanthophylls and anthocyanins. Note that gibberellins and ABA are synthesised from isoprenoid precursors, whilst cytokinins are formed via the addition of an isopentenyl unit to the N6 position of purines.

BOX 10.4 Flower colour

Flower colours are due mainly to flavonoids, carotenoids and betalains. The most important flavonoids in this respect are the anthocyanins, which contribute to red and blue flowers. Part of the anthocyanin pathway is shown below.

The first enzyme indicated is chalcone synthase, which was discussed in Chapter 8 in the context of gene silencing. The anthocyanin synthesis pathway is quite complex, at various stages the compounds can be modified by enzymes such as dihydroflavonol-4-reductase (DFR) to produce different coloured pigments. However, not all plants have all the enzymes of the pathway, so it is impossible to get some colours of flowers in some species. A combination of mutation and



Anthocyanin synthesis pathway. (CHS, chalcone synthase; CHI, chalcone flavanone isomerase; F3H, flavanone-3-hydroxylase; F3'H, flavonoid-3'-hydroxylase; F3'5'H, flavonoid-3',5'-hydroxylase; DFR, dihydroflavonol-4-reductase; ANS anthocyanidin synthase; 3GT, UDP-flavonol 3-O-glucosyltransferase.)

BOX 10.4 Continued

genetic manipulation of the anthocyanin pathway has the potential to generate colours in plants that previously were not possible. For example, petunias are unable to produce the pelargonidin-related pigments (brick red/orange) as their DFR is unable to use dihydrokaempferol. However, a line of petunias has been isolated with mutations in the flavonoid-3'-hydroxylase (F3'H) and the flavonoid-3',5'-hydroxylase (F3'5'H) genes. The plants are unable to use dihydrokaempferol in the other parts of the pathway, so there is a build-up of the substrate. This plant line has been transformed with a maize DFR gene. The expression of the maize protein allows the dihydrokaempferol to be utilised to produce the brick-red pigment pelargonidin-3-glucoside, so producing brick-red petunia flowers.

Companies such as Florigene have engineered the anthocyanin biosynthetic pathway in plants such as roses, carnations, chrysanthemums and gerberas to produce flowers in the blue spectrum (mauve-violet-blue). These flowers normally lack the anthocyanin, delphinidin, due to the absence of flavonoid-3',5'-hydroxylase (F3'5'H). The enzyme has been cloned and introduced into a range of flower types to produce blue flowers. Florigene introduced its first mauve-coloured carnation on to the market in 1996—called Moondust™—it was the world's first genetically modified flower on sale. Subsequently, other flower products have been marketed.

One interesting twist to this line of commercialisation, particularly for the regulators, is one of cultural practice. In the West, cut flowers would seem to a relatively safe product for marketing, they are often grown in greenhouses and do not represent a threat to human health or the environment. However, in countries such as Japan, many flower petals are used as decorations for foods and are often eaten. This is actually an important point, which indicates that care must be taken when considering what risks are involved when marketing/releasing transgenic plants. However the flip side to this is that some anthocyanin molecules are seen as nutraceuticals and eating anthocyanin-enhanced plant material may actually be a good thing.

pathway, and the resultant dwarfing. This is of interest because dwarfing is an important agronomic trait. Dwarf plants require less of the plant's resources to be committed to the growth of the stem, allowing more dry matter to be partitioned to the grain (the main commercial product) rather than the straw. This increase in harvest index directly improves yield, and, in addition, the dwarf crops are less prone to damage by wind and rain. The dwarf character was a major feature of the high-yielding varieties of wheat and maize that gave rise to the 'Green Revolution' (see Box 10.5).

The 'Green Revolution' strains of wheat (Box 10.5) were shown to be short because they respond abnormally to gibberellin. This trait is conferred by mutant dwarfing alleles at two reduced height-1 loci (*Rht-B1* and *Rht-D1*). Recent work has shown that the *Rht-B1/Rht-D1* and the maize dwarf-8 genes are orthologues of the *Arabidopsis* gibberellin insensitive (*gai*) gene. The *gai*

BOX 10.5 The 'Green Revolution'

The development of dwarf, high-yielding varieties of wheat and rice by conventional breeding during the 1960s enabled several developing countries, most notably in the Indian subcontinent, to move from a position of food scarcity to become net exporters of these cereals. This step-change in crop yield was termed the 'Green Revolution', and the key wheat breeder involved, Norman Borlaug, was awarded the Nobel Peace Prize for his contribution to global food security. The key to his success lay in the development of semi-dwarf, high-yielding, winter wheat varieties adapted to heavy rain conditions in the USA after the Second World War. This wheat was unsuitable for growing in the tropics or semi-tropical climates because, as a winter wheat, it required a period of cold during early growth. It was also found to be susceptible to rust fungus. Using the semi-dwarf wheat lines as a starting point, a breeding programme was established in Mexico to develop wheat strains that were adapted to the growing conditions found in poorer countries. This was done by crossing a number of desirable traits, from wheat varieties collected from around the world, into the semi-dwarf lines. The result was the production of spring wheat varieties that matured quickly and were insensitive to photoperiod, allowing them to be grown more than once per year. They were also rust-resistant and adapted to a variety of warm climates. These characteristics meant that yields could be doubled compared to those of traditional varieties, given adequate fertiliser.

At about the same time, similar breeding programmes were implemented for rice at the International Rice Research Institute (IRRI) in The Philippines. Some of the most successful crosses were between semi-dwarf japonica and relatively high-yielding indica lines, and rice varieties that were semi-dwarf, high yielding, early maturing, photoperiod-independent and blast fungus-resistant were produced. At IRRI, the new varieties could yield 3–4-fold more than traditional varieties under appropriate experimental conditions.

The Green Revolution required that farmers not only adopted the new cereal seeds, but also signed up to a high-input method of agriculture—including the use of nitrogen fertilisers, herbicides and pesticides as well as equipment for tilling and irrigation—to get the maximum yields from these varieties. There were significant cereal yield improvements, though often not as much as predicted from the experimental studies. Many small farmers were unable to afford to get the best out of the new seeds, whilst the richer farmers were able to produce more crops for less input of labour and thus started to acquire more land. The result was that smaller farmers were displaced from the land. Thus, the impressive increase in global food production was achieved at some cost in terms of higher agrochemical inputs, the loss of locally adapted crop varieties and increased poverty for some.

Analyses of current trends in population growth indicate the need for another step-change in crop yields comparable to the Green Revolution. It is clear that the genetic modification of crops could make an important contribution to raising the 'yield ceiling' reached by conventional breeding efforts. However, it is also important that the implementation of this 'Second Green Revolution' takes account of the negative aspects of the first Green Revolution.

genes encode transcription factor-like proteins that contain domains indicative of phosphotyrosine signalling. The mutant *gai* genes encode proteins altered in a conserved gibberellin signalling domain. Transgenic rice and wheat plants into which a mutant *gai* gene was introduced were shown to have a reduced response to gibberellin, and to have a dwarf phenotype. This result is particularly significant because, whilst the dwarf character has been found in rice, attempts to breed a dwarf strain of basmati rice have failed to date because the resulting short plants have lost their characteristic flavour. The success of the *gai* gene experiments suggests that the dwarf character could be introduced into a wide range of crop species to improve crop yield, without having to shuffle all the other genes that contribute to the desirable characteristics of the current elite lines. An important lesson here is that transgenic approaches to crop improvement, particularly those that use well-characterised genes, can be much more precise and predictable than conventional breeding. However, as we look at effects of modifying phytochromes in a later section, it will become clear that this is not always the case.

The final example of crop quality enhancement that flows from the study of carotenoid biosynthesis is that of provitamin A production in rice grains—so called ‘Golden Rice’.

CASE STUDY 4 **Golden Rice**

Rice is the most important food crop in the world, and is eaten by some 3.8 billion people. In some regions of the world where rice forms a staple component of the diet, vitamin A deficiency is a major nutritional problem. Deficiency of this vitamin can cause symptoms ranging from night blindness to total blindness as a result of xerophthalmia and keratomalacia. It has been estimated that around 124 million children are vitamin-A deficient, causing about 500 000 children to go blind each year. Vitamin A deficiency also exacerbates other health problems, for instance diarrhoea, respiratory diseases and childhood diseases such as measles. In consequence, it is estimated that improved vitamin A nutrition could prevent 1–2 million childhood deaths per year. One of the causes of vitamin A deficiency in regions where the majority of calories consumed comes from rice, is that milled rice contains no β -carotene (provitamin A). One of the solutions proposed for this problem is to engineer rice to produce provitamin A in the rice endosperm. Since the successful manipulation of β -carotene synthesis in the rice grains gives them a characteristic yellow/orange colour, rice which is genetically enriched in provitamin A has been described as ‘Golden Rice’.

Figure 10.9 shows that the biosynthetic pathway of provitamin A is a continuation of the lycopene pathway already discussed in the previous section on tomato ripening. Immature rice endosperm is capable of synthesising geranyl geranyl diphosphate, but subsequent stages of the pathway are not expressed in this tissue. Early transformation experiments with a phytoene synthase (*psy*) gene from daffodil fused to a rice endosperm-specific promoter indicated that phytoene could be synthesised from GGDP in the rice grain. However, three subsequent steps are required to convert phytoene to β -carotene; phytoene desaturase and ζ -carotene desaturase to introduce the double bonds to form lycopene, and

lycopene β -cyclase to form the rings in β -carotene. Fortunately, a bacterial carotene desaturase gene capable of introducing all four double bonds can be substituted for the phytoene desaturase and ζ -carotene desaturase. Nevertheless, the manipulation of Golden Rice requires the introduction of three genes: phytoene synthase, carotene desaturase and lycopene β -cyclase.

The constructs used to target expression of the appropriate genes to the rice endosperm are shown in Figure 10.10. The daffodil *psy* gene::rice glutelin promoter construct was inserted into the vector pZPsC, along with the bacterial carotene desaturase gene (*ctr1*) from *Erwinia uredovora* controlled by the 35S promoter. Both enzymes were targeted to the plastid (the site of GGDP synthesis): the *psy* gene by its own transit peptide, and the *ctr1* gene by fusion to a pea *rbcS* (ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit) transit peptide sequence. The lycopene β -cyclase gene from daffodil with a functional transit peptide was inserted into the vector pZLcyH under the control of the rice endosperm-specific glutelin promoter, along with a hygromycin-resistance selectable marker gene.

Rice immature embryos were inoculated with a mixture of *Agrobacterium* LBA4404 containing each of the two plasmids (see Chapter 3). A total of 60 hygromycin-resistant regenerated lines were selected at random, all of which were shown to contain the pZCyH construct. Of these, 12 were also found to contain the pZPsC cassette. Most of the seeds from these transgenic lines containing both constructs were found to be yellow, indicating carotenoid synthesis. A range of carotenoids was found in some of these lines, whereas β -carotene was the only carotenoid in others. The highest producing line of this type was found to contain $1.6 \mu\text{g } \beta\text{-carotene g}^{-1}$ endosperm from a mixed population of segregating grains. It is therefore calculated that the homozygous grains of the same line should produce at least $2 \mu\text{g g}^{-1}$ provitamin A, which corresponds to $100 \mu\text{g}$ retinol equivalents daily intake, assuming 300 g rice is consumed per day. This is sufficient to make a significant contribution to the daily intake of vitamin A, though it is probably not enough to provide the complete dietary requirement of the vitamin (see Box 10.6)

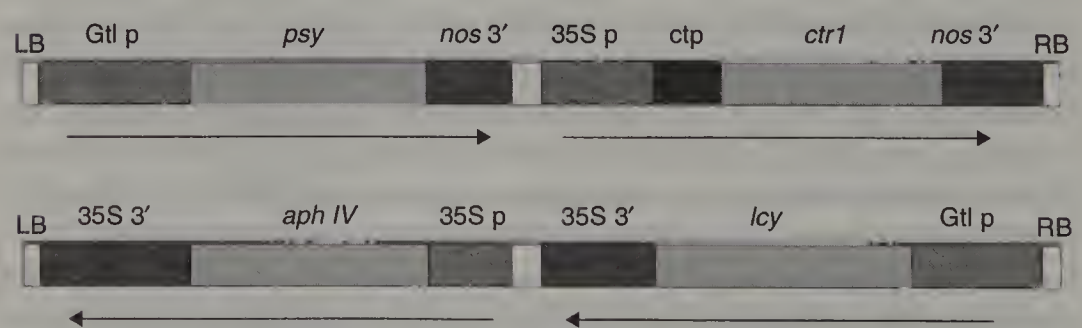


Figure 10.10 Constructs for the production of Golden Rice. The most successful strategy for the production of Golden Rice involved transformation with two independent constructs. The first one contains a daffodil phytoene synthase (*psy*) gene fused to a rice glutelin promoter in tandem with a bacterial carotene desaturase gene (*ctr1*) driven by the 35S promoter. The second construct contains the hygromycin-resistance *aphIV* selectable marker gene in tandem with the lycopene β -cyclase gene (*lcy*) of daffodil also fused to an endosperm-specific rice glutelin promoter.

Before carrying on to discuss the exploitation of Golden Rice, it is important to emphasise the scientific significance of these results. As shown in Figure 10.9, four enzymes are involved in catalysing five reactions from GGDP to β -carotene. The genetic manipulation of this multistep pathway, via the insertion of three genes into rice, required several years of intensive work and represents a considerable achievement. However, obtaining the laboratory-scale result is only the first of a series of hurdles that have to be negotiated before provitamin A-enriched rice reaches those who could benefit from it. Subsequent challenges included the clarification of intellectual property rights (IPR) and the mechanisms for technology transfer.

Golden Rice, IPR and technology transfer

One of the major obstacles to the delivery of Golden Rice to its potential beneficiaries proved to be the protection of intellectual property rights (IPR). Since Golden Rice was developed in public laboratories with public funding for humanitarian purposes, the aim was to patent-protect the process in order to fulfil the goal of supplying it to subsistence farmers free of charge and restriction. However, in compiling the international patent application, it was discovered that the research underpinning the development of Golden Rice had involved the use of procedures and technology protected by 70 IPRs and TPRs (technology protection rights) belonging to 32 different companies and universities! (One way of protecting technology rights is by the use of material transfer agreements (MTAs), which require that certain conditions are met when DNA constructs are requested from the originating laboratory. These conditions typically require that the DNA is used for research purposes only.) In order to be able to grant 'freedom to operate' licences to public research institutions in developing countries so they could introduce the trait into local varieties, it was necessary to obtain free licences from all the patent holders.

The size and complexity of the task of drawing up the patent application and obtaining these licences led the public scientists to involve a commercial partner (Zeneca) in the patent application, in return for an exclusive licence for commercial exploitation of the technology. (Commercial exploitation being defined as income from Golden Rice exceeding \$US10 000.) Fortunately, free licences for humanitarian purposes have been largely granted by the IPR holders, at least in part because of the enlightened self-interest of the biotechnology companies involved. Golden Rice is seen by the biotechnology industry as a valuable 'flagship' project to improve the public perception of GM technology (see Box 10.6). Conversely, there would be a lot of negative publicity for a company that was seen to be exploiting or standing in the way of the humanitarian aims of this project.

The next challenge faced by the developers of Golden Rice was that of technology transfer. It is important to realise that Golden Rice originated in Switzerland as a series of laboratory lines. These lines would be poorly adapted to the range of conditions found in different rice growing countries, for which locally adapted varieties and ecotypes have been developed. There is therefore an enormous task involved in transferring the Golden Rice trait into local varieties. In order to aid this technology transfer process, a 'Golden Rice Humanitarian Board' has been established. This board will ensure that, for each region, there is a proper assessment of needs and will set a framework for the optimal use of Golden Rice. It will also coordinate the assessments of bioavailability, food safety and allergenicity required before

BOX 10.6 The anti-GM opposition to Golden Rice

One of the surprising features of the Golden Rice programme has been the opposition it has generated from opponents of GM food, because Golden Rice appears to meet all the objections to GM crops that have been presented by anti-GM groups. Thus, as its originator, Ingo Potrykus, has pointed out:

1. It benefits primarily the poor and disadvantaged.
2. It will be given free of charge and restrictions to subsistence farmers.
3. It can be re-sown every year from the saved harvest.
4. It does not create advantages for rich landowners.
5. It was not developed by or for the biotechnology industry, and industry does not benefit from it.
6. It fulfils an urgent need by complementing (rather than displacing) traditional interventions.
7. It is a sustainable, cost-free solution to vitamin A deficiency, not requiring other resources for extraction, manufacture and distribution.
8. It avoids the unfortunate negative side-effects of the Green Revolution (see Box 10.5).
9. It creates no new dependencies.
10. It can be grown without any additional inputs.
11. It does not reduce agricultural biodiversity.
12. It does not affect natural biodiversity.
13. There is no negative environmental impact envisaged.
14. There is no conceivable risk to consumer health.
15. It was impossible to develop the trait by traditional methods.

However, Golden Rice has been attacked by groups such as Greenpeace who seem to view it as a 'Trojan horse' for other GM crops. That is, if they concede that Golden Rice is acceptable, they may open the door to other GM crops. Thus, a number of arguments have been deployed against Golden Rice. One is that the level of provitamin A in the rice is only sufficient to meet about 20% of the daily requirement for vitamin A, and that the vitamin A deficiency would be better met by encouraging a better, more varied diet of local vegetables (many of which used to be grown more widely, but have been replaced by rice). This argument seems to be self-defeating, since Golden Rice is being attacked for not meeting the complete dietary requirements for vitamin A, but also for encouraging a dependence on rice as the sole constituent of a less varied diet.

the rice is made available for human consumption. The Golden Rice material will only be given to public research institutions that can ensure proper handling and use of GM plants according to local rules and regulations. These institutions will transfer the Golden Rice trait into the best locally adapted lines by conventional breeding methods and/or *de novo* transformation of elite varieties. Socioeconomic and environmental impact studies will also be conducted to ensure that the technology does reach the poor without damage to the envi-

ronment. India will take a leading role in this technology transfer process and serve as a model for other countries to follow.

Engineering plant protein composition for improved nutrition

The example of Golden Rice has dealt with one aspect of nutritional quality. Another that has become a target for genetic engineering is the amino acid content of plant foods. Humans are only capable of synthesising 10 of the 20 naturally occurring amino acids. The other, 'essential', amino acids are obtained from the diet. Obtaining balanced quantities of the amino acids can be problematic if certain foods predominate within a diet. This is the case with cereal grains, which are commonly used as the principal energy source, and with legume seeds, which are important sources of proteins in the diet of humans and livestock. Cereal grains are often limiting for lysine, while the legume seeds have an adequate level of lysine but are limiting for the sulphur-containing amino acids, methionine and cysteine. It is important to note that animals can convert methionine into cysteine, but not the reverse, so it is possible to make up any short fall in the S-amino acids by increasing the level of the methionine. Transgenic routes have been used to increase the content of several different amino acids and some of the approaches will be discussed here.

One approach used has been to isolate the gene for a sulphur-poor protein and modify and enrich its nucleic acid sequence for the S-amino acid. This has been attempted for proteins such as (-phaseolin from *Phaseolus vulgaris* and vicilin from *Vicia faba*, but the modified proteins were either unstable or contained too little methionine to make them useful. A more successful approach has been to construct totally artificial genes that code for proteins containing a high S-amino acid content. One such totally synthetic protein, containing 13% methionine residues, has been successfully expressed in sweet potato (*Ipomoea batatas*). Several methionine-rich proteins have been identified in maize (21-kDa zein—28% methionine, 10-kDa zein—23% methionine); rice (10-kDa prolamin—20% methionine); sunflower (2S sunflower seed albumin—16% methionine, 8% cysteine); Brazil nut (Brazil nut albumin—18% methionine, 8% cysteine). The genes for these proteins have been introduced into a number of crops (maize, soybean, lupin, canola) to increase the level of S-amino acids in blended stock feeds containing cereal and legume grains. One major problem associated with moving proteins between species has become apparent with the Brazil nut albumin (BNA), i.e. the protein responsible for the potent allergenicity of Brazil nuts. This property is maintained in the seeds of BNA-containing transgenic plants, which makes them unacceptable for human consumption and has been a timely warning that safety issues should be paramount when dealing with food crops.

Some of the problems associated with these studies relate to the stability of the proteins and their localisation within the plant. This point can be illustrated by considering the expression of high-lysine proteins in the seed endosperm of cereals such as rice. These studies raise important issues about the types of protein that can be used and how they are localised in the grain, issues that should be considered when ectopically expressing proteins for whatever reason. In cereal endosperm, storage proteins accumulate primarily in protein storage vacuoles (PSVs) of terminally differentiated cells and as protein bodies (PBs) assembled directly within the endoplasmic reticulum (ER). Specific storage proteins are stored in these structures. In rice for example, there are two main types of storage protein: prolamins, which form accretions within the lumen of the ER, and glutelins (related to 11S globulins). Glutelins comprising up to 80% of the total seed protein are synthesised on rough ER and then transported to PSVs. It has recently been found that mRNAs for both types of protein are found in rough ER polysomes, but that the prolamins transcripts are preferentially located on membranes surrounding prolamins-containing protein bodies; while glutelin mRNAs are predominantly associated with polysomes on the cisternal ER.

One interesting study has been carried out with rice, where a lysine-rich protein from soybean has been expressed in the endosperm. To optimise expression and localisation, the soybean protein chosen was a glycinin, which is a member of the 11S globulin family. The important feature of the glycinin is that it is processed and stored in soybean by a process very similar to that found for the rice glutelins. When the transgenic rice plants were analysed it was found that both globulins (glutelin and glycinin) were expressed and formed complexes of discrete sizes that were targeted correctly. If this had not been the case then it would not be possible to obtain optimum levels of proteins. The other interesting point of glycinin proteins in soybean is that they are associated with a hyporcholesterolaemic effect (reduction in cholesterol) if the dietary intake of these proteins is in excess of 6 g/day. Therefore, expressing the soybean gene in rice potentially has two beneficial effects: it increases the amount of lysine in the endosperm protein; and it would contribute to the dietary levels of soybean globulins that lead to a reduction in the level of cholesterol.

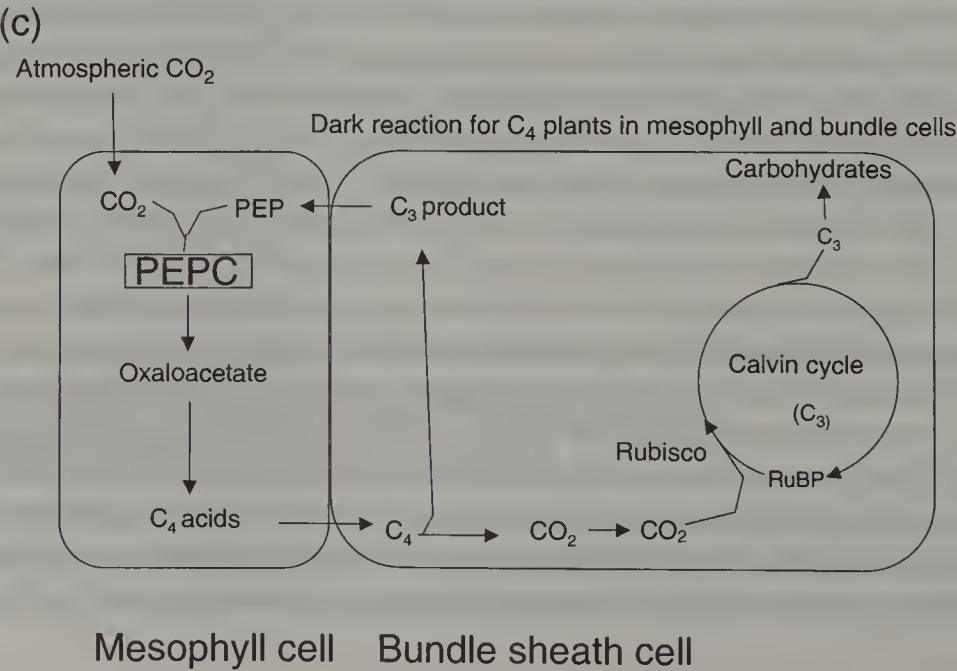
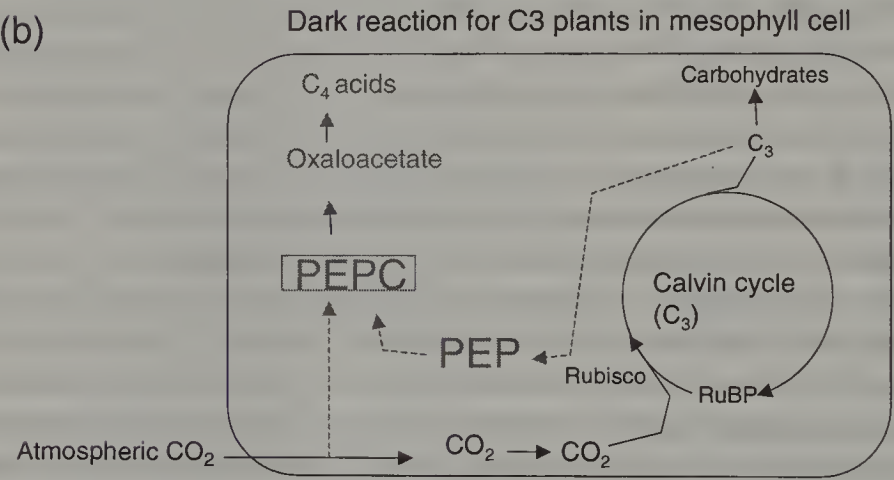
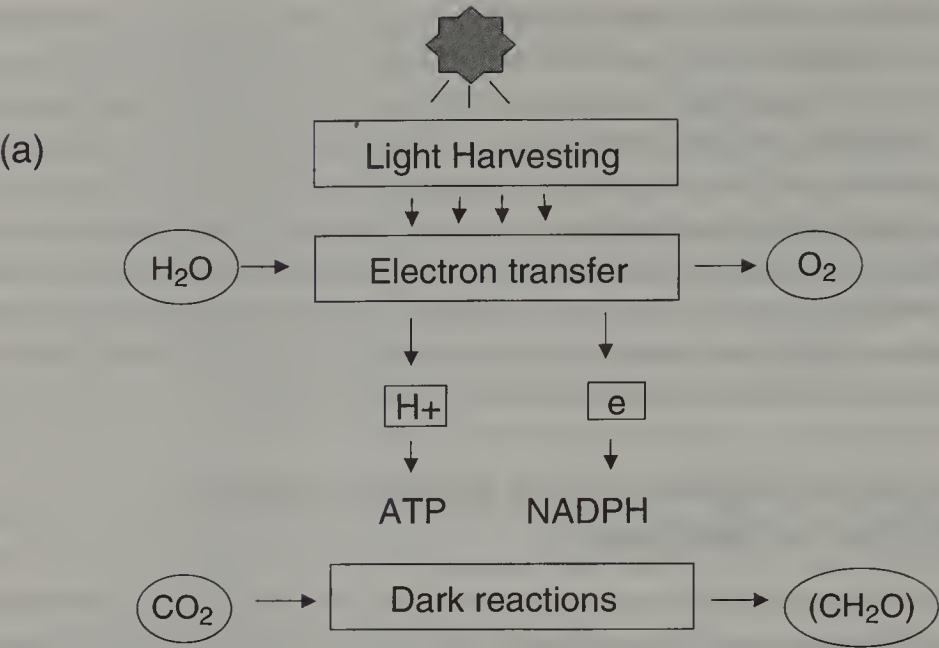
The genetic manipulation of crop yield by enhancement of photosynthesis

Crop yield is totally dependent upon light and its conversion (light harvesting and electron transport) into the usable energy (ATP, NADP) that drives the dark reactions of photosynthesis (carbon dioxide conversion into carbohy-

drates)—Figure 10.11. These reactions usually take place in chloroplasts within the leaves, the source, but the efficiency of the process is also dependent upon the capacity of sink tissues and organs to assimilate this fixed carbon. These are complex processes and outside the scope of this book, but some studies have been carried out to investigate the potential for enhancing photosynthesis by biotechnological means. Although these have been very preliminary in nature, several examples will be briefly discussed to describe the potential for manipulating complex physiological interactions. Direct manipulation of light harvesting, electron transfer or the processes of photoprotection and photoacclimation will not be considered here.

Manipulation of light harvesting and the assimilate distribution—phytochromes

When grown in the field or in their natural environments, plants are not always under optimum light conditions, even during daylight hours. They are faced with changes in the amount, direction, duration and quality of incident light radiation. Plants have evolved mechanisms that optimise the acquisition of available light energy for photosynthesis. In dense populations or in shady conditions, plants display the shade response. The plants use available resources to increase stem and petiole elongation to outgrow any shading plant; there is also a reduction in chlorophyll synthesis, leaf thickening and an increase in apical dominance. From an agricultural point of view, this shade response has important implications for yield, as assimilates are used to re-establish optimal light conditions for photosynthesis rather than being stored. The shade response is regulated through a series of photoreceptors. The phytochrome family of proteins are the best characterised of these proteins (there are at least five Phy proteins in dicots). They are able to detect the level and quality of the incident light and then control growth and development via a series of signal transduction pathways that regulate gene expression. Phy proteins are photochromic proteins, they have a protein moiety connected to a tetrapyrrole chromophore, and they can exist in two forms: physiologically inactive (P_R , red-light absorbing) and active (P_{FR} , far-red light absorbing). These forms are interconvertible by red or far-red light, respectively. The photosynthetic pigments in leaves absorb most visible radiation (400–700 nm) but reflect light beyond 700 nm (far red, 700–800 nm). This leads to a high proportion of FR radiation being found in dense plant stands. Plants are able to perceive this radiation and the proximity of other plants because the equilibrium of the phytochromes is shifted towards the inactive form. Two of the major phytochromes involved in this response are PhyA and PhyB. PhyA accumulates in the dark and is rapidly degraded upon conversion, by the absorption of red light, into the labile P_{FR} form. Despite this, it is responsible for the detection of continuous FR light and dampens the shade-avoidance response. PhyB is responsible for the detection of red light, so



under high light conditions (high relative proportions of red light) it is converted to the active form, in which form it suppresses the shade-avoidance response.

The manipulation of phytochromes is a general approach that can have pleiotropic effects on photosynthesis and assimilate distribution. Analysis of these systems suggests that phytochrome manipulation is likely to have important consequences on photosynthesis (the phytochromes coordinately regulate many of the genes involved in photosynthesis) and the sink–source relationship. So, a number of different transgenic plant types, ectopically expressing either PhyA or PhyB, have been constructed and some interesting effects have been observed. Plants expressing the phytochromes at very high levels were seriously affected in their growth, but it was possible to identify plants in which the phytochromes improved the harvest index under high densities in the field. Several examples will be discussed in more detail below.

In one set of experiments, tobacco plants, engineered to overexpress an oat *PHYA* cDNA, were grown at five different planting densities. At the higher planting density the wild-type tobacco demonstrated the shade response, in that the plants were normally taller than those planted at low densities. With the transgenic lines, they elongated at lower rates—so, at the high planting densities, relative to the lower planting densities, the result was dwarfed plants. The ectopically expressed PhyA was not only damping the shade response in response to far-red light, but was actually reducing stem elongation to below the ‘normal levels’ found in plants grown in optimal light conditions. Overall yield was not affected, but assimilates showed an enhanced allocation to leaves with increases in the harvest index approaching 20%. So the beneficial outcome of this approach was that at high planting densities assimilates were not being utilised in stem growth but were being utilised in leaf production or were being stored.

Similar dwarfing effects were found in an experiment with transgenic potato plants containing an ectopically expressed *Arabidopsis PHYB* sequence. However, overexpression of the phytochrome also led to an increase in

Figure 10.11 Photosynthesis. (A) Photosynthesis can be divided into several separate steps: light harvesting and electron transfer, in which light energy is converted into useful chemical energy (ATP, NADPH); and the dark reaction where CO_2 is fixed. (B) In C_3 plants, the fixation is carried out by Rubisco (ribulose 1,5-bisphosphate carboxylase), as part of the Calvin cycle (C_3 cycle), in the chloroplasts of mesophyll cells (solid line). (C) In C_4 plants, CO_2 fixation occurs in mesophyll cells, via phosphoenolpyruvate carboxylase (PEPC), to form C_4 acids such as malate and aspartate. These are transferred to adjacent bundle-sheath cells where metabolism of the C_4 acids generates high concentrations of CO_2 that are used by Rubisco in a normal C_3 cycle. In C_3 transgenic plants containing PEPC (B), there is a conflict between the existing C_3 cycle and the presence of the PEPC, which is likely to drain C_3 precursors from the Calvin cycle so limiting fixation into carbohydrates (hatched line). PEP, phosphoenolpyruvate. (In part redrawn with permission from Figure 1, p. 22 in Edwards, G. (1999).)

chloroplast number and an overall increase in photosynthesis per leaf area. Photosynthesis was also less sensitive to photoinactivation under prolonged light stress. Although leaf senescence was not delayed, chlorophyll degradation was, leading to a prolonged period of photosynthetic activity. The increase in the rate of photosynthesis and the prolonged period of photosynthetic activity resulted in an increase in biomass. This was demonstrated by extended underground organs (roots/shoots) and with a large number of small tubers. In fact, tuber yields were increased by as much as 50% in weight.

The pleiotropic nature of this approach has some overlap with strategies to prolong the photosynthetic period of plants by delaying leaf senescence. One such strategy has involved the production of transgenic plants containing the senescence-related *SAG12* promoter (*SAG*, senescence activated gene) driving expression of the *Agrobacterium* cytokinin biosynthetic gene *ipt* (see Chapter 3). As the leaves begin to senesce, the expression of the *ipt* gene is switched on, thereby increasing the cytokinin level in the leaf. This delays senescence and maintains chlorophyll content. The prolonged period of photosynthetic activity is expected to increase assimilate production. However, it is not clear if this strategy will be successful as there may be problems with nitrogen metabolism. As the nitrogen released from senescing leaves is required for other developmental processes, such as grain filling in cereals, its maintenance in the leaf is unlikely to be beneficial.

Direct manipulation of photosynthesis—enhancement of dark reactions

The sheer complexity of photosynthesis and assimilate storage suggests the task of increasing yield will require whole suites of genes—nuclear and chloroplast—to be enhanced. The phytochrome and the *ipt* approaches allow for this, in that they have pleiotropic effects. It is likely that different crops will require different strategies to be adopted. This is the case for C_3 and C_4 plants (this refers to whether the plant uses a 3-carbon or a 4-carbon route in the fixation of CO_2), which will be discussed in the next section. However, identifying limiting steps in pathways can allow increases in yield to be obtained. This has recently been found with tobacco where two enzymes in the Calvin cycle (C_3 cycle, Figure 10.11(B)) have been shown to be limiting. Most of the enzymes in the Calvin cycle are present at levels in excess of those required to sustain a continued rate of CO_2 fixation. However, levels of fructose-1,6-bisphosphatase (FBPase) and sedoheptulose-1,7-bisphosphatase (SBPase) are extremely low compared with those of other enzymes in the Calvin cycle. Transgenic tobacco plants expressing a single gene for a chloroplast-targeted, dual-function cyanobacterial fructose-1,6/sedoheptulose-1,7-bisphosphatase show enhanced photosynthetic efficiency and growth characteristics. Dry matter and photosynthetic CO_2 fixation were 1.5- and 1.24-fold higher in the transgenic plants compared with wild-type tobacco. It was also found that

ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) activity was increased, and that various Calvin cycle intermediates and the accumulation of carbohydrates were also higher. These results seem to relate to an increased metabolic flux, and support the notion that the enzymes were limiting the photosynthetic pathway in tobacco.

Perhaps the most challenging approach attempted has been to introduce various enzymes from the C_4 photosynthesis pathway into C_3 crop species. The C_3 plants—which include wheat, oats and soybean—have chloroplast-containing mesophyll cells in the leaves. CO_2 is assimilated in these chloroplasts, via Rubisco, as part of the Calvin cycle, and the first product from this is the C_3 molecule 3-phosphoglyceric acid (Figure 10.11(B)). This is not a very efficient system as the CO_2 concentration in C_3 plants can be a rate-limiting factor. Also, Rubisco can use O_2 as an alternative substrate for reactions with ribulose-1,5-bisphosphate, which ultimately leads to photorespiration (metabolism of C_3 molecules leading to the release CO_2 in mitochondria). In essence, O_2 is a competitive inhibitor with respect to CO_2 fixation by Rubisco. In addition, high temperature decreases the availability of CO_2 to Rubisco, and water stress (by closing stomata) increases the resistance to CO_2 diffusion into the leaf. In C_4 plants, such as maize and sugar cane, two different types of chloroplast-containing cells are involved in CO_2 assimilation. CO_2 is first fixed in mesophyll cells by phosphoenolpyruvate carboxylase (PEPC) (Figure 10.11(C)). C_4 acids then migrate into associated bundle-sheath cells where the C_4 pathway generates a high concentration of available CO_2 which is then used by Rubisco. This system is very efficient under all the conditions where the C_3 system fails.

Recently, a complete maize PEPC gene (maize is a C_4 plant) has been introduced into rice (a C_3 plant) with surprising results. Many of the plants obtained showed high levels of expression, two or three times that found in maize, and the enzyme accounted for up to 12% of the total leaf protein. The most important observation was that the O_2 sensitivity of photosynthesis in rice was decreased in the transformed plants. Why this should be the case was not clear, but it suggested that carbon fixation by the maize enzyme increased the supply of CO_2 to Rubisco. This experiment raises the possibility of manipulating other stages of the photosynthetic pathways. However, it is also likely to be a very difficult process, even the expression of the PEPC gene had drawbacks. Expression of PEPC in the absence of the other enzymes of the C_4 pathway led to a drain of carbon into C_4 acids rather than the carbohydrates, so reducing plant yield (Figure 10.11(B)).

Summary

This chapter has reviewed a number of specific examples of the genetic modification of yield and quality traits. The manipulation of tomato ripening has been considered in

some detail because it is possible to demonstrate a number of techniques and approaches with this topic. Tomato ripening was one of the first systems in which antisense technology was used to manipulate endogenous gene expression, such that levels of enzyme activity of <1% of the wild type were achieved. Antisense techniques were exploited to confirm the role of specific genes in the ripening process (PG) and identify ripening-related genes (ACC oxidase). The technique was also applied to the modification of ripening for commercial targets. Subsequent work on this system led to the recognition that endogenous gene expression was also inhibited by sense constructs, and that the phenomenon of co-suppression could also be used for genetic manipulation purposes. Three of the important genes that were identified and exploited exemplify three distinct target areas for genetic modification. The manipulation of PG demonstrates the possibility of manipulating the expression of a single gene that has very little detectable effect on related processes in the plant—softening can be inhibited without affecting the other ripening events. In contrast, the manipulation of a regulatory pathway (ethylene production) has profound effects on the entire ripening process. The third example, phytoene synthase, demonstrates the potential for, and pitfalls of, manipulating complex biochemical pathways in plants. This potential has been exploited most impressively in the engineering of Golden Rice. Enhancement of nutritional quality has also been discussed in respect of amino acid content. This area is likely to become more accessible to manipulation as metabolomic studies provide information on plant pathways.

Several preliminary studies on the enhancement of photosynthetic pathways were presented in this chapter. These indicate that it might be possible increase yield levels by manipulating the dark reactions. However, it is clear that this will not be a rapid process for a lot more fundamental work is required in this area. Some work is also underway in understanding how it may be possible to manipulate the light reactions of photosynthesis. Again these are complex areas, and a lot more basic research is required before it will be possible to optimise yield production through the manipulation of light harvesting or the energy production aspects of photosynthesis.

Further reading

Crop yield and quality

- Chapple, C. and Carpita, N. (1998). Plant cell walls as targets for biotechnology. *Current Opinion in Plant Biology*, 1, 179–85.
- Mazur, B., Krebbers, E. and Tingey, S. (1999). Gene discovery and product development for grain quality traits. *Science*, 285, 372–5.

Tomato ripening

- Bouzayen, M., Hamilton, A., Picton, S., Barton, S. and Grierson, D. (1992). Identification of genes for the ethylene-forming enzyme and inhibition of ethylene synthesis in transgenic plants using antisense genes. *Biochemical Society Transactions*, 20, 76–9.
- Grierson, D. and Covey, S. (1988). *Plant molecular biology* (2nd edn). Blackie, Glasgow and London, UK
- Grierson, D. and Schuch, W. (1993). Control of ripening. *Philosophical Transactions of the Royal Society, London*, 342, 241–50.
- Hamilton, A. J., Lycett, G. W. and Grierson, D. (1990). Antisense gene that inhibits synthesis of the hormone ethylene in transgenic plants. *Nature*, 346, 284–7.

Smith, C. J. S., Watson, C. F., Ray, J., Bird, C. R., Schuch, W. and Grierson, D. (1988). Antisense inhibition of polygalacturonase expression in transgenic tomatoes. *Nature*, 334, 724–6.

Flower colour

Florigene website—web-link 10.1: <http://www.florigene.com>

The ‘Green Revolution’

Chrispeels, M. J. and Sadava, D. (1994). *Plants, genes and agriculture*. Jones and Bartlett Publishers, London.

Norman Borlaug. A listing of the best and most comprehensive online resources pertaining to Norman Borlaug—web-link 10.2: <http://www.ideachannel.com/Borlaug.htm>

Peng, J. R., Richards, D. E., Hartley, M. N., Murphy, G. P., Devos, K. M., Flintham, J. E., Beales, J., Fish, L. J., Worland, A. J., Pelica, F., Sudhakar, D., Christou, P., Snape, J. W., Gale, M. D. and Harberd, N. P., (1999). ‘Green Revolution’ genes encode mutant gibberellin response modulators. *Nature*, 400, 256–61.

Golden Rice

Potrykus, I. (2001). Golden Rice and beyond. *Plant Physiology*, 125, 1157–61.

Ye, X., Al-Babili, S., Kloti, A., Zhang, J., Lucca, P., Beyer, P. and Potrykus, I. (2000). Engineering the provitamin A (β -carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science*, 287, 303–5.

Enhancement of photosynthesis

Ku, M. S. B., Agarie, S., Nomura, M., Fukayama, H., Tsuchida, H., Ono, K., Hirose, S., Toki, S., Miyao, M. and Matsuoka, M. (1999). High-level expression of maize phosphoenolpyruvate carboxylase in transgenic rice plants. *Nature Biotechnology*, 17, 76–80.

Miyagawa, Y., Tamoi, M. and Shigeoka, S. (2001). Over expression of a cyanobacterial fructose-1,6-sedoheptulose-1,7-bisphosphatase in tobacco enhances photosynthesis and growth. *Nature Biotechnology*, 19, 965–9.

Robson, R. H., McCormac, C., Irvine, S. and Smith, H. (1996). Genetic engineering of harvest index in tobacco through overexpression of a phytochrome gene. *Nature Biotechnology*, 14, 995–8.

Thiele, A., Herold, M., Lenk, I., Quail, H. and Gatz, C. (1999). Heterologous expression of *Arabidopsis* phytochrome B in transgenic potato influences photosynthetic performance and tuber development. *Plant Physiology*, 120, 73–81.

Enhancement of protein content

Herbers, K. and Sonnewald, U. (1999). Production of new/modified proteins in transgenic plants. *Current Opinion in Biotechnology*, 10, 163–8.

Herman, E. M. and Larkins, B. A. (1999). Protein storage bodies and vacuoles. *Plant Cell*, 11, 601–14.

Katsube, T., Kurisaka, N., Ogawa, M., Maruyama, N., Ohtsuka, R., Utsumi, S. and Takaiwa, F. (1999). Accumulation of soybean glycinin and its assembly with glutelins in rice. *Plant Physiology*, 120, 1063–73.

Tabe, L. and Higgins, T. J. V. (1998). Engineering plant protein composition for improved nutrition. *Trends in Plant Science*, 3, 282–6.

Molecular farming/‘pharming’

Introduction

‘Molecular farming’ is a term coined to describe the application of molecular biological techniques to the synthesis of commercial products in plants. The term can be applied to a broad spectrum of activities, from the enhanced production of products that are already extracted from plants through to the manufacture of compounds that are completely novel to plants. A wide range of products have already been identified as likely targets for molecular farming, these include a variety of carbohydrates, fats and proteins, as well as secondary products. This chapter will consider examples of some of the major classes of molecules produced in transgenic plants. The first part of this chapter will look at some examples of molecular farming as applied to carbohydrates and oils. In the second part of the chapter, examples of molecular farming in which it is the transgenic protein itself that is the product will be looked at. Many of the proteins being ‘farmed’ in plants are antibodies, vaccines or biopharmaceuticals aimed at improving human and animal health, and it is in the area of molecular farming that many of the most exciting and potentially beneficial developments in plant biotechnology are taking place. At the end of the chapter, we will also review some of the economic reasons why molecular farming might, in some cases, be an attractive alternative to current forms of manufacture of these compounds.

Carbohydrates and lipids

In this first section we will look at some examples (see Table 11.1 for a summary) of the production of novel or modified carbohydrates and oils (and their derivatives) in plants. Many of these examples have multiple applications (both as modified food- or feedstuffs and industrially) that will be highlighted here.

Table 11.1 Major targets for carbohydrate and lipid molecular farming

Compound	Origin of genes	Application	Transgenic plant
<i>Carbohydrates</i>			
Amylose-free starch	Potato	Food, industrial	Potato
High amylose starch	Potato	Food, industrial	Potato
Cyclodextrins	<i>Klebsiella pneumoniae</i>	Food, pharmaceutical	Potato
Fructans	<i>Bacillus subtilis</i>	Industrial	Tobacco, potato, maize
	Jerusalem artichoke	Food	Sugar beet
Trehalose	<i>Escherichia coli</i>	Food	Tobacco
<i>Lipids</i>			
Medium-chain fatty acids	California bay tree	Food, detergent, industrial	Oilseed rape
Saturated fatty acids	<i>Brassica rapa</i>	Food	Oilseed rape
Mono-unsaturated fatty acids	Rat	Food	Tobacco
Polyhydroxybutyrate	<i>Alcaligenes eutrophus</i>	Biodegradable plastics	<i>Arabidopsis</i> , soybean, oilseed rape, cotton

Carbohydrate production

Plants produce a range of commercially valuable carbohydrates. The two most abundant carbohydrates are cellulose (used for fibres from cotton and flax, for paper-making from trees and for a range of industrial products such as paints and polymers) and starch (used for food, feed and industrial purposes). Some biotechnological effort is directed towards improving the yield and quality of these bulk carbohydrates. However, there are a number of other carbohydrates that it could be attractive to produce in transgenic plants, including oligofructans, cyclodextrins and trehalose (Figure 11.1).

Figure 11.2 indicates the biosynthetic pathways and cellular locations of these compounds in plants. One obvious point is that carbohydrates are synthesised and stored in a number of different cellular compartments. Thus, starch and its derivatives are synthesised in plastids, whilst the fructans, another important class of storage carbohydrates in plants, are synthesised and stored in the vacuole. Sugars and sugar alcohols are synthesised in the cytosol and accumulate throughout the cell, often in response to abiotic stress (see Chapter 9). Note also that Figure 11.2 shows a stylised cell, and that in practice different carbohydrates tend to accumulate in different tissues and organs. For example, starch is stored transiently in the leaves of cereals, and starch reserves accumulate in the amyloplasts of the grain endosperm. However, the fructans are also a significant storage carbohydrate in some cereals, and accumulate in the leaves and stems.

CASE STUDY 1 Starch

The major starch-producing crops—cereals and potatoes—are already widely grown to produce starch as a chemical feedstock. Approximately 70% of the starch produced in Europe and the USA is used for a variety of industrial purposes, with only 30% used for human consumption and animal feed. In Figure 11.2, the pathway for starch synthesis in chloroplasts is shown, starting with triose-phosphate, which is generated in the Calvin cycle. This is subsequently converted to hexose phosphate and then ADP-glucose, which is the substrate for starch synthesis. In the case of amyloplasts, which are plastids specialised for starch accumulation, the hexose molecules are transported directly into the amyloplast. In the cereal endosperm, ADP-glucose may be synthesised in the cytosol and transported directly into the plastids. Starch synthesis involves two classes of enzymes. Starch synthase (SS) catalyses the addition of glucose residues from ADP-glucose to the non-reducing end of the growing chain, forming $\alpha(1\rightarrow4)$ links. One form of starch molecule, amylose, is composed entirely of unbranched chains of glucose of about 1000 residues long. There are several SS isoenzymes, some of which are soluble in the plastid stroma, and others which are bound to the insoluble starch granules (granule-bound starch synthase, or GBSS).

The other class of starch biosynthetic enzymes is the starch branching enzymes (SBE) which create $\alpha(1\rightarrow6)$ branches in the starch molecule. These branched starch molecules are called amylopectin and typically contain 10^4 – 10^5 glucose residues. The branches are

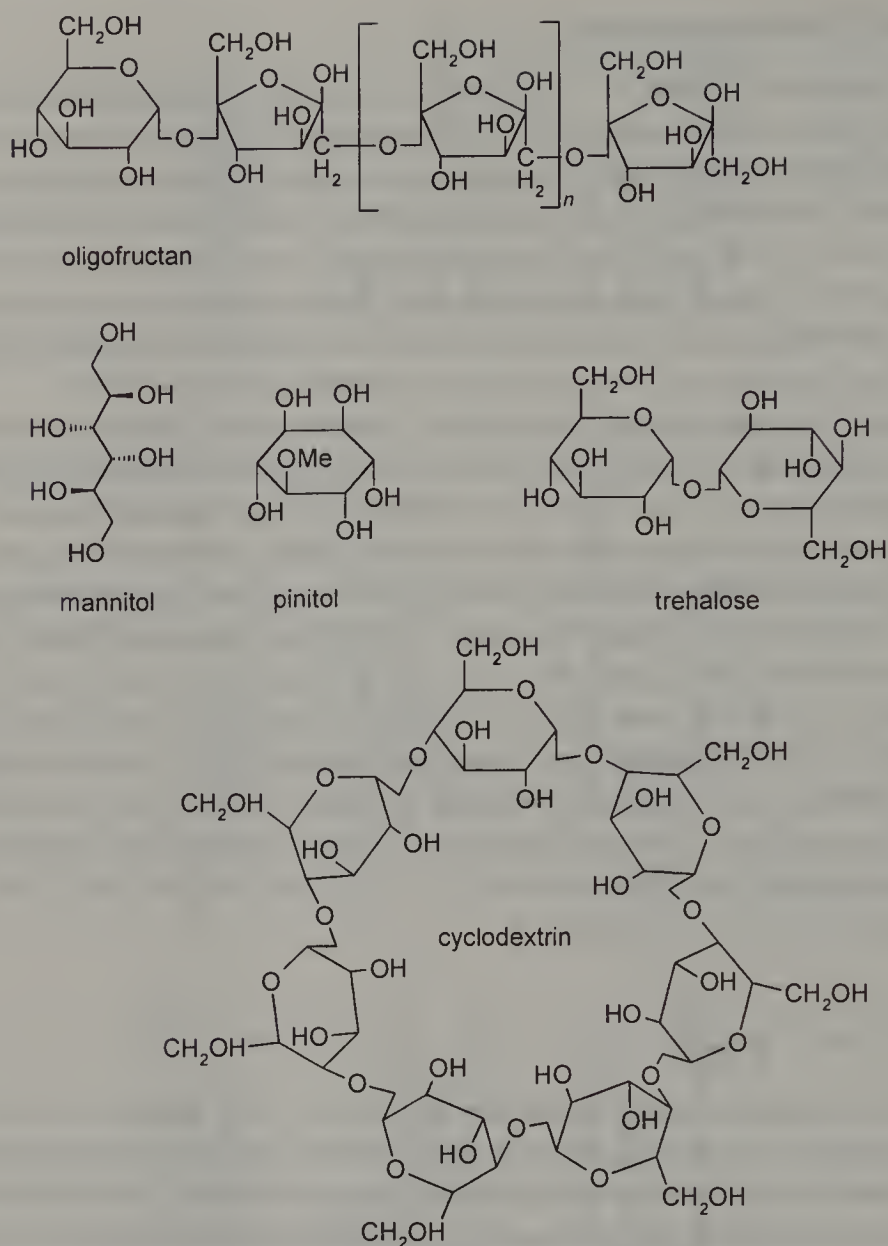


Figure 11.1 Target carbohydrate products for molecular farming. The structure of some potential targets for molecular farming are shown. Oligofructans are polymers produced by the sequential addition of fructose on to an initial sucrose molecule. The sugar alcohols mannitol and pinitol were discussed in Chapter 9 with reference to their roles as osmolytes, and the potential for producing them in crops to enhance stress tolerance. Trehalose is a disaccharide with useful properties as a food preservative. Cyclodextrins are rings of 6, 7 or 8 glucose residues. The 7-sided ring shown here has a number of pharmaceutical applications for the solubilisation of small hydrophobic molecules.

actually formed by cleaving an $\alpha(1\rightarrow4)$ linkage in the chain and joining the reducing end of the cut fragment to the 6-carbon position of a glucose residue some 20 residues downstream. This mode of action ensures that the branches in amylopectin are not random, but occur roughly every 20 glucose residues.

The simple diagram of starch biosynthesis shown in Figure 11.2 illustrates some important principles in the genetic engineering of metabolic pathways. There is a single step from

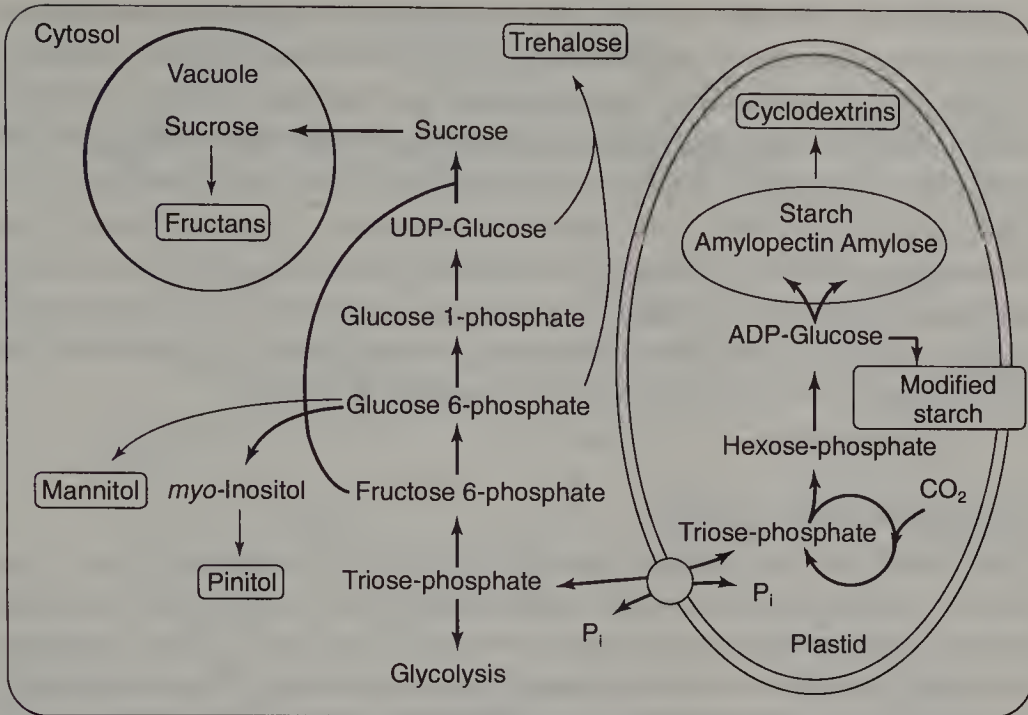


Figure 11.2 Metabolic pathways for the biosynthesis of carbohydrate products for molecular farming. The biosynthetic pathways for the production of the compounds shown in Figure 11.1 are shown schematically in this stylised cell. Starting from the pool of hexose-phosphate sugars in the cytoplasm, sucrose is synthesised via UDP-glucose and transported into the vacuole, where it may be stored as fructans. UDP-glucose and glucose 6-phosphate are used to form trehalose in the cytoplasm, and glucose 6-phosphate is also the precursor of the sugar alcohols. In leaf cells, triose-phosphates are transported in and out of the chloroplast, where they can be converted to hexose-phosphate and then ADP-glucose for starch synthesis. Cyclodextrins are not found in plants, but starch is a substrate for some bacterial biosynthetic enzymes. (Redrawn with permission from Goddijn, O. J. M. and Pen, J. (1995).)

a pool of common primary metabolites (hexose phosphate sugars) to the immediate precursor of starch, ADP-glucose, catalysed by the enzyme ADP-glucose pyrophosphorylase. This precursor molecule is specific to the starch pathway, so this step commits hexose sugars to starch biosynthesis. Starch comprises amylose and amylopectin, so effectively there is a branch in the biosynthetic pathway leading to one or the other form of starch. Thus, this pathway could be manipulated at three points, in a number of different ways. The production of ADP-glucose could be altered, which would affect the overall level of starch biosynthesis. Thus, increasing the amount or activity of the enzyme ADP-glucose pyrophosphorylase should increase the amount of starch produced. On the other hand, manipulation of one of the branches in the pathway would affect the ratio of amylose to amylopectin.

The proportion of amylose:amylopectin is normally about 20–30% amylose to 70–80% amylopectin, and it is this ratio that has the greatest influence on the physicochemical properties of the starch. For some applications it would be advantageous to increase the proportion of amylopectin. For example, many of the uses of starch in food production involve the formation of a gel after heating the starch in water and cooling. Amylose molecules tend

to aggregate and crystallise on cooling, whereas amylopectin gels are more stable and generably more desirable for food processing. On the other hand, a high-amylose starch with limited branching would be a valuable feedstock for industrial purposes. In potatoes, starch synthesis involves the activities of three isoforms of soluble starch synthase (SS), one granule-bound starch synthase (GBSS1) and two isoforms of the starch branching enzyme SBE (SBE A and B). The GBSS is responsible for the synthesis of amylose chains; the successful antisense inhibition of GBSS1 in potato produced an amylose-free starch. On the other hand, the engineering of a high-amylose starch required the antisense inhibition of both SBE A and SBE B. The resulting high-amylose starch also had much higher phosphate levels and could prove to be useful for food and industrial applications.

Cyclodextrins from starch

Whilst starch itself is a bulk feedstock for industry, the potential exists for carrying out further biotransformations on the starch in the plant, rather than by chemical or fermentation processes after extraction. One type of high-value product that could be made from starch is the cyclodextrins. These compounds are typically 6-, 7- or 8-membered rings comprising glucopyranose subunits attached in $\alpha(1\rightarrow4)$ linkages (Figure 11.1). These compounds are normally produced by bacterial fermentation of maize starch, particularly for pharmaceutical applications. The structure of the cyclodextrins is such that they form a cone-shaped ring, with the hydrophilic residues on the exterior and a hydrophobic pocket in the centre of the ring. The 7-membered ring has the ideal dimensions to form a pocket for small hydrophobic compounds. Thus, in a concentrated suspension, the cyclodextrins will effectively solubilise hydrophobic pharmaceuticals such as steroids. At lower concentrations, for example after injection into the bloodstream, the therapeutic agent is released.

Having been identified as a suitable target for molecular farming, it is striking that only one report, in 1991, of an attempt to produce cyclodextrins has been published. A bacterial cyclodextrin glycosyltransferase gene from *Klebsiella pneumoniae* was fused to a plastid-targeting sequence and placed under the control of the promoter from the patatin gene. Patatin is a protein that accumulates in potato tubers, and the promoter directs high levels of expression in the tubers of transgenic potatoes. However, transformation of potatoes with this construct resulted in very little conversion (0.001–0.01%) of starch to cyclodextrins. It was concluded that the insoluble starch granules may have been inaccessible to the bacterial enzyme, or that the enzyme became trapped in the growing granule. Whatever the reason, no subsequent attempts to produce this commodity have appeared in the literature.

CASE STUDY 2 **Polyfructans**

Another example of a carbohydrate targeted for production in transgenic plants is polyfructans. These compounds are soluble polymers of fructose that are synthesised and

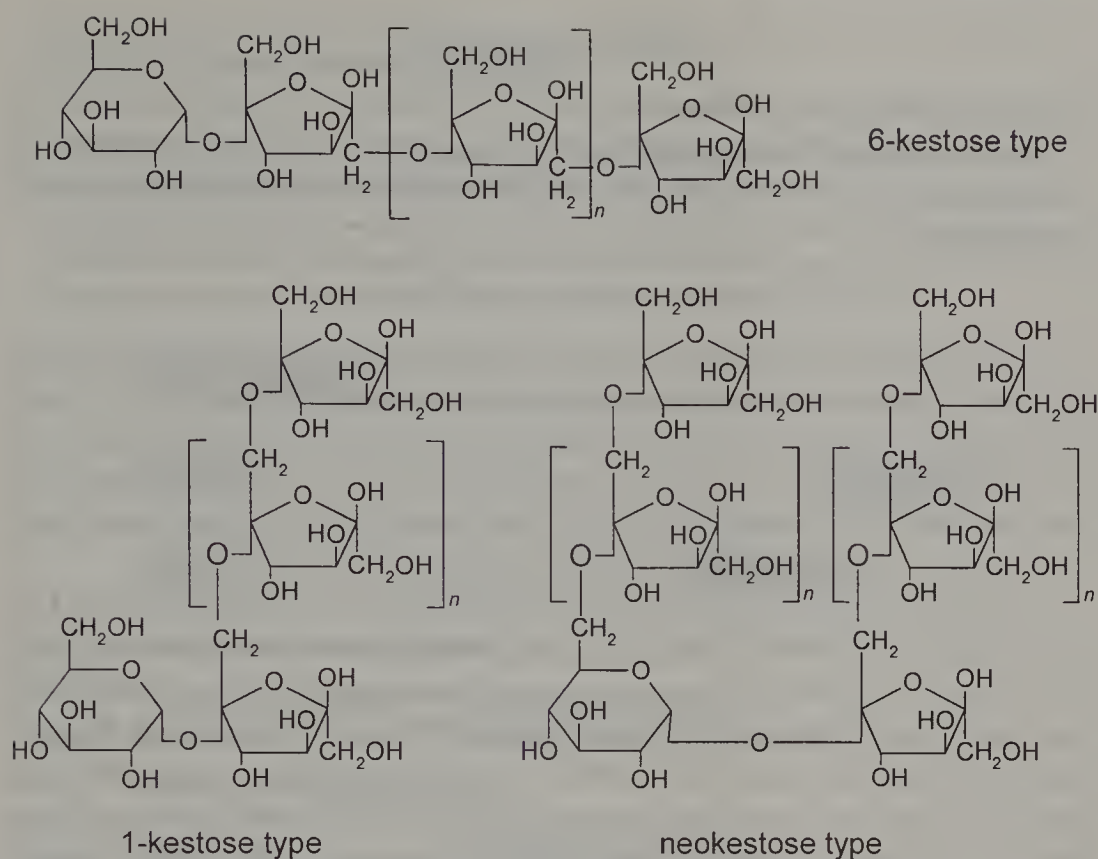
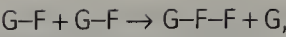


Figure 11.3 Structure and synthesis of polyfructans. The structures of three types of polyfructan are shown. The levan-type fructans of the 6-kestose class are formed by fructose-(6→2 β)-fructose linkages. These fructans are typically found in grasses, with $n = \sim 200$. (In bacteria, levans are much larger polymers than this). The 1-kestose type of fructans are also called inulins, and are formed by fructose-(1→2 β)-fructose linkages. These are typically found in bulbs and the storage roots of plants such as Jerusalem artichoke and chicory, with $n < 50$. Fructans of the neokestose type, with 1-kestose chains attached to both the glucose and fructose residues of the initial sucrose moiety, are the smallest fructans, with only 5–10 fructose residues.

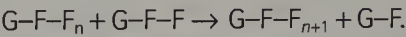
stored in the vacuole. They have a typical structure of glucose–fructose–(fructose) $_n$ (G–F–F $_n$) as shown in Figure 11.3. The use of fructans as a carbohydrate reserve is widespread throughout the plant kingdom. As with glucose polymers, there are different glycosidic linkages possible between the fructose residues, giving different straight and branched polymers. The inulins are the major storage carbohydrate found in bulbs such as onion, and storage roots such as chicory and Jerusalem artichoke, and are formed by (1→2 β) linkages. Levans are widespread in the leaves and stems of grasses, including major cereal crops such as wheat, and comprise (6→2 β) linkages. Graminae-type fructans found, for example, in grasses, are a mixed type and have both (1→2 β) and (6→2 β) linkages.

The biosynthetic pathway of fructans in plants is a two-stage process. The first step involves the transfer of fructose from a donor sucrose molecule to an acceptor sucrose molecule to form kestose by the enzyme sucrose–sucrose fructosyltransferase (SST).



where G = glucose and F = fructose.

In the second step, the kestose (GFF or GF₂) acts as the fructose donor to the growing fructan chain, via fructan–fructan fructosyltransferase (FFT) activity, and a sucrose molecule is recycled.



In certain bacteria, such as *Bacillus subtilis*, very high molecular weight levans are produced by a single reaction in which sucrose acts directly as the fructose donor to the growing chain (sucrose–fructan fructosyltransferase, SFT). Note that in all cases, sucrose is the initial acceptor molecule of the chain, so the first sugar in the fructan chain is always glucose. However, for each remaining fructose residue added to the chain a glucose residue is released, which in plants is transported back out of the vacuole into the cytosol.

A number of transgenic plants producing polyfructans have now been developed. In some of the first experiments with tobacco, two different bacterial genes were compared. The *sacB* gene of *B. subtilis* encodes a levansucrase catalysing a 6→2β linkage, whilst the *ftf* gene from *Streptomyces* spp. encodes a fructosyltransferase that forms 1→2β linkages. The construct used to express *sacB* is shown in Figure 11.4, demonstrating how the *sacB* gene was modified with a vacuolar targeting sequence from the yeast carboxypeptidase gene (*cpy*). The features of vacuolar signal sequences (see also Chapter 1) are described in Box 11.1 The targeting sequence directs the enzyme to the vacuole, where sucrose is stored and polyfructan synthesis normally occurs, in those plants that store this carbohydrate reserve.

In tobacco, both transgenes were responsible for the production of significant amounts of fructans, with the *sacB* gene proving more effective than *ftf*. The transgenic tobacco plants carrying the *sacB* gene were found to be more tolerant to drought stress induced by growth in medium containing polyethylene glycol to reduce the water concentration. In Chapter 9, the production of osmoprotectants to provide stress tolerance against the effects of water deficit was described, and this provides another example of this principle in action. More recently, expression of the *sacB* gene in sugar beet has also been shown to improve tolerance to polyethylene glycol-mediated drought stress.

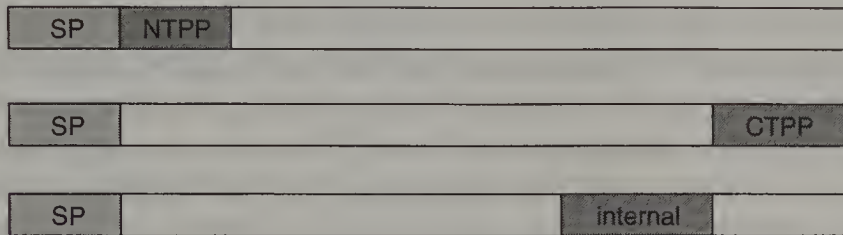
Polyfructans have also been produced in potatoes using the same vacuolar-targeted *sacB* gene, but under the control of the patatin gene promoter (see the earlier example of cyclodextrins). In potato, the formation of a new sink, by diversion of sucrose away from starch accumulation in the tuber and towards fructans in the vacuole, has provided a useful model system for studying the regulation of sucrose metabolism and the partitioning of



Figure 11.4 Construct for the synthesis of fructans in the vacuole of transgenic plants. A construct of the type used to transform tobacco in the first experiments to engineer fructan synthesis is shown. The bacterial *sacB* gene is flanked by an N-terminal vacuolar targeting sequence from yeast carboxypeptidase. Note the tandem-repeated CaMV 35S promoter and AIMV translational enhancer to obtain high levels of expression.

BOX 11.1 Vacuolar targeting sequences

There have been several examples of protein targeting to the chloroplast in previous chapters, but this is probably the first in which the final destination of the transgene product is the vacuole. A number of proteins are known to be targeted to the vacuole, but the targeting sequence appears to vary considerably between proteins. Examples of vacuolar targeting sequences are shown below.



NTPP (N terminal) sequences

Sweet potato sporamin	HSRFNPRLPTTHEPA
Barley aleurain	SSSFADSNPIRPVTDRAASTLE

CTPP (C terminal) sequences

Barley lectin	VFAEAIANSTLVAE
Tobacco chitinase	NGLLVDTM
Tobacco β -1,3-glucanase	VSGGVWDSSVQTNATASLVSQM
Tobacco AP24	QAHPNFPLEMPGSDEVAK

assimilates. In subsequent work, the levansucrase gene from *Erwinia amylovora*, which was fused to different targeting sequences (apoplast, vacuole and cytosol), was used to transform starch-deficient potatoes. The apoplastic- and vacuolar-targeted enzyme produced fructan to levels of 12% and 19% of tuber dry weight, respectively. However, the fructan did not behave as a new sink for increased allocation of carbohydrates in the starch-deficient tubers, and tuber yield was actually further decreased.

In maize, a *B. amyloliquifaciens sacB* gene was expressed in the endosperm under the control of a zein (a maize seed-storage protein) promoter. The enzyme was targeted to the vacuole of the endosperm cells using two different vacuolar targeting sequences. In one construct, the sweet potato sporamin signal peptide and vacuolar targeting sequences were fused to the N-terminal end of the enzyme (see Box 11.1). In the other, the barley lectin signal peptide was fused to the N-terminal end and the barley lectin vacuolar targeting sequence was fused to the C-terminal end of the enzyme. Seed from transgenic plants containing these constructs appeared no different from the controls. Fructan accumulated to 7–9 mg g⁻¹ of seed in several different lines from either construct. This level could be increased by up to eight- to ninefold when the transgene was crossed into maize starch-mutant lines, which accumulate much higher levels of sucrose. On the other hand, expression of a *sacB* construct without vacuolar targeting produced kernels that were clearly distinguishable from the controls in terms of their small size and dry mass, and their darker colour. These factors are all indicative of very low starch accumulation, indicating a

competition for sucrose in the cytoplasm between the fructosyl transferase and the starch biosynthetic pathway.

More recently, short oligofructans (GF₂, GF₃ and GF₄) have been produced in sugar beet using a gene encoding the 1-SST enzyme from Jerusalem artichoke. This enzyme catalyses the production of not only GF₂ (kestose) but also GF₃ and GF₄. It was found that the sugar stored in the storage root of sugar beet was almost totally converted into these oligofructans. This transgenic sugar beet has therefore been renamed the 'fructan beet'. Interestingly, there is no obvious change to the phenotype or growth rate of the fructan beet storage root. The fructan beet has direct applications in the 'nutraceuticals' market. Short-chain fructans are almost as sweet as sucrose, and can be used as substitute sugars in foods and drinks. Fructans are not digested in the gut, and can therefore be marketed as a 'low-calorie sweetener'. The added health benefits are that they encourage the growth of beneficial gut flora (*Lactobacillus* and *Bifidus* species) and cause acidification of the digestive tract, which inhibits bacteria such as *Escherichia coli*. In the future, it is proposed that longer chain fructans produced in sugar beet could be used as base materials for glues, textile coatings and polymers.

Trehalose

Trehalose is produced in some plants and microorganisms (e.g. yeasts), often in response to osmotic stress (see Figure 11.1). It is therefore another potential target for the genetic manipulation of tolerance to abiotic stresses that create water deficit (see Chapter 9). However, it is also a valuable commodity for food processing, dehydration and flavour retention. Genes for trehalose synthesis from yeast and *E. coli* have already been introduced into transgenic tobacco, but the purpose of these experiments was to manipulate drought tolerance (Chapter 9) rather than to manufacture bulk quantities of the sugar.

Metabolic engineering of lipids

The next class of compounds that we will explore as targets for molecular farming is the lipids and their derivatives. As with the carbohydrates, lipids are already produced in large quantities from major crops such as oilseed rape (canola), soybean and maize, for industrial as well as food purposes. Molecular farming will therefore have a role in the same broad areas as shown for carbohydrates, that is:

- improvement of existing lipid products;
- engineering of novel lipid products.

Improvement of plant oils

In daily life, our closest association with plant oils comes by way of their rapid displacement of animal fats as cooking oils. This association alone should indicate that oils are already extracted from a number of different types of crop.

Oilseed rape or canola, and soybean are major sources of oil that are often unspecified, but sunflower oil, corn oil and olive oil are often identified because of the improved characteristics of these oils for particular purposes. Plant oils are also used in a wide range of processed foods, and for animal feeds. They also have an increasing number of applications in industry for the manufacture of soaps and detergents, lubricants and biofuel. However, the non-food applications currently comprise only about 10% of the total vegetable oils produced. One of the reasons for this relatively low level of industrial use is the complex and variable nature of plant oils, which generally makes them less suitable for oleochemical applications than cheaper alternatives from mineral oils. A major long-term goal of this sector is therefore the improvement of plant oils for industrial applications, since they must eventually replace non-renewable, petroleum-based products.

Storage oils that are used as carbon/energy reserves in plant seeds and some fruits (such as olives and avocados) are generally glycerol esters of fatty acids, called triacylglycerols (TAGs) (Figure 11.5). The differences between oils from different sources are largely due to the particular proportions of different fatty acids comprising TAGs (Table 11.2). A few of the common fatty acids are also shown in Figure 11.5. These differ in their length, and the number and positions of double bonds. Saturated fats contain no double bonds—the more double bonds there are, the greater the degree of unsaturation.

The site of *de novo* fatty acid biosynthesis is exclusively in the stroma of the plastid (see Box 5.1), whereas most modification of the fatty acids occurs in the cytoplasm and on the endoplasmic reticulum (ER). Assembly of TAGs takes place in the membrane of the ER. The TAGs are stored in oil bodies, which are essentially oil droplets surrounded by a lipid monolayer formed from one half of the ER membrane lipid bilayer. The oil bodies of seeds (and other tissues that undergo extreme desiccation) also contain proteins, called ‘oleosins’, in the lipid monolayer (oleosins are discussed later in this chapter).

The first committed step in fatty acid biosynthesis in the plastid is the formation of malonyl-coenzyme A (CoA) from acetyl-CoA, in an ATP-dependent reaction catalysed by acetyl-CoA carboxylase (ACCase). The CoA residue is then exchanged for the acyl carrier protein (ACP). Acetyl-CoA is then condensed to the malonyl-ACP, forming acetoacetate-ACP, and liberating CO₂. Three subsequent steps (reduction of the carbonyl group, removal of water to form a double bond and reduction of the double bond) produce an acyl-ACP, which is two carbons longer than the original. The entire sequence of elongation reactions from the initial binding to ACP is catalysed by a fatty acid synthase multienzyme complex. The elongated fatty acid chain is then transferred to another ACP protein and this is then condensed to a new malonyl-ACP. Thus, the formation of fatty acids occurs by the stepwise addition of 2-carbon units at the carboxyl end, hence the even numbers of carbon atoms in the fatty acids shown in Figure 11.5. In many of the oil producing crops, this process stops at the 16-carbon stage, and the palmitoyl-ACP (16:0) is elongated to

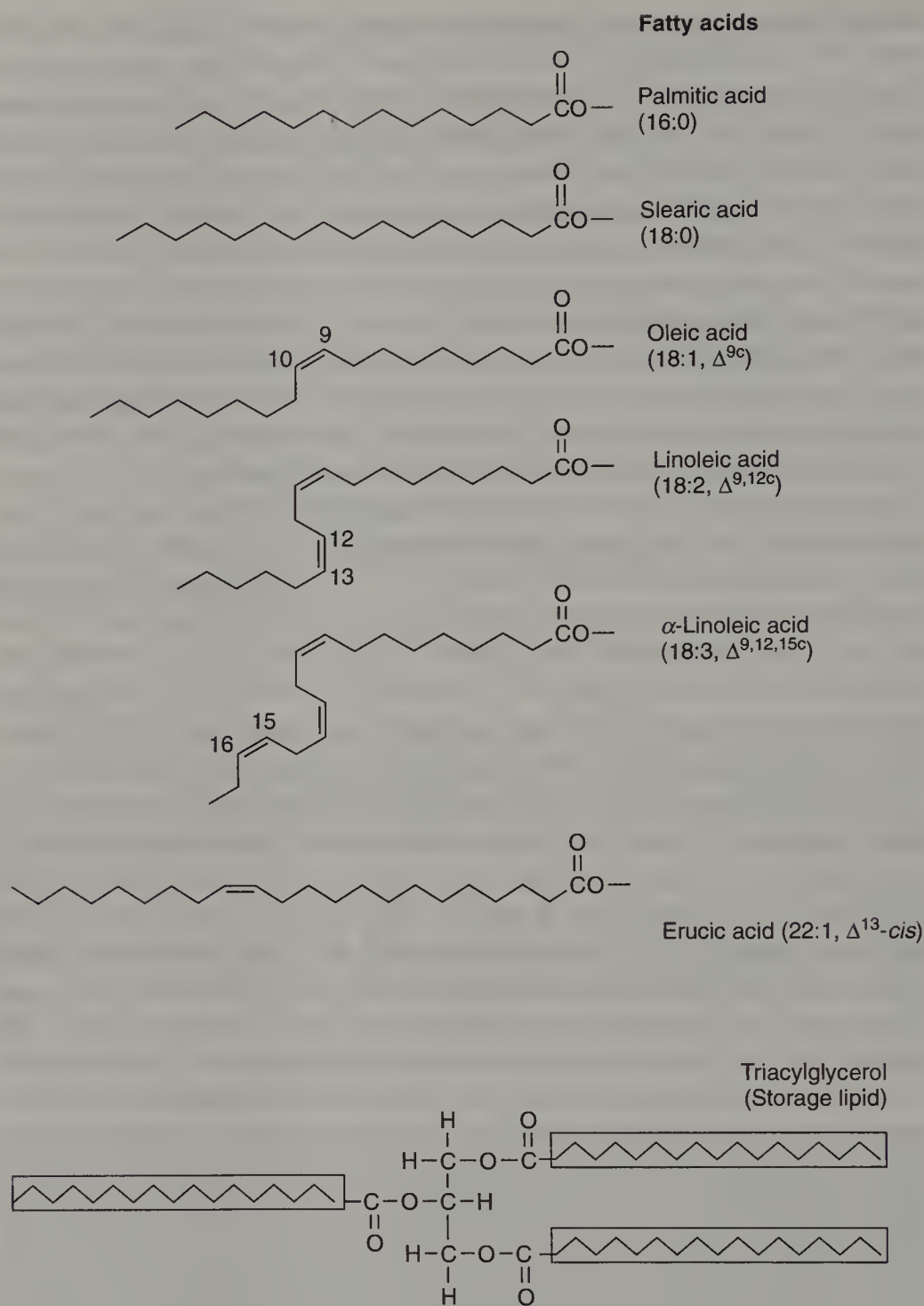


Figure 11.5 Fatty acid and triacylglycerol structure. The structures of some of the common fatty acids found in plant oils are shown. The standard nomenclature in brackets denotes the number of carbon atoms, the number of double bonds and the position of the double bonds, counting from the carboxyl-, or Δ -end. Thus, linoleic acid (18:2 $\Delta^{9,12}$) has 18 carbons and 2 double bonds, one from C9–C10, and the other from C12–C13. The structure of a triacylglycerol storage lipid is also shown, indicating the attachment of three fatty acid chains to a glycerol molecule. Note that apart from the glycerol residue, these molecules are completely hydrophobic—compare this structure with the phospholipids of biological membranes, where one of the fatty acids is replaced by a polar ‘head’ group.

Table 11.2 The fatty acid composition of major oil crops

Fatty acid	Soya oil	Palm oil	Rape-seed oil (LEAR)	Sunflower oil	Peanut oil
16:0	11	42	4	5	10
18:0	3	5	1	1	3
18:1	22	41	60	15	50
18:2	55	10	20	79	30
18:3	8	0	9	0	0
20:1	0	0	2	0	3
Others	1	0	2	0	0

LEAR, low erucic acid rape.

stearoyl-ACP (18:0) by a specific synthase. (The standard nomenclature of fatty acids indicates the number of carbon atoms and the number of double bonds.) At this stage, desaturation by a soluble Δ^9 C-stearoyl desaturase in the plastid stroma converts most of the stearoyl-ACP to oleoyl-ACP (18:1 Δ^9). (The Δ^9 indicates that the double bond starts at the 9th carbon counting from the carboxylic acid 'front end' of the molecule.)

After termination of synthesis in the plastid, the fatty acids (mainly palmitic, stearic and oleic acids) are released from the ACP and exported to the cytoplasm, where they are converted to acyl-CoA esters. TAGs are formed by stepwise acylation of glycerol-3-phosphate in the ER membrane. Further modifications to the fatty acids normally occur after they have been attached to various glycerophosphatides. Additional double bonds may be added (typically to Δ^{12} , Δ^{15} and Δ^6 as well as Δ^9). Other modifications may include hydroxylation by the addition of water across a double bond.

Production of shorter chain fatty acids

The oils produced by the major oil crops of the world consist mainly of palmitic, stearic, oleic, linoleic and linolenic acids, which are all C16 or C18 in length (Table 11.2). Coconut and palm kernel oils are largely C8–C14, and lauric acid (12:0), in particular, is an important raw material for the production of soaps, cosmetics and detergents (think of the widespread use of sodium lauryl sulphate, or SDS, in these products). The synthesis of fatty acids at a particular length is terminated by the hydrolysis of the acyl-ACP by a thioesterase. The California bay tree contains a very high proportion of lauric acid in its seeds, and an acyl-ACP thioesterase that specifically hydrolyses lauroyl-ACP has been cloned from this source. The introduction of this gene into oilseed rape causes fatty acid synthesis to terminate at the 12:0 stage, and a high proportion of lauric acid to accumulate in the seed oil. Most importantly, field tests show that these plants grow normally and produce normal yields.

Production of longer chain fatty acids

One of the important targets is to produce fatty acids longer than C18 for use as industrial oils. In oilseed rape and other *Brassica* species, there is a two-step elongation pathway from oleoyl-CoA (18:1 Δ^9) to erucoyl-CoA (22:1 Δ^{13}), such that erucic acid is one of the constituents of brassica oils. However, whilst erucic acid is valuable as an industrial oleochemical, it is nutritionally unsuitable for human consumption. Conventional breeding has led to the development of two distinct oilseed rape crops—high-erucic acid rape (HEAR) for industrial purposes, and low-erucic acid rape (LEAR) with virtually no erucic acid, for food products (see Table 11.2). (NB: the existence of two distinct crops that need to be kept apart means that systems are in place in the seed production, agriculture and processing industries to ensure minimal cross-contamination, a process called 'identity preservation'. Thus, requirements to keep GM and non-GM seeds and products separate do not create entirely new demands on this sector.) However, the highest erucic acid content of HEAR is about 50% of the total fatty acids, which makes the cost of separating out and disposing of the other fatty acids uncompetitive with mineral oil sources. Attempts are being made to overexpress the genes that encode the elongases, and also to transfer enzyme activities that preferentially incorporate erucic acid into TAGs.

Modification of the degree of saturation

As the erucic acid example shows, the ideal products for industrial purposes are as uniform as possible. This requires uniformity of desaturation as well as length. (The same does not necessarily apply to food constituents, where complex mixtures may produce the ideal product. In this case it is consistency, rather than uniformity, that is required.) One of the first attempts in this area was to increase the level of stearic acid (18:0) at the expense of oleic acid (18:1 Δ^9) by inserting an antisense construct of the Δ^9 desaturase gene into oilseed rape. The antisense gene was under the control of the napin (a seed-specific protein of brassica) gene promoter. There was a marked decrease in the amount of the desaturase enzyme, resulting in a decreased formation of oleic acid, and a rise in the proportion of stearic acid from 1–2% up to 40% of total fatty acids. This high stearic acid oil has potential as a cocoa butter substitute.

On the other hand, there is considerable potential in the production of a very high (>90%) oleic acid oil for food purposes and as a uniform oleochemical feed-stock. Conventional breeding has produced mutant oilseed rape lines with an oleic acid content close to 80%, but attempts to increase this level have produced plants with poor cold tolerance, presumably as a result of the lack of unsaturated fatty acids in the cellular membranes (see Chapter 9). Antisense repression and co-suppression of the Δ^{12} desaturase in the seeds of oilseed rape have, however, facilitated the raising of oleic acid levels to 87–88%, without affecting cold tolerance. Similar experiments in soybean increased the oleic acid level from 22% to 79%.

Production of rare fatty acids

There are 210 known types of fatty acids produced in plants, but most of these are not found in the major crop plants and would be difficult to produce commercially in their host plants. However, it is possible to transfer genes from these plants in order to manipulate the profile of fatty acids produced in the major oil crop plants. One such fatty acid is petroselenic acid ($18:1\Delta^6$), which is found in certain Umbelliferae, such as coriander, and has potential as a raw material for industry. Oxidation of petroselenic acid by ozone produces lauric acid ($12:0$) (for soaps and detergents) and adipic acid ($6:0$), which can be used for nylon production. Transformation of tobacco with a coriander acyl-ACP desaturase cDNA led to the production of petroselenic acid in calli to a level of 5% of total fatty acids. (See Box 11.2 for a discussion of the economics of the process.)

Certain polyunsaturated fatty acids have pharmaceutical or nutraceutical value. These include γ -linolenic acid ($18:3\Delta^{6,9,12}$) and arachidonic acid ($20:4\Delta^{5,8,11,14}$), which are essential fatty acids for humans and precursors of eicosanoids (including prostaglandins, leukotrienes and thromboxanes). Recently, it has become clear that the synthesis of these compounds involves a specific subclass of microsomal desaturases called 'front-end' desaturases. These enzymes insert additional double bonds between existing bonds and the carboxyl 'front-end' of the fatty acid, whereas most desaturases in plants add sequentially towards the methyl end. The transformation of a front-end desaturase from borage (which is one of a few plant species to produce γ -linolenic acid) into tobacco resulted in the accumulation of high levels of Δ^6 unsaturated fatty acids. A Δ^5 desaturase has been cloned from the filamentous fungus *Mortierella alpina* (which accumulates arachidonic acid) and expression in oilseed rape produced significant levels of polyunsaturated fatty acids.

Ricolenic acid (Δ^{12} -hydroxyoleic acid) is produced in castor beans to a level of 90% of the total fatty acids. However, the castor oil crop has a number of problems, including the presence of toxic compounds such as ricin in the residual meal. The synthesis of ricolenic acid involves the direct hydroxylation of oleic acid bound to phosphatidylcholine on the ER membrane. The cDNA for the 12-hydroxylase has been cloned from castor bean and could be used to produce 'castor oil' in major oil crops. Another fatty acid modifying enzyme has been cloned from *Crepis acetylenics*. This non-haem, di-iron protein catalyses triple bond and epoxy group formation in fatty acids, and would be a valuable way of inserting chemically reactive sites into oils.

CASE STUDY 3 **Bioplastics**

One of the most imaginative examples of molecular farming has been the attempt to produce biodegradable plastics ('bioplastics') in plants. These compounds are currently produced by microbial fermentation, but a number of experimental studies have been carried out to determine the feasibility of producing them in bulk in plants. The structure of the polyhydroxyalkanoates (PHAs) which form bioplastics is shown in Figure 11.6. The length of

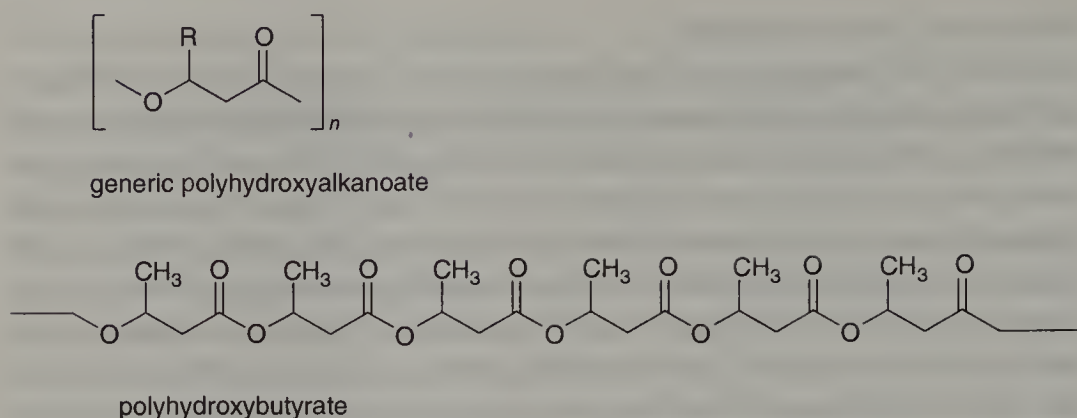


Figure 11.6 Chemical structure of the major polyhydroxyalkanoates. The repeating subunit of polyhydroxyalkanoates is shown, along with a short section of a polyhydroxybutyrate chain (where R = methyl group). The R group can be anything from 0 carbons (i.e. H) in hydroxypropionate, to >10. The length of the R-group side chain affects the physical properties of the plastic formed by these polymers.

the R side chain alters the properties of the plastics, and can vary from 0 carbon (3-hydroxypropionate), 1 carbon (3-hydroxybutyrate) or 2 carbons (3-hydroxyvalerate) up to long carbon chains. Polyhydroxybutyrate (PHB) is the best-characterised PHA and is found as intracellular inclusions in a wide variety of bacteria. In *Alcaligenes eutrophus*, PHB accumulates as a high molecular weight polymer up to 80% of the bacterial dry weight.

The pathway for polyhydroxybutyrate synthesis is shown in Figure 11.7. It can be seen that this is a relatively simple three-stage pathway starting from acetyl-CoA. The genes for the three enzymes involved in the pathway (successively the *phaA*, *phaB* and *phaC* genes) have been cloned from *Alcaligenes eutrophus*. In the initial experiments, it was recognised that the first step of the pathway, production of acetoacetyl-CoA, occurs in the cytoplasm of plants, at the start of the pathway to isoprenoids (see Chapter 10). Thus, only *phaB* and *phaC*, encoding acetocetyl-CoA reductase and PHB synthase, respectively, were transformed into *Arabidopsis*, without targeting sequences (Figure 11.8(A)). Microscopic observation of the *Arabidopsis* leaves indicated the formation of microbodies of bioplastics accumulating in the cytosol, nucleus and vacuole. However, the amount of bioplastic was relatively low (20–100 µg g⁻¹ fresh weight) and the plants were severely stunted in growth.

Subsequently, all three genes were transformed into *Arabidopsis*, and targeted to the chloroplast (Figure 11.8(B)). In the first generation of experiments, each gene was separately fused to a sequence encoding the transit peptide plus N-terminal fragment of the Rubisco small subunit protein, and expression of each construct was directed by the CaMV 35S promoter. Each construct was transformed into separate *Arabidopsis* plants, and brought together by a series of sexual crosses between the individual transformants. In this case, the bioplastics accumulated as 0.2–0.7-µm granules in the plastids, to levels of up to 14% of plant dry weight, and there was no observable effect on growth or fertility.

Increased PHB production in *Arabidopsis* has been achieved using a triple construct, so that all three genes are transferred into the plant in a single transformation event. A rapid gas chromatography—mass spectrometry (GC–MS) procedure was used to screen a large number of transgenic plants, permitting the selection of plants accumulating PHB in the leaf

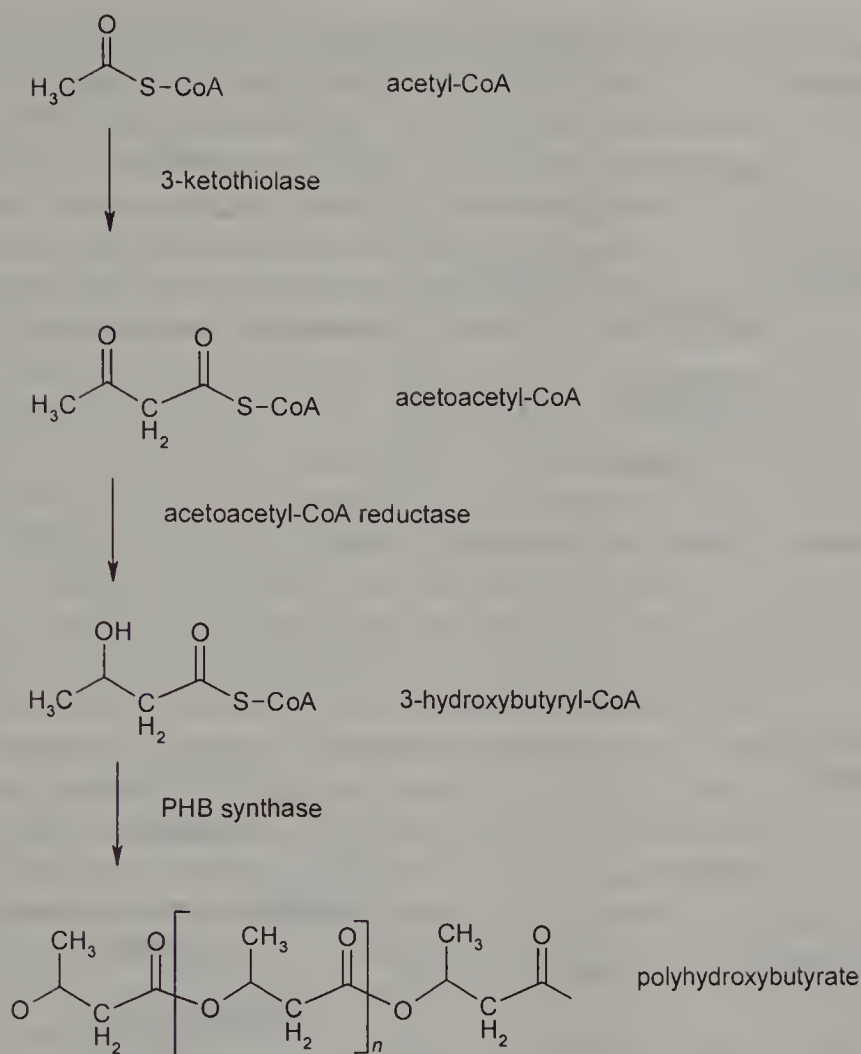
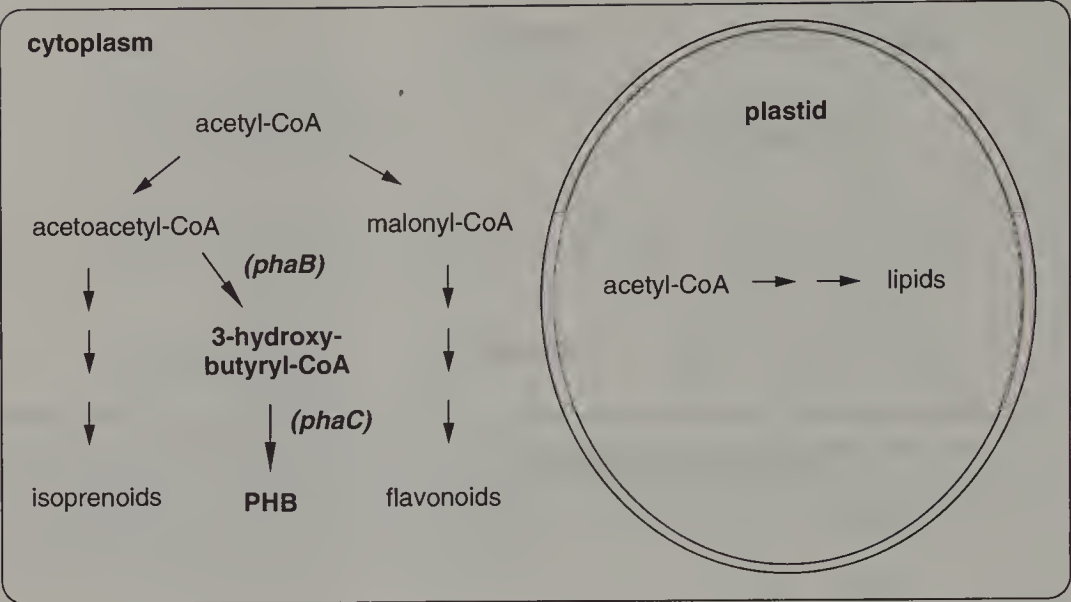


Figure 11.7 The PHB biosynthetic pathway of bacteria. The biosynthesis of polyhydroxybutyrate requires only three steps from acetyl-CoA. In the first step, two acetyl-CoA residues are condensed by 3-ketothiolase to form acetoacetyl-CoA. This is reduced by acetoacetyl-CoA reductase to give 3-hydroxybutyryl-CoA. In the final stage, this compound is polymerised by PHB synthase.

chloroplasts to more than 4% of their fresh weight (40% dry weight). However, the high-producing lines showed stunted growth and loss of fertility. Interestingly, the production of PHB did not affect fatty acid accumulation or composition, but there was a significant impact on the levels of various organic acids, amino acids, sugars and sugar alcohols.

Whilst *Arabidopsis* is a valuable model species for these analyses, some research has been focused on the potential for the commercial production of plastics in oil crops, effectively diverting the pool of acetyl-CoA away from fatty acid biosynthesis (for oil body production—see above) towards bioplastic production. A team from Monsanto used the three genes for the PHB biosynthetic pathway from the bacterium *Ralstonia eutropha* and fused each one to a seed-specific promoter (from *Lesquerella fendleri* oleate 12-hydroxylase). These were transferred into a single, multigene vector, which was used to transform oilseed rape. PHB was found to accumulate in mature oilseed leucoplasts to levels up to 7.7% of fresh seed weight. However, apart from in *Arabidopsis*, it has proved difficult to obtain very

A



B

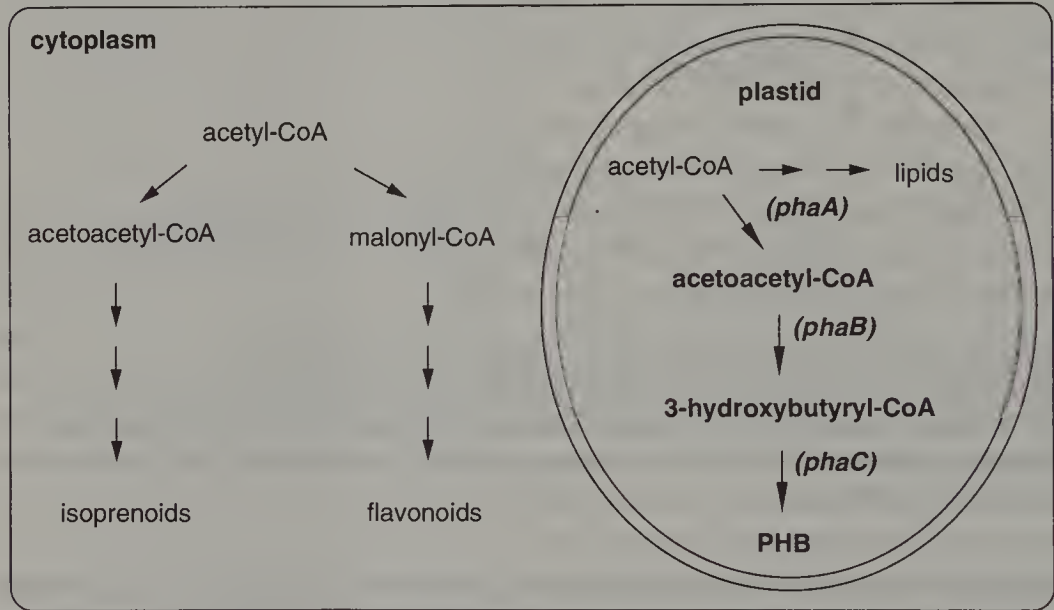


Figure 11.8 Expression of the PHB biosynthesis pathway in transgenic *Arabidopsis*. In scheme A, the presence of 3-ketothiolase activity in the plant cytoplasm removes the necessity to transfer three genes to the plant. The *phaB* and *phaC* genes from *Alcaligenes eutrophus* were transformed into *Arabidopsis* without protein targeting sequences. PHB granules were produced in the cytoplasm, but they also accumulated in the nucleus and vacuole. In scheme B, all three genes of the pathway are targeted to the chloroplast using Rubisco small-subunit transit peptide sequences. In this case, PHB accumulates in the plastids. (Redrawn with permission from Poirier, Y., et al. (1995).)

high levels of PHB in transgenic plants. Recent work indicates that constitutive expression of the β -ketothiolase gene is detrimental to the efficient production of PHB in some plant species, and the use of inducible or developmentally regulated promoters to drive the *phaB* gene permitted some production of PHBs in tobacco and potato.

Another interesting direction taken by this research has been the production of bioplastics in cotton fibres. At the start of this work it was recognised that cotton fibres contain β -ketothiolase activity (the first enzyme in the PHB pathway, see Figure 11.7). Therefore, the *A. eutrophus phaB* and *phaC* genes were transformed into cotton by particle bombardment (see Chapter 3) of seed axis meristems. The *phaB* gene was driven by a cotton fibre-specific gene promoter, whilst *phaC* was fused to the 35S promoter. Clusters of small PHB granules were found in the cytoplasm of the fibre cells (i.e. in the fibre lumen). The thermal properties of these fibres were found to be altered, indicative of enhanced insulation characteristics.

One of the limitations of this approach to bioplastic production is that polyhydroxybutyrate is a highly crystalline polymer, which produces rather stiff and brittle plastics. Polyhydroxyalkanoate co-polymers made from longer monomers have better physical properties in terms of being less crystalline and more flexible. Pseudomonads accumulate medium chain-length PHAs when grown on fatty acid substrates. These PHAs are synthesised from 3-hydroxy-acyl-CoA intermediates generated by the β -oxidation of the fatty acids, hence the PHA monomer size is related to the length of the fatty acid substrate. Medium chain-length PHAs have been produced in *Arabidopsis* using the *phaC1* gene from *Pseudomonas aeruginosa* with a peroxisome targeting sequence from an oilseed rape isocitrate lyase. The enzyme was targeted to leaf-type peroxisomes in light-grown plants and glyoxisomes in dark-grown plants. These plants accumulated PHAs in the glyoxisomes and peroxisomes, and in the vacuole, to a level of 4 mg g⁻¹ dry weight. The PHAs contained saturated and unsaturated 3-hydroxyalkanoic acids ranging from 6 to 16 carbons, with 41% being a mixture of the 8-carbon saturated and mono-unsaturated monomer.

Molecular farming of proteins

Throughout this book we have presented to you examples where transgene expression is used as a tool to, for example, modify a biosynthetic pathway, manipulate growth and development or alter some other property of the plant such as herbicide resistance. However, in some cases it is the transgenic protein itself that is important. In Chapter 10 we saw how attempts to improve the amino acid content of grain are being addressed by expression of proteins with a high content of lysine. In this section of the chapter, we will consider the production of bulk enzymes and potentially high-value proteins such as antibodies and vaccines. The production of functional proteins in plants on an industrial scale does not, at first sight, suggest too many difficulties. However, the production of a complex protein that is folded, processed, easily purified and can be shown to be safe for pharmaceutical use, all at low cost, is a real challenge. The economics of protein production in plants is complicated. The

actual cost will depend on many factors, amongst them being the cost of growing the plant, transport costs and processing and protein purification costs. The costs of proteins produced in plants may undercut the costs of producing proteins by existing methods, but if too much (in economic terms) is produced, then the market price will fall. This has already proved to be the case for some vaccines produced by existing methods.

In this remaining section, the protein products that will be considered fall into two main types. First, enzymes for industrial and agricultural uses will be looked at. Second, we will consider medically related proteins such as:

- antibodies—full immunoglobulins and engineered types such as scFvs (single chain antibodies);
- subunit vaccines—vaccines based upon short peptide sequences that act as antigens rather than whole organism-based vaccines; and
- protein antibiotics.

Production systems

The interest in producing such proteins in plants comes in part from the problems associated with existing fermentation or bioreactor systems. Mammalian cell systems are expensive and cannot easily be scaled up (input costs can approach \$US1 million per kg of crude product). Bacterial systems can be scaled up, but often the recombinant proteins are not properly processed (they are not properly folded and disulphide bridges are not formed), so intracellular precipitation of non-functional proteins can occur. Plant systems can be scaled up so that large amounts of material can be harvested and processed, allowing industrial-scale amounts of protein to be purified. In some cases it may be possible to omit purification, as is the case with vaccines from plant sources, and where plant material containing recombinant enzymes can be added directly to animal feed or industrial processes. There are some differences between plant and animal post-translational modification systems, and the difference in glycosylation patterns between plants and animals has caused some concerns, particularly when expressing antibodies in plants. However, plants are able to fold, cross-link, and post-translationally modify (by, for example, glycosylation) non-plant proteins sufficiently well to ensure that, in most cases, functional proteins are obtained. The production of recombinant proteins in plants also benefits from the ability to target the proteins to cellular compartments where processing occurs, or where the proteins may be more stable. Another advantage of using plant systems is that the potential for contamination of the protein products with toxins or pathogens of animals and humans is very much reduced.

The plants initially chosen for expression studies related to the ability of scientists to transform the plant material. A lot of work was first done with tobacco, but this was not a suitable plant for feeding to animals. For work in which edible products are being developed, plants such as potatoes, tomato,

Table 11.3 Potential biomass yields from various plants considered as potential vehicles for protein production

Crop	Potential annual biomass yield (tonne per hectare)
Tobacco	>100
Alfalfa	25
Maize	12
Rice	6
Wheat	~3

maize and lettuce have been used. For bulk production, the leaves of plants such as tobacco and alfalfa remain a favourite, because they can be harvested several times a year. However, plants such as rice, wheat, maize and soybean have also been used, although their biomass yields are lower (Table 11.3).

Strategies for protein production

Two major strategies have been developed for the production of proteins in plants: the stable integration approach, which has been used for many systems; and the use of plant viruses as transient vectors.

Stable expression

The stable transgene expression approach, in which the transgene is regulated by a strong, constitutive promoter (such as the 35S promoter), is perhaps the most suitable for the bulk production of soluble proteins in leaves, although yields can be low using this approach. A more sophisticated approach has been to target gene expression and protein production to specific tissues (see Figure 11.9). Restricting gene expression to particular organs or tissues often leads to higher yields of recombinant proteins. Targeting expression has a number of advantages, some of which are related to the increased yield. The resources of the plant are effectively wasted if proteins are produced in parts of the plant that are not harvested. So, for vaccines that are to be ingested, after the minimum amount of processing, it makes sense to target the protein only to the parts of the plant that are normally eaten. Another important point is that storage organs, such as tubers and seeds, are designed to maintain their biological integrity over long periods. Targeting proteins to these structures has been shown to increase the stability of recombinant proteins from weeks to years. Seeds are also of great interest because they can offer simple strategies for protein purification. In tobacco, seeds contain a simpler mixture of proteins and lipids, with fewer phenolic compounds, than do green leaves.

As well as targeting expression to particular organs or tissues, intracellular targeting has been used to optimise the production of active proteins. Post-translational modification has been shown to be associated with the ER

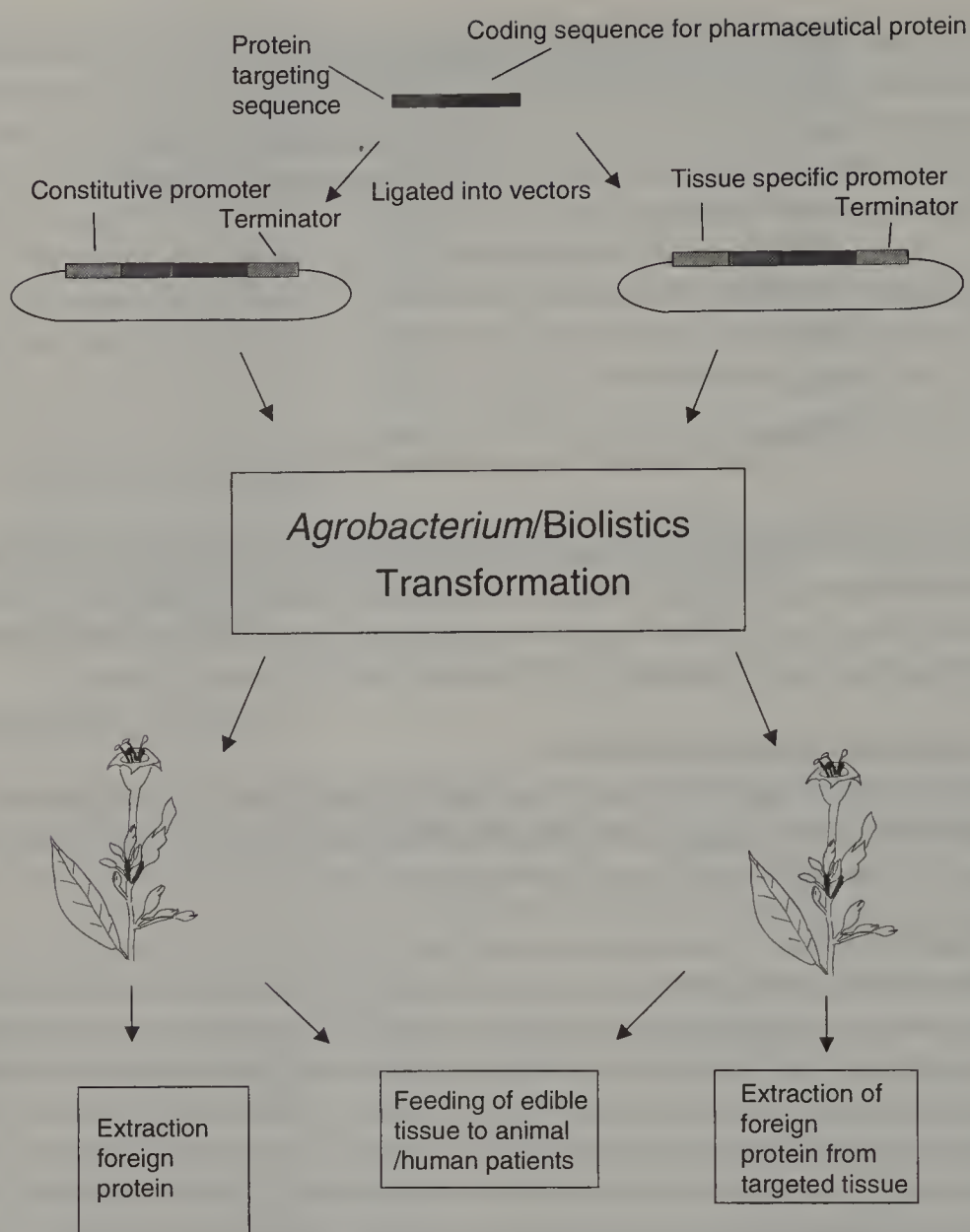


Figure 11.9 Different routes of protein manufacture in stably transformed plants. Constitutive (from, for instance, the 35S promoter) or targeted (to particular organs or tissues) expression and accumulation results in recombinant proteins that can either be used after purification or directly as plant products.

and/or apoplastic export, and the inclusion of transit peptides has been shown to improve both yield and activity. Tissue and intracellular targeting are brought together in the use of oleosins as vehicles for protein purification.

CASE STUDY 4 Hirudin production in *Brassica napus*

In this case study we will look at a purification system that was developed to improve the efficiency and economic viability of pharmaceutical proteins produced in plants. We will look at the application of the technology to the production of hirudin, an anticoagulant peptide.

Although pharmaceutical protein production is discussed in more detail later in this chapter, the case study here illustrates how novel purification procedures may help to make pharmaceutical proteins produced in plants a widely adopted method of production.

Background

Oil bodies, the subcellular organelles which store triacylglycerols (TAGS) have already been met in this chapter (see the section on improvement of plant oils). They are especially abundant in oilseeds, like the seeds of canola, where they are the primary source of energy for seedling germination and the early stages of growth. Oil bodies from seeds are composed of TAGS surrounded by a phospholipid monolayer. Proteins called oleosins are associated with the surface of the oil bodies. Studies had shown that if the *gusA* reporter gene (see Chapter 4) was fused to the C-terminus of either a native (or C-terminal deletion mutant) oleosin, then GUS activity partitioned with the oil body fraction, indicating that it was targeted to the oil body by the oleosin. It was found that the N-terminal and central domains of the oleosin were required for correct targeting of GUS to the oil body.

Purifying oil bodies

Oil bodies can be purified extremely simply by a series of centrifugations. Seeds are ground in an aqueous buffer and centrifuged. After centrifugation, three fractions are evident. A pellet of insoluble material is formed at the bottom of the tube. An aqueous phase, consisting of the soluble components of the seed is present; and, floating on top of this, is a layer composed almost exclusively of oil bodies with their associated oleosins. This oil body layer can be removed, resuspended in buffer and re-centrifuged to increase the purity of the oil bodies. This process is illustrated in Figure 11.10 below.

Hirudin

Hirudin is a small (~7-kDa) anticoagulant peptide found in the salivary glands of the medicinal leech *Hirudo medicinalis*. Hirudin is the most potent inhibitor of thrombin (a serine protease that catalyses the initial steps in blood clotting) so far identified, and has high therapeutic value in the treatment of thrombosis. Thrombosis (in heart attacks and strokes) is the biggest killer in the western world. Hirudin has several advantages over other, more commonly used, anticoagulants such as heparin. It has a low immunogenicity, does not require endogenous co-factors for activity and does not interact with other proteins in the blood. Its use is limited, however, due to difficulties in obtaining large quantities. Recombinant hirudin has been produced in both *E. coli* and yeast. Production in *E. coli* resulted in low yields of the protein and reduced activity. Some yeast production systems have proved to quite efficient, producing large amounts of functional protein.

Production and purification in plants

A synthetic gene, with codon usage optimised for expression in *Brassica napus*, encoding hirudin was fused to the 3' end of an *Arabidopsis* oleosin gene with its own promoter. An endoprotease Factor Xa cleavage site was engineered between the C-terminus end of the oleosin and the N-terminus end of the hirudin. This construct was used to transform *Brassica napus*.

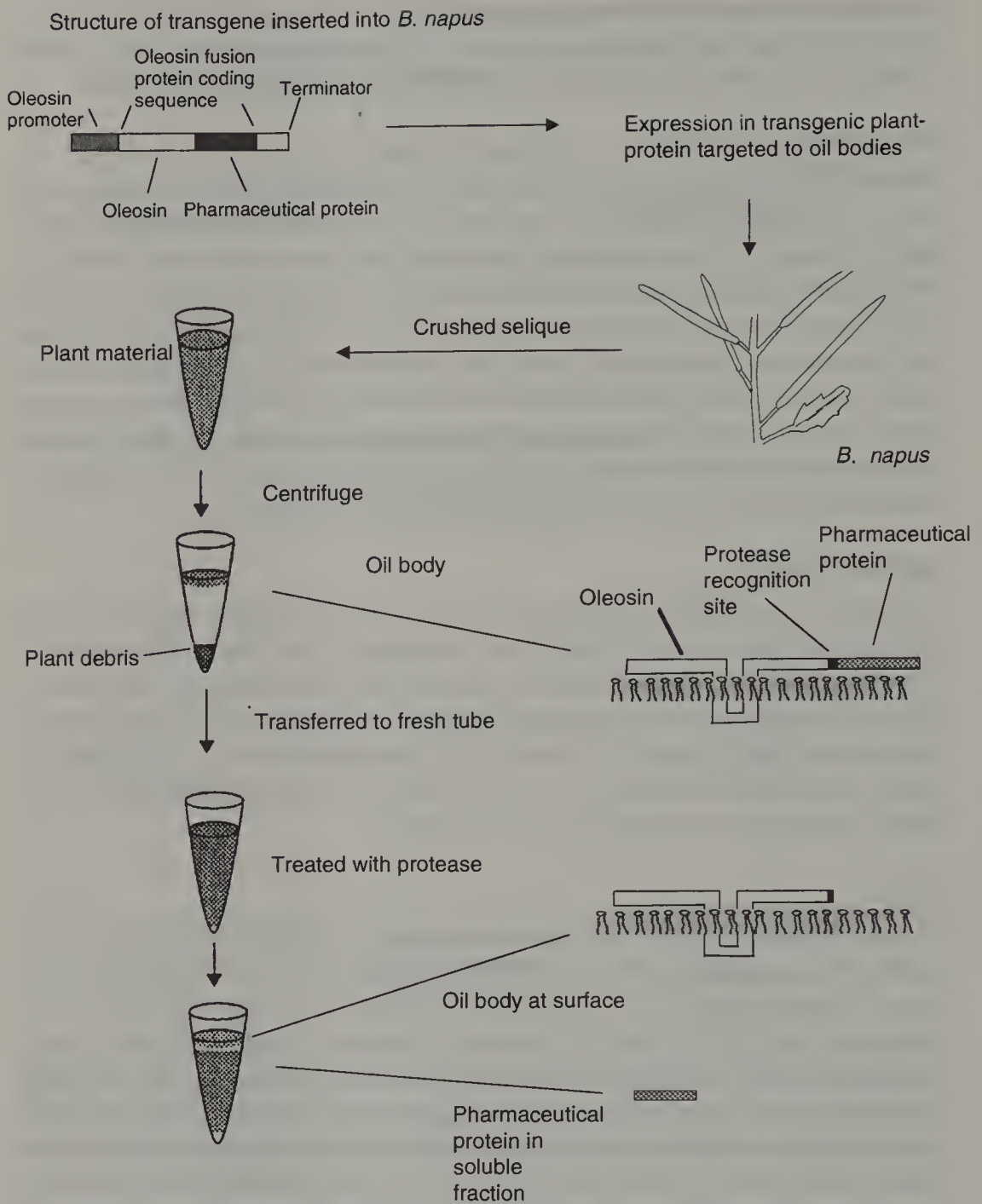


Figure 11.10 A scheme for pharmaceutical protein purification by production in oil bodies. The pharmaceutical protein is produced as a C-terminus fusion to oleosin (an oil body associated protein). The pharmaceutical protein and the oleosin are separated by a protease cleavage site. Seeds are crushed in an aqueous buffer and centrifuged, resulting in three fractions; the oil bodies, with the oleosins and the pharmaceutical protein, are the topmost layer (which floats on the aqueous layer). This oil body layer can be removed, re-centrifuged and then treated with a protease, which cleaves the pharmaceutical protein from the oleosin. Re-centrifugation results in highly purified protein being present in the aqueous phase. Full details can be found in the text. Redrawn from Goddijn OJM and Pen J. 1995 TIBTECH 13 379–387.

Hirudin was purified using a modification of the method for oil body purification outlined earlier in this case study. After the oil bodies have been centrifuged, they are treated with the endoprotease Factor Xa, which cleaves the hirudin from the oleosin, as shown in Figure 11.10. The oil bodies can then be re-centrifuged, after which the hirudin will be present, highly purified, in the aqueous phase (Figure 11.10).

Prospects

There appears to be a good prospect that purification systems based on oil body/oleosin technology could make pharmaceutical protein production in plants an economic reality, especially if cheaper alternatives to Factor Xa, for example, are found. A company called SemBioSys has been established in Canada that hopes to successfully market pharmaceutical proteins, like hirudin, produced using this technology.

One drawback of the technology is that it cannot be used to produce proteins that require extensive processing and modification in the endoplasmic reticulum though, which unfortunately applies to many clinically important proteins.

Transient expression

Plant virus capsids have also been used as carriers of recombinant proteins, particularly vaccines. In one approach, coding sequences for epitopes or proteins have been introduced into the coat protein gene of the virus genome so that it is 'presented' externally when virus particles are made. This approach is illustrated in Figure 11.11. Another approach that has been used is to construct viral vectors to produce recombinant proteins that are targeted to the ER for processing. The virus can be replicated in the host plant, and through serial passage enough protein can be generated for clinical use.

The production of enzymes for industrial uses

The global industrial enzyme market has been estimated to be worth something approaching \$US2 billion per annum. Proof of concept incursions by plant biotechnology into this market were on a small scale. The first commercialised 'industrial proteins' produced from transgenic plants (Table 11.4 summarises recombinant enzyme production in plants) were two proteins familiar to the molecular biologist—avidin (from chicken) and β -glucuronidase (from *E. coli*)—both of which were produced in maize. Following on from these successes, ProdiGene Inc., the company responsible for their development, is leading the race to large-scale production of plant-derived transgenic enzymes for commercial use. They announced in Spring 2002 that they would be going into large-scale production of trypsin, a proteolytic enzyme, which by its very nature is difficult to produce in conventional recombinant systems. This protein is currently harvested from bovine and porcine pancreases, and is used for a number of applications, including the production of pharmaceuticals such as insulin, vaccine production and wound care. The company has predicted that the worldwide demand for the enzyme will increase fivefold in

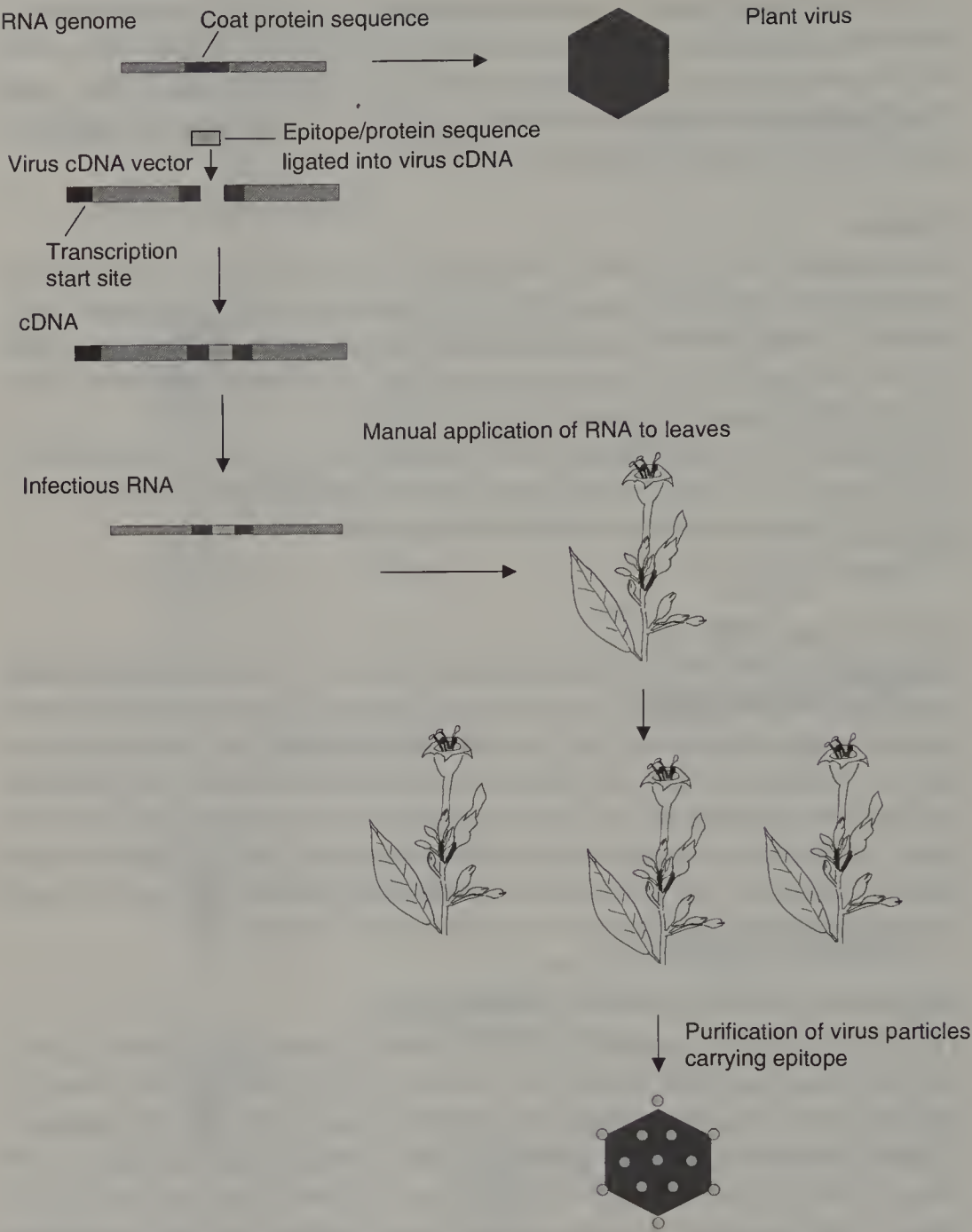


Figure 11.11 Construction of a recombinant plant virus carrying a protein epitope. The gene for the epitope sequence was introduced into the coat protein sequence within the vector. Infectious RNA was made and then introduced into plants (transfection) where virus capsids, displaying the expressed protein, were produced. This is known as the ‘overcoat system’ and can be used for a range of proteins.

Table 11.4 Examples of enzymes produced in plants

Enzyme	Use
Avidin	Diagnostic kits
β -Glucuronidase	Diagnostic kits
Trypsin	Pharmaceuticals, wound care
Cellulase	Ethanol production from cellulose waste
Thermostable xylanase	Biomass processing
Phytase	Phytate breakdown, improved phosphate utilisation
α -Amylase	Food processing
(1–3) (1–4) β -Glucanase	Brewing
Lignin peroxidase	Paper manufacture

(Data taken with permission from Daniell, H., *et al.* (2001); Giddings, G., *et al.* (2000); Kusnadi, A. R., *et al.* (1997) and Hemming, D. (1995).)

the next 5 years. Existing sources may be unable to supply this increase, and there is the added worry of the spread of pathogens (like BSE) with products derived from animals. ProdiGene Inc. aims to meet this demand by producing kilogram quantities of the protein from plant sources.

Some of the other important enzymes that may be produced in plants are, interestingly, normally targeted against components of the plant cell. Cellulases and xylanases, which are normally produced by the microorganisms associated with digestion in ruminants, have been produced in a number of different plants. There is a large market for these enzymes in the bioethanol, textile, pulp and paper industries and for the production of animal feed. It may, at first sight, seem a strange concept to express such enzymes in plants, due to the risk of ‘autodigestion’, but engineered, thermostable, forms of the enzymes, with high temperature optima were used. Therefore, no deleterious effects were detected at the temperatures at which a plant normally grows, but high levels of enzyme activity were detected when plant extracts were heated to the temperature optima of the enzyme. These were only small-scale initial trials, but the process could easily be scaled up. An interesting point to note is that these enzymes would require only the minimum of purification (an important consideration when determining applicability and economic viability). This is also the case for another important hydrolytic enzyme, phytase. This enzyme catalyses the hydrolysis of inositol hexaphosphate (phytate) to its constituents, the sugar-alcohol inositol and inorganic phosphate (see Figure 11.12). Phytase is a very useful enzyme as it releases phosphate from a substrate that is normally indigestible to monogastric animals. As phytate is found in significant quantities in seeds used to formulate feed meal for pigs and poultry, these animals do not benefit from this source of phosphate, which

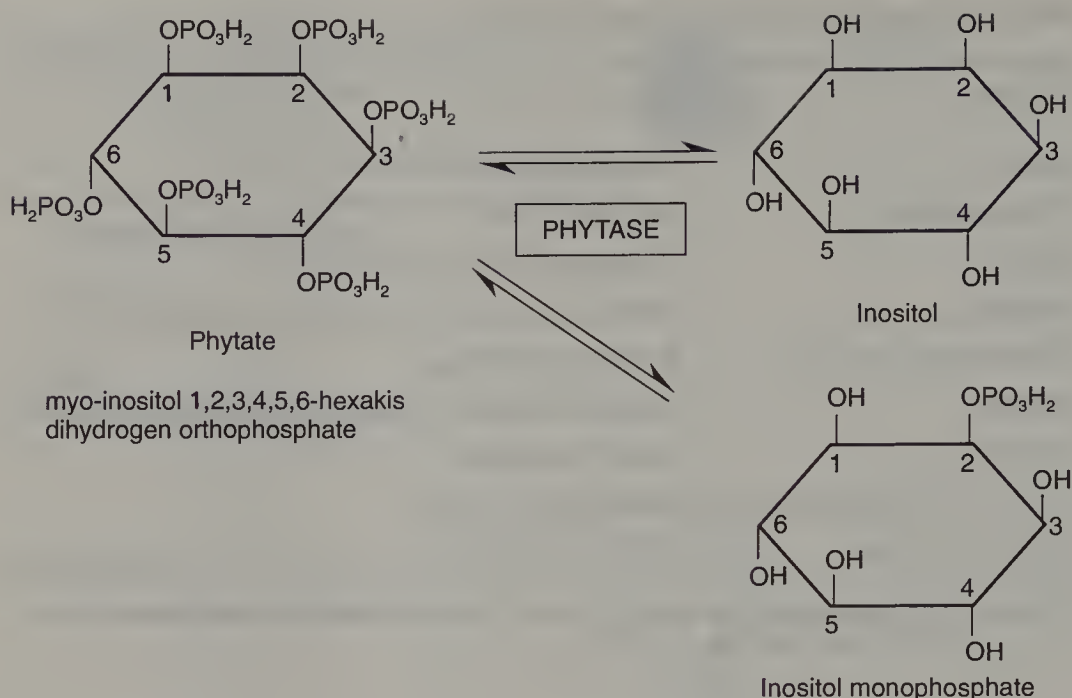


Figure 11.12 The action of phytase. Phytate is broken down either to inositol monophosphate (which is believed to be the main product of phytate breakdown) or to inositol.

is required for proper skeletal growth. Problems with phytate are exacerbated by the fact that monogastric animals excrete the indigestible phytate in their faeces, which contributes to an excessive phosphate build-up in ground water, leading to eutrophication. Transgenic plants expressing phytase in, for example, seeds have been produced. When these transgenic seeds are added to the diet of pigs and poultry, phosphate utilisation is increased (negating the need for supplements) and phosphate excretion is decreased. Thus, the addition of the enzyme in a simple plant-derived formulation solved both a nutritional and an environmental problem.

Medically related proteins

This is an important, expanding, area for plant biotechnology, in which there is great hope that recombinant proteins made *in planta* will be accepted and widely utilised. This section of the chapter is divided into three sections:

- (1) antibodies;
- (2) vaccines; and
- (3) other medically related proteins.

Antibodies or 'plantibodies'

In mammals there are a number of different types of antibodies, or immunoglobulins (Ig). The immunoglobulins have been divided into five major

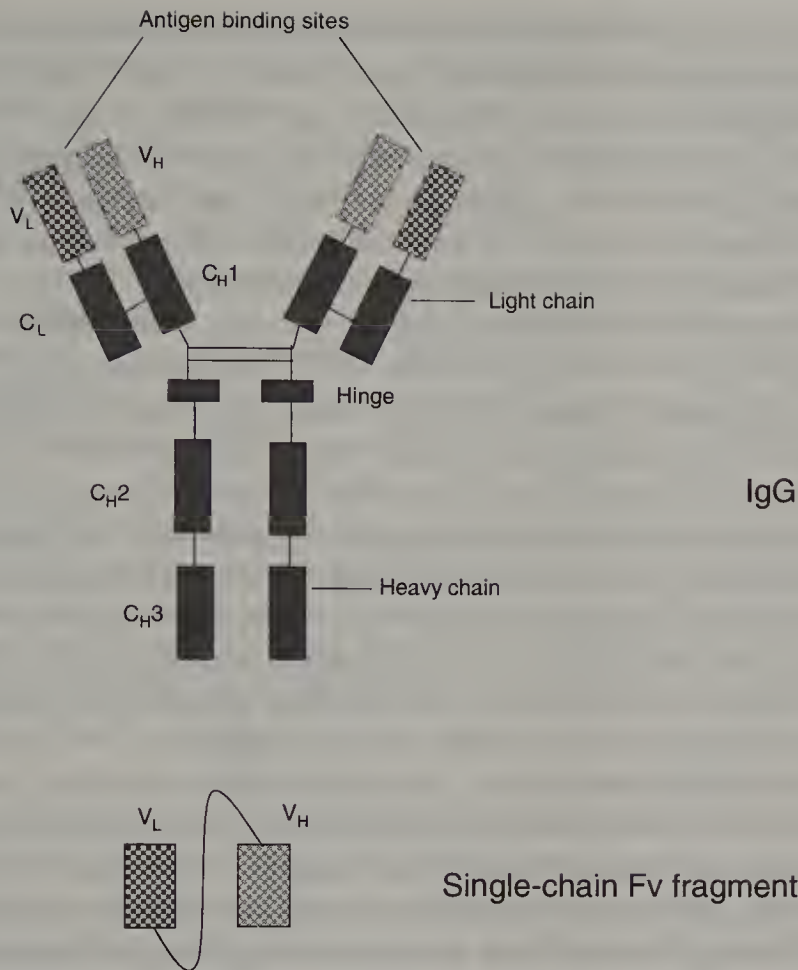


Figure 11.13 Structures of IgG and an engineered single-chain variable fragment antibody. A typical antibody is composed of two identical heavy (H) chains and two identical light (L) chains. Antigen-binding sites are found at the amino-terminal regions of both the H and L chains. The hinge and tail regions are formed only from the H chains. The light and heavy chains are composed of constant (C) and variable (V) domains. The variable domains of the heavy (V_H) and light (V_L) chains make up the antigen-binding sites. Single-chain variable fragment (scFv) antibodies are engineered by joining together light and heavy chain variable regions. Antigen-binding is determined solely by these variable regions.

classes on the basis of their physical, chemical, and immunological properties: IgA, IgD, IgE, IgG and IgM. Figure 11.13 shows the structure of IgG (IgGs are discussed later in this chapter) which can be used as an example. Immunoglobulins are composed of two heavy chains and two light chains. Each class of Ig has its own class of H (heavy) chain, termed α , δ , ϵ , γ and μ , respectively (IgA molecules have two α -class heavy chains, for example). Each Ig also has two light (L) chains, which are either κ or λ , but not a mix of the two. IgA, the main antibody found in secretions, can have a more complicated structure. In secretions, IgA exists as a dimer of two IgA molecules joined by a J (or joining) chain and associated with a secretory component. Each antibody has two antigen-binding sites, formed from the variable (V) regions of the light and heavy chains. Also shown on the figure is the structure of a

single chain antibody (a single-chain variable fragment antibody or scFv). As the variable regions of the light and heavy chains are solely responsible for antigen binding, simpler, smaller molecules that still bind antigens can be produced. Single-chain variable fragment antibodies are produced from synthetic genes made by fusing the sequences for light- and heavy-chain variable regions. Although not the only example of engineered antibodies, scFv antibodies are the most commonly used and most successful.

The first example of a functional antibody (a mouse immunoglobulin IgG1) being produced in plants was reported in 1989. The construction of this antibody was a two-step process. In the first stage, two separate transgenic tobacco lines were produced, one contained the gene for the γ -heavy chain the other contained the gene for the κ -light chain. The second stage of the process was to cross these plants to generate progeny that expressed both genes. It was found that in the F1 progeny expressing both chains, significant amounts (up to 1.3% of total leaf protein) of assembled antibody were produced. These transgenes contained native immunoglobulin signal sequences, which targeted the nascent light and heavy chains to the ER and were found to be necessary for efficient assembly of the antibody in plants.

Several important observations, which underpinned much of the future work in this area, were made in this original research. Targeting of the antibody chains to the lumen of the ER was found to be necessary for efficient chain assembly and stability. Conditions in the lumen of the ER favour the correct processing of the antibody chains, especially disulphide bridge formation, and molecular chaperones also help to direct antibody chain folding and assembly. Another important observation was that the formation of the complete antibody in the F1 plants increased the protein yield over that of either of the parents that contained single chains, reflecting the stable nature of the complete antibody in the plant system.

Although a pioneering study, demonstrating that functional antibodies could be assembled in plants, the antibody in question had no practical applications.

Antibodies produced in plants are thought to be particularly suitable for topical immunotherapy. In an extension of the strategy of producing the various antibody subunits in different plant lines that are subsequently crossed to produce a functional antibody, a secretory IgA (sIgA) that protects against dental caries produced by *Streptococcus mutans* has been expressed in plants. The antibody does this by recognising the native streptococcal antigen (SA) I/II cell-surface adhesion molecule, thereby preventing colonisation.

This multiple-component sIgA antibody required four parental transgenic tobacco lines to be produced, in which cDNAs for either the H, L or J chains (the H and L chains were derived from a murine antibody termed 'Guy's 13' that recognises the SA I/II cell-surface adhesion molecule), or the secretory component, were expressed under control of the 35S promoter, and then crossed. The strategy used is illustrated in Figure 11.14. The success of the approach has been quite an impressive technical achievement. In mammalian

Table 11.5 Examples of antibodies that have been produced in plants

Application	Plant	Antibody	Signal sequence
Immunoglobulins			
<i>S. mutans</i> SA I/II (dental caries)	Tobacco	sIgA (hybrid)	Murine IgG
<i>S. mutans</i> SA/II (dental carries)	Tobacco	IgG (guy's 13)	Murine IgG
Surface antigen (colon cancer)	Tobacco	IgG Co17–1A	Murine IgG/KDEL ^a
Herpes simplex virus	Soybean	IgG (anti HSV-2)	Tobacco extensin
Single-chain Fv			
Lymphoma	Tobacco	scFv (38C13)	Rice α -amylase
Carcinoembryonic antigen (cancer)	Cereals	scFvT84.66	Murine IgG/KDEL ^a

(Data taken with permission from Daniell, H., *et al.* (2001); Giddings, G., *et al.* (2000); Kusnadi, A. R., *et al.* (1997) and Hemming, D. (1995).)
^a ER targeting sequence (see Chapter 1).

systems, both plasma cells and epithelial cells are involved in the synthesis and assembly of secretory antibodies, yet single plant cells containing all four transgenes are able to produce them efficiently. Secretory antibodies have many theoretical advantages in situations like the one described here. Secretory IgAs are the predominant antibody type that protects against microbial infections at mucosal sites. They are also more resistant to proteolysis (which contributes to high yield), and they bind antigens with a higher efficiency than do the original antibodies produced in plants.

These secretory antibodies have now been tested on humans. They have been topically applied to teeth and found to be effective at preventing re-colonisation by *S. mutans* for up to 4 months, which is as effective as an IgG produced in a murine hybridoma. Despite the differences in structure, there appear to be no differences in binding properties between the proteins.

Plants have been used to produce a variety of antibodies, including whole antibodies, antigen-binding fragments and single-chain variable fragment antibodies. Despite the now well-characterised differences in the glycosylation pattern seen between antibodies produced in plant and mammalian expression systems, antibodies produced in plants generally seem to exhibit similar properties to antibodies produced in other, mammalian, systems. Some of the most important examples are listed in Table 11.5.

CASE STUDY 5 Custom-made antibodies

The strategy used to generate most of the antibodies listed in Table 11.5 is to generate stably transformed plants that express the protein. This has not been the approach used by Large Scale Biology Corp. for the development of an idiotypic vaccine. A tobacco mosaic

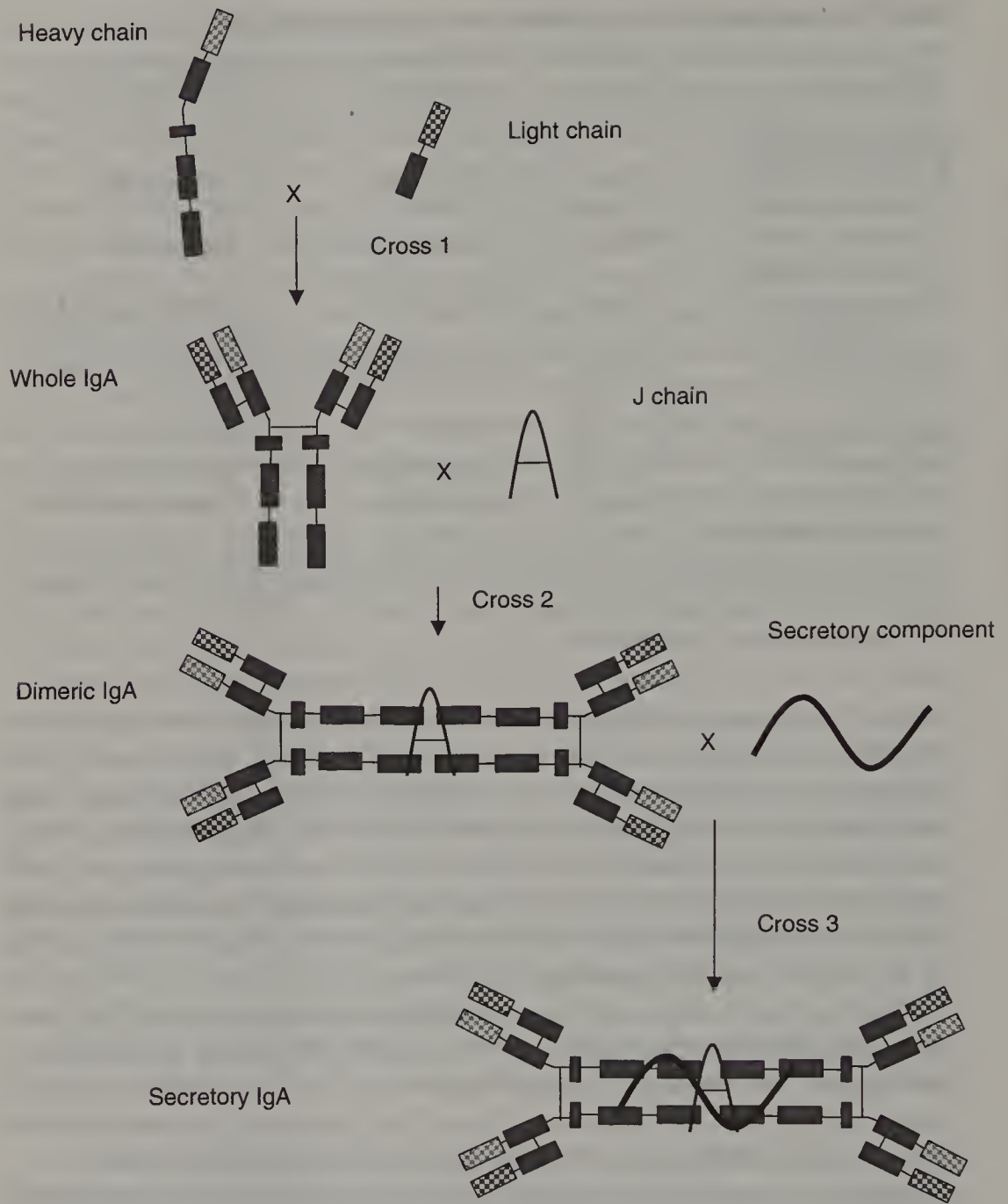


Figure 11.14 Synthesis of a secretory IgA (sIgA) molecule in transgenic plants. Four separate parental plant lines were constructed. Each expressed a separate piece of the antibody (the H, L or J chain, or the secretory component). The lines expressing the H and L chains were crossed, giving F1 plants that accumulated IgA. These plants were further crossed with plants expressing the J (joining) chain, which, in progeny, resulted in the accumulation of dimeric IgA. These plants were finally crossed with plants expressing the secretory component. Progeny of this cross accumulated functional sIgA.

tobamovirus (TMV) -based vector has been developed for the production of a secreted scFv protein during virus infection of non-transgenic tobacco plants. In this vector, a virus promoter regulates the expression of the scFv gene and the antibody coding sequence is inserted downstream from a rice α -amylase leader sequence, which targets the protein to the extracellular compartment of the plant (see Figure 11.15). The advantage of this transient expression system is that it is very rapid and can be used to produce small batches of customised antibody. The company has targeted its use towards the treatment of B-cell lymphoma—in this disease, these cells produce a particular antibody that is unique to each patient. This idiotype vaccine system is one where the patient is treated with a copy of the lymphoma antibody to generate immunity against the original antibody-producing lymphoma cells. This works because each individual antibody has a variable region that has a unique antigenic constitution, the idiotype. Thus, the idiotype vaccine will only raise a specific response against the lymphoma cells and not against any of the patient's other cells. It is now possible, using the TMV vector system, to generate the idiotype region of a tumour-specific immunoglobulin from a patient, as a scFv, in a matter of 6 to 8 weeks. The system was developed using mice as a model, but it has now entered clinical trials with 16 patients. This phase 1 testing is the first and smallest required by the US Food and Drug Administration (FDA) for any drug. It is interesting that each of the antibodies in the clinical trials is unique, and yet can be considered as a single drug. One important advantage of this system is that it is possible to harvest enough material for clinical use from greenhouse-grown plants, thus ensuring high levels of containment.

Plant-derived vaccines

Both stable and transient expression systems have been used to produce vaccines in plants. One of the first successful attempts in this area used a transient expression system. Recombinant cowpea mosaic *comovirus* (CPMV) was used as a vector for a linear epitope (antigenic determinant) from the VP2 capsid protein of mink enteritis virus (MEV). The recombinant CPMV was replicated in black-eyed beans (*Vigna unguiculata*), from which chimeric virus particles (CPMV-VP2), displaying the VP2 epitope, were isolated and used for the subcutaneous injection of mink. The vaccine was able to induce resistance to clinical infection by MEV. Many other vaccines have now been produced in plants and Table 11.6 lists many of these.

The main aim of many of these studies is the generation of edible vaccines. The driving force behind the use of edible vaccines has been the need for cheap, effective treatments for enteric disease in the developing world. Plant-based systems have the potential to provide low-cost, easily administered, locally produced edible vaccines that could provide mucosal immunity against the infectious agents that are responsible for the deaths of millions of people, particularly children, annually. To bring this about it is necessary to present, in a practical, edible form, a quantity of vaccine that will stimulate the GALT (gut-associated lymphoid tissue) system in the gut to generate secretory antibodies. The questions that have had to be considered relate to delivery of the

Modified TTO1A TMV chimeric vector containing scFv sequence

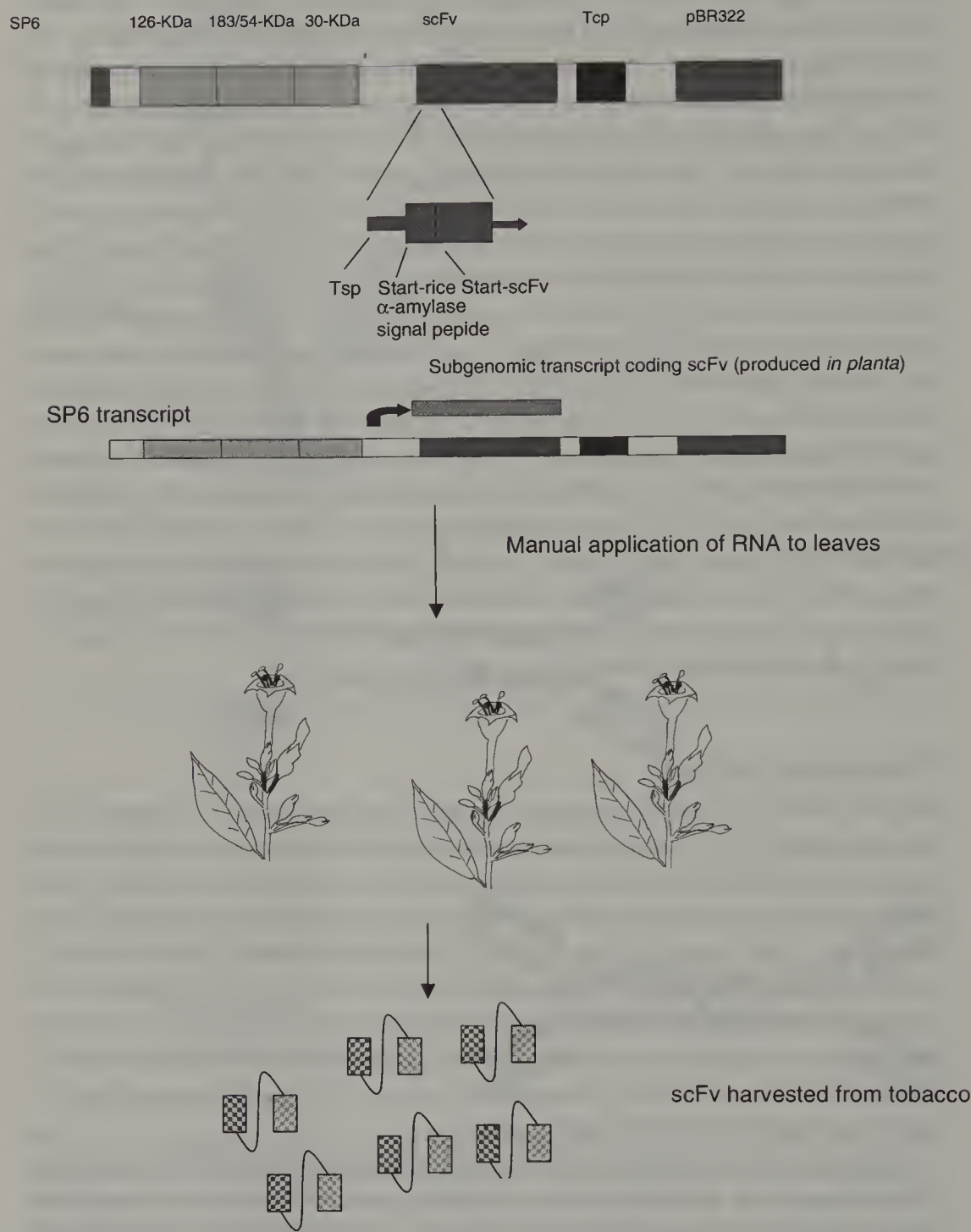


Figure 11.15 The use of engineered tobacco mosaic virus (TMV) vectors for the expression of pharmaceutical proteins. An scFv antibody sequence has been located adjacent to a translation start sequence and a rice α -amylase signal peptide. Infectious RNA is produced from an SP6 promoter and used to infect plants (transfection). Subgenomic RNA (see Chapter 8) is produced, and this acts as a template for expression of the scFv. scFv proteins can be isolated. (Tsp, translation start position; Tcp, tomato coat protein; pBR322, bacterial replication sequence.)

Table 11.6 Examples of vaccines produced in plants

Origin	Recombinant protein	Plant	Production level ^a
Human			
<i>Escherichia coli</i>	Heat-labile enterotoxin B	Tobacco	0.001% SLP
<i>Vibrio cholerae</i>	Cholera CtoxA and Ctox B subunits	Potato	0.3% TSP
Hepatitis B	Envelope surface protein	Tobacco/potato	<0.1% FW
Norwalk virus	Capsid protein	Tobacco/potato	0.23%/0.37% TSP
Rabies virus	Rabies virus glycoprotein	Tomato	1% TSP
Animal			
Foot and mouth virus	Virus epitope VP1	Alfalfa/ <i>Arabidopsis</i>	N/A
Porcine coronavirus	Viral glycoprotein	Tobacco/maize	0.2% TSP/0.01% FW
Mink enteritis virus	Viral epitope VP2	Blackeyed bean	N/A (CPMV) ^c
Canine parvovirus	Peptide from VP2 capsid protein	<i>Arabidopsis</i>	3% SLP

(Data taken with permission from Daniell, H., *et al.* (2001); Giddings, G., *et al.* (2000); Kusnadi, A. R., *et al.* (1997) and Hemming, D. (1995).)

^aSLP, soluble leaf protein; TSP, total soluble protein; FW, fresh weight; N/A, not available.

^cProduced using CPMV expression System.

vaccine to the gut through the stomach. As proteins, these vaccines are likely to be degraded in the stomach. The experiments carried out so far indicate that orally administered plant material expressing a vaccine can induce an immune response, with particulate material the most effective. Interestingly, when virus coat proteins are expressed in plant leaves they can form virus particles that may also improve the stability and effectiveness of the vaccine by presenting the protein in its native form.

Much of the initial work on producing vaccines in plants was carried out in inedible plants, such as tobacco. There has been a move by some scientists towards producing vaccines in plants that could be eaten raw, such as tomato and banana. Bananas are considered to be the ideal production system by some scientists as they are grown widely in the countries of the developing world and can be eaten raw. For use in animals, it is possible to consider the use of fodder crops or other food crops.

These so-called ‘edible vaccines’ do work in the laboratory, but would they really work as originally envisaged? The major flaw in the process is that one needs to control very carefully the dose being given to patients if effective and safe protection against disease is to be ensured. One of the drawbacks of

edible vaccines is that it may prove very difficult to control the dose administered if consumption as part of a foodstuff is the means of delivery. Charles Arntzen, one of the main lights in this field, has recently withdrawn from the concept of edible vaccines. He believes they should be considered vaccines that come from a plant source. His group produces vaccines in tomato, but they are partially processed so that food-based tablets containing a known dose of vaccine can be produced. Not only is this better science but it will be easier to license the use of these vaccines from plant sources. Bananas may, of course, still prove to be a useful system that allows locally based production facilities to be established in developing countries.

Other proteins—biopharmaceuticals

A wide range of proteins of pharmaceutical interest has been expressed in transgenic plants in the hope of producing an economically viable system for large-scale production (Table 11.7). Although active recombinant proteins have been produced, one problem associated with production in plant systems is that these often give a relatively low product yield and recovery. Various strategies to overcome this problem are being developed, amongst them the use of novel purification systems (see the section on oleosins for an example) and chloroplast transformation (which is discussed in Chapters 3 and 12).

Despite some difficulties, plants hold out great promise as production systems for pharmaceutical proteins, and we will look at some examples in a little more detail.

Trichosanthin

Trichosanthin (TCS) is a component of the tuber of the plant *Trichosanthes kirilowii*, which is used in Chinese medicine. A ribosome-inactivating protein, TCS inhibits tumour growth and the immune response, and may be useful as a treatment for HIV/AIDS. Relatively high levels of TCS accumulation (2% of total soluble protein 2 weeks after inoculation) have been achieved in *Nicotiana benthamiana* using a viral RNA-based transfection system. This system offers the possibility of rapid, large-scale production of TCS. The rapid production also offers the opportunity to rapidly produce mutated forms of TCS for screening in trials aimed at identifying forms with increased efficacy or reduced side-effects.

Glucocerebrosidase

Gaucher's disease is an inherited disorder in which glucocerebroside (a component of red blood cells that is normally broken down into glucose and ceramide as old and damaged red blood cells are removed by the body) accumulates in lysosomes, due to a deficiency in the enzyme glucocerebrosidase. Gaucher's disease, which results in swelling of the spleen and liver and severe bone damage, can be extremely debilitating and painful. Currently, treatment is based on managing the symptoms or by treatment with a drug developed from glucocerebrosidase purified from human placentas. However,

Table 11.7 Examples of biopharmaceuticals produced in plants

Recombinant protein	Origin	Plant	Application
Protein C	Human	Tobacco	Anticoagulant
Hirudin	<i>Hirudo medicinalis</i>	Canola	Anticoagulant
Somatotrophin	Human	Tobacco	Growth hormone
β -Interferon	Human	Rice/turnip/tobacco	Treatment for hepatitis B + C
Serum albumin	Human	Tobacco	Burns/fluid replacement, etc.
Haemoglobin- α and - β	Human	Tobacco	Blood substitute
Homotrimeric collagen	Human	Tobacco	Collagen
α_1 -Antitrypsin	Human	Rice	Cystic fibrosis, haemorrhages
Aprotinin (trypsin inhibitor)	Human	Maize	Transplant surgery
Lactoferrin	Human	Potato	Antimicrobial
ACE	Human	Tobacco/tomato	Hypertension
Enkephalin	Human	<i>Arabidopsis</i> /canola	Opiate
Trichosanthin- α	<i>Trichosanthes kirilowii</i>	Tobacco	HIV therapy, cancer

(Data taken with permission from Daniell, H., *et al.* (2001); Giddings, G., *et al.* (2000); Kusnadi, A. R., *et al.* (1997) and Hemming, D. (1995).)

ACE, angiotensin-converting enzyme; HIV, human immunodeficiency virus.

a huge amount is required (10–12 tons of placentas a year for one patient), making this one of the world's most expensive drugs. Recombinant approaches to production in mammalian cell cultures have reduced the cost to some extent, although it remains an extremely expensive drug.

A process to produce glucocerebrosidase in tobacco has recently been patented, which will hopefully result in cheaper glucocerebrosidase being available for patients with Gaucher's disease.

Human serum albumin

Human serum albumin (HSA) has many potential medical applications (the treatment of burns and in liver cirrhosis, for example). HSA has been expressed in tobacco and potato under the control of the 35S promoter. Two forms of the HSA protein were expressed, with different signal sequences to ensure secretion of the HSA. One form had the human prepro-sequence, whilst the other had the signal sequence from the tobacco extracellular PR-S

protein. Although both forms of the HSA were secreted, analysis showed that the form with the human prepro-sequence was not properly processed, leading to proHSA being secreted. However, the form with the PR-S signal sequence was correctly processed, and mature HAS, that was indistinguishable from the human protein, was produced.

Arginine vasopressin

The examples we have looked at demonstrate that both plant and non-plant pharmaceutical proteins, which are of high value both clinically and economically, can be successfully produced in plants. However, in what is a typical situation in developing technologies, not all attempts at protein production in plants have been successful.

Attempts to express arginine vasopressin (AVP) in tobacco cultures, for example, have proved unsuccessful due to protein instability and retarded cell growth.

Economic considerations for molecular farming

Whilst there has been much talk in this chapter of suitable targets for molecular farming, few of the examples described have yet been produced commercially. This may result from problems with the technology itself (see the section on cyclodextrins), or it could reflect the long time between initial laboratory-based demonstration and subsequent optimisation and regulatory approvals for field release and approval for food or drug use. However, it may be the case that levels of production that represent major scientific breakthroughs are none the less insufficient to be economically viable. This may seem surprising, because it is tempting to make the naïve assumption that the production of compounds in a field, with free energy from the sun, must be cheaper than production in a chemical factory or industrial fermenter (Box 11.2 gives an interesting example of the economics involved in petroselenic acid production in transgenic soybean). This is often not the case, because farming is not a cost-free activity, the extraction of the product from the plants may be a difficult and expensive process and the price of the product must be significantly higher than that of the crops it has replaced. This may mean that the current levels of synthesis in transgenic plants coupled to the costs of production make a number of the examples described above uneconomic when compared to current production methods. On the other hand, continuing improvements in plant biotechnology, coupled to the inevitable rise in costs of non-renewable resources, such as petrochemicals, make it likely that many of these products will eventually be made in plants.

The economics of producing medically related proteins in plants has been much discussed, particularly because of the high unit price of many of these proteins and the expansion of the market for medically related proteins. It is

BOX 11.2 The cost of producing petroselenic acid in transgenic soybeans

A number of factors need to be taken into consideration when determining the economics of molecular farming. These can be exemplified by a calculation of the cost of producing a petroselenic acid-rich oil in transgenic soybean. A major factor that makes soybean an attractive crop for molecular farming is the current high volume and low cost of production. Thus, in the USA, about 6.5×10^{10} kg of soybeans are produced annually, at a cost of about \$US0.24 kg⁻¹. The low cost of soybean production permits soybean oil to be priced in the range of \$0.41–0.53 kg⁻¹, which compares favourably with many petrochemical products. There are therefore considerable benefits in using a high-volume/low-cost crop like soybean as a platform for molecular farming. However, using a highly developed crop to produce a speciality product also incurs costs.

One factor to be taken into account is that the productivity of major crops like soybean is maintained by a continuous breeding programme that constantly improves traits like yield and pest resistance, and develops different cultivars for different environments. It is not then sufficient to introduce a transgenic trait, even in an elite background, without considering the effects of this constant improvement. The transgenic crop must either accept an ever-decreasing relative yield, or try to keep pace with the ongoing breeding programmes. Traits like yield are generally polygenic, so they need to be selected empirically from large breeding populations. Since each unlinked locus that is required to be present in the crop increases the size of the breeding population by a factor of two, the introduction of three genes (to produce bioplastics, for example) would need an eightfold increase in the size of the breeding programme. Even if the transgenic crop keeps pace with the rate of improvement, the initial introduction of the transgene into a particular elite line is estimated to result in a yield penalty of 3% by the time that outdated cultivar reaches the field. This yield penalty will need to be paid to the crop producer to compensate for the loss of yield from growing normal elite soybean. At a 3% reduction in yield, the farmer would need to be paid an extra 3% of \$0.24 kg⁻¹. The speciality crop therefore costs an extra \$0.007 kg⁻¹ to produce.

Another cost associated with the production of a transgenic compound in a major commodity crop is the cost of what is called 'identity preservation', that is keeping the speciality crop separate from the major bulk crop. This is estimated to be about 6% of the price of the bulk crop (for simple identity preservation), so for soybean this would be about \$0.015 kg⁻¹. The total cost is therefore $\$0.24 + 0.007 + 0.015 = \0.262 kg⁻¹.

Since the seed oil content of soybean is 20%, and since only the oil in this speciality crop is of increased value, these increased direct costs will be added only to the oil. Thus, with an added production cost of \$0.022 kg⁻¹ for the unprocessed bean, the oil will cost an added $5 \times \$0.022 = \0.11 kg⁻¹. For example, if the normal cost of soybean oil was \$0.58 kg⁻¹, the speciality oil would cost \$0.69 kg⁻¹ to produce.

The next consideration is the cost of extraction and processing. Petroselenic acid (18:1 Δ^6) can be cleaved by ozonolysis to form adipic acid (which is a potential precursor to the nylon 6,6 monomers) and lauric acid. Assuming that petroselenic acid could be engineered to the same level as high oleic acid soybeans (85%), it is estimated that 1 kg of oil would produce 0.4 kg of adipic acid, 0.55 kg

BOX 11.2 Continued

of lauric acid, 0.1 kg glycerol and 0.26 kg other fatty acids. Given a market value of \$1.48 kg⁻¹ each for adipic acid and lauric acid, \$1.01 kg⁻¹ for glycerol, and a minimal value for the other fatty acids (\$0.07 kg⁻¹), the total return is \$1.52 kg⁻¹ oil (\$0.59 for adipic acid + \$0.81 for lauric acid + \$0.10 for glycerol + \$0.02 for the other fatty acids).

However, the cost of similar ozonolysis processes is calculated to be \$1.29 kg⁻¹. This means that with a starting material cost of \$0.69 kg⁻¹, the total production costs would be \$1.98 kg⁻¹ compared to a total return of \$1.52 kg⁻¹. Thus, production of petroselenic acid in soybeans, even at levels of 85% of total fatty acids, would not be profitable based on these figures. (Calculations from Hitz, B. (1999).)

in this area that many companies and people have invested money, in the hope of achieving economically viable molecular farming in plants. Fortunately for these companies and investors, the figures seem to suggest that there is a future for molecular farming. It has been estimated that the cost of producing immunoglobulins in alfalfa, in a 250-m² greenhouse, are \$US500–600 g⁻¹ compared with the figure of \$US5000 g⁻¹ for a hybridoma-produced antibody. The company Planet Biotechnology has compared the costs of producing an IgA antibody in four different systems. The comparison was between production in a cell culture system, transgenic goats, grain (assuming a yield of 7.5 tonnes ha⁻¹) and green biomass, such as leaves (assuming a yield of 120 tonnes ha⁻¹). The cost of an IgA produced in leaves was estimated to be below \$US50 g⁻¹. This compared very favourably with the costs for production in a cell culture system (\$US1000 g⁻¹) and in a transgenic animal (\$US100 g⁻¹). It must be remembered that these figures are fairly crude, being based on projected yields, etc. However, they do illustrate that costs of protein production in plants are at least comparable with other systems of production and further developments are now coming on stream. As we have seen in chapter 3 transformation systems are now available for the introduction of foreign genes into chloroplast, giving an estimated 10,000 copies per cell. Using such systems it is claimed that recombinant protein content has reached as much as 47% of the total soluble protein. An example given in Table 11.7 Somatotrophin (-growth hormone), produced in chloroplasts, reached about 7% of the TSP. This compares favourable with that of a nuclear transgene (0.01% TSP). Using Chloroplast based plant systems have other potential advantages over existing systems. Despite the very high number of gene copies there have been no reports of gene silencing that could effect expression levels. Also, protein processing is one of the most expensive parts of producing proteins in existing systems such as *E. coli*. It seems that chloroplasts can assemble and fold complex foreign proteins, but whether they can cope with the other post transla-

tional modification problems remain to be seen. Another advantage of chloroplast is that they are maternally inherited. This means that they will not be spread via pollen to non-transgenic systems. This type of clean gene technology is most important for getting approval for growth in non-contained facilities. Other Health related proteins produced in chloroplasts include: human proinsulin for the treatment of autoimmune disease, Interferon IFN α 2b, human serum albumen, synthetic human haemoglobin genes, Guys 13 Monoclonal antibody, and a *Bacillus anthracis* protective antigen.

Summary

This chapter has illustrated the wide range of compounds that have already been produced in transgenic plants and those that might be valuable targets for plant biotechnology in the future. The chapter has concentrated on three major classes of biological macromolecule—carbohydrates, lipids and proteins. It should be borne in mind that there is an extremely broad range of secondary metabolites that could also be produced in transgenic plants, but these are somewhat beyond the scope of this book.

In the case of carbohydrates, two types of product were identified. On the one hand, existing bulk carbohydrate products such as starch may be modified to enhance their utility. On the other hand, there is the potential to produce novel, high-value carbohydrate products in high-yielding crops. The most successful of these to date has been the development of the 'fructan beet', but many of the other attempts to produce novel carbohydrates have met with less success. Nevertheless, many of these studies have been useful in terms of the growing body of knowledge concerning the regulation of carbohydrate metabolism and the partitioning of carbon between different compartments and tissues in the plant.

The precise manipulation of lipid metabolism requires a detailed understanding of the role of different enzymes in producing the profile of different fatty acids in different plants. There is a considerable amount of fundamental research still required in this area before 'designer oils' with a uniform composition become commonplace. On the other hand, the progress towards the production of bioplastics in plants has been impressive, partly because this can be achieved by importing well-characterised pathways from bacteria that do not exist in the plant, rather than having to 'fine-tune' complex plant pathways. Nevertheless, the import of this pathway into plants does have a significant impact on the host plant metabolism, and it is still proving difficult to produce commercially significant amounts of this product in crop plants.

A wide variety of proteins have been expressed in plants, although the majority of these have not been produced on a commercial scale, and therefore many questions about the viability of using plants for producing proteins remain unanswered. Improvements in the technology of protein production in plants will undoubtedly come and, like lipid metabolism, fundamental research into protein processing, post-translational modification and stability will make a contribution. These developments should see the costs of protein production in plants being reduced, although development and regulatory approval costs will still be significant. It has been estimated that a plant-derived pharmaceutical will cost \$US50 million merely to develop.

Although the future of protein production in plants may be uncertain, the production of recombinant proteins in plants provides some of the best examples of where plant biotechnology can make a positive impact on human (and animal) health and quality of life.

Further reading

Plant biochemistry

- Buchanan, B. B., Gruissem, W. and Jones, R. L. (2000). *Biochemistry and molecular biology of plants*. American Society of Plant Physiologists, Rockville, Maryland, USA.
- Heldt, H.-W. (1997). *Plant biochemistry and molecular biology*. Oxford University Press.

Molecular farming

- Goddijn, O. J. M. and Pen, J. (1995). Plants as bioreactors. *Trends in Biotechnology*, 13, 379–87.
- Hitz, B. (1999). Economic aspects of transgenic crops which produce novel products. *Current Opinion in Plant Biology*, 2, 135–8.

Carbohydrates

- Caimi, P. G., McCole, L. M., Klein, T. M. and Kerr, P. S. (1996). Fructan accumulation and sucrose metabolism in transgenic maize endosperm expressing a *Bacillus amyloliquefaciens* *SacB* gene. *Plant Physiology*, 110, 355–63.
- Ebskamp, M. J. M., Van der Meer, I. M., Spronk, B. A., Weisbeek, P. J. and Smeekens S.C.M. (1994). Accumulation of fructose polymers in transgenic tobacco. *Biotechnology*, 12, 272–5.
- Oakes, J. V., Shewmaker, C. K. and Stalker, D. M. (1991). Production of cyclodextrins, a novel carbohydrate, in the tubers of transgenic potato plants. *Biotechnology*, 9, 982–6.
- Schwall, G. P., Safford, R., Westcott, R. J., Tayal, A., Shi, Y. C., Gidley, M. J. and Jobling A. S. (2000). Production of very-high-amylose potato starch by inhibition of SBE A and B. *Nature Biotechnology*, 18, 551–4.
- Sevenier, R., Hall, R. D., Van der Meer, I. M., Hakkert, H. J. C., van Tunen, A. J. and Koops, A. J. (1998). High level fructan accumulation in a transgenic sugar beet. *Nature Biotechnology*, 16, 843–6.
- Van der Meer, I. M., Ebskamp, S. C. M., visser, R. G. F., Weisbeek, P. J. and Smeekens, S. C. M. (1994). Fructan as a new carbohydrate sink in transgenic potato plants. *Plant Cell*, 6, 561–70.

Lipids and bioplastics

- Bohmert, K., Balbo I., Kopka, J., Mittendorf, V., Nawrath, C., Poirier, Y., Teschendorf, G. and Trethewey, R. N. (2000). Transgenic *Arabidopsis* plants can accumulate polyhydroxybutyrate to up to 4% of their fresh weight. *Planta*, 211, 841–5.
- Bohmert, K., Balbo, I., Steinbuchel, A., Tischendorf, G. and Willmitzer, L. (2002). Constitutive expression of the β -ketothiolase gene in transgenic plants. A major obstacle for obtaining polyhydroxybutyrate-producing plants. *Plant Physiology*, 128, 1282–90.
- Houmiel, K. L., Slater, S., Broyles, D., Casagrande, L., Colburn, S., Gonzalez, K., Mitsky, T. A., Reiser, S. E., Shah, D., Taylor, N. B., Tran, M., Valentin, H. E. and Gruys K.J. (1999). Poly(beta-hydroxybutyrate) production in oilseed leucoplasts of *Brassica napus*. *Planta*, 209, 547–50.

- John, M. and Keller, G. (1996). Metabolic pathway engineering in cotton: biosynthesis of polyhydroxybutyrate in fiber cells. *Proceedings of the National Academy of Sciences, USA*, **93**, 12768–73.
- Mittendorf, V., Robertson, E. J., Leech, R. M., Kruger, N., Steinbuchel, A. and Poirier, Y. (1998). Synthesis of medium-chain-length polyhydroxyalkanoates in *Arabidopsis thaliana* using intermediates of peroxisomal fatty acid β -oxidation. *Proceedings of the National Academy of Sciences, USA*, **95**, 13397–402.
- Napier, J. A., Michaelson, L. V. and Stobart, A. K. (1999). Plant desaturases: harvesting the fat of the land. *Current Opinion in Plant Biology*, **2**, 123–7.
- Poirier, Y., Nawrath, C. and Somerville, C. (1995). Production of polyhydroxyalkanoates, a family of biodegradable plastics and elastomers, in bacteria and plants. *Biotechnology*, **13**, 142–50.
- Topfer, R., Martini, N. and Schell, J. (1995). Modification of plant lipid synthesis. *Science*, **268**, 681–5.

Proteins

- Boothe, J. G., Saponja, J. A. and Parmenter, D. L. (1997). Molecular farming in plants: oilseeds as vehicles for the production of pharmaceutical proteins. *Drug Development Research*, **42**, 172–81.
- Daniell, H., Streatfield, S. J. and Wycoff, K. (2001). Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. *Trends in Plant Science*, **6**, 219–26.
- Fischer, R., Schumann, D., Zimmermann, S., Drossard, J., Sack, M. and Schillberg, S. (1999). Expression and characterisation of bispecific single-chain Fv fragments produced in transgenic plants. *European Journal of Biochemistry*, **262**, 810–16.
- Giddings, G., Giddings, G., Allison, G., Brooks, D. and Carter, A. (2000). Transgenic plants as factories for biopharmaceuticals. *Nature Biotechnology*, **18**, 1151–5.
- Gil, F., Brun, A., Widgorovitz, A., Catalá, R., Martinez-Torrecedrada, J. L., Casal, I., Salinas, J., Borca, M. V. and José, M. E., (2001). High-yield expression of a viral peptide vaccine in transgenic plants. *FEBS letters*, **488**, 13–17.
- Goddijn, O. J. M. and Pen, J. (1995). Plants as bioreactors. *Trends in Biotechnology*, **13**, 379–87.
- Hemming, D. (1995). Molecular farming: using transgenic plants to produce novel proteins and other chemicals. *AgBiotech News and Information*, **7**, 19N–29N.
- Hiatt, A. (1990). Antibodies produced in plants. *Nature*, **344**, 469–70.
- Kusnadi, A. R., Nikolov, Z. L. and Howard, J. A. (1997). Production of recombinant proteins in transgenic plants: practical considerations. *Biotechnology and Bioengineering*, **56**, 473–84.
- Ma, J. K.-C. and Hein, M. B. (1995). Immunotherapeutic potential of antibodies produced in plants. *Trends in Biotechnology*, **13**, 522–7.
- Ma, J. K.-C., Hiatt, A., Hein, M. B., Vine, N. D., Wang, F., Stabila, P., van Dolleweerd, C., Mostov, K. and Lehner, T. (1995). Generation and assembly of secretory antibodies in plants. *Science*, **268**, 716–19.
- Miele, L. (1997). Plants as bioreactors for biopharmaceuticals: regulatory considerations. *Trends in Biotechnology*, **15**, 45–9.

- Peeters, K., De Wilde, C. and Depicker, A. (2001). Highly efficient targeting and accumulation of a Fab fragment within the secretory pathway and apoplast of *Arabidopsis thaliana*. *European Journal of Biochemistry*, 268, 4251–60.
- Staub, J. M., Carcia, B., Graves, J., Peter, T. J., Hunter, P., Nehra, N., Paradkar, V., Schilittler, M., Carroll, J. A., Spatola, L., Ward, D., Ye, G. and Russell, D. A. (2000). High-yield production of a human therapeutic protein in tobacco chloroplasts. *Nature Biotechnology*, 18, 333–7.

Future prospects for GM crops

Introduction

In this last chapter we will look briefly at some of the broader issues and concerns surrounding the use of GM crops. Some of these are scientific; others are economic, social, ethical and moral. Complete volumes could, and should, be written about each issue and its influence on plant biotechnology. It is only by careful consideration of all the issues that any meaningful dialogue can be entered into and decisions made. Some of these issues may appear more relevant than others, but it is worth remembering that very few decisions are taken in isolation, and decisions made in one part of the world can influence, for better or worse, the whole world.

The first part of this chapter will take a brief look at the current status of GM crops in world agriculture. We will then look at some of the concerns that surround the growing of GM crops and GM food, discuss some measures that can be taken to minimise the risk and then see how GM crops and food are regulated to ensure safety. The final section of this chapter will, in broad terms, look at some of the future developments in plant biotechnology.

The current state of transgenic crops

It seems appropriate to spend a little time reflecting on the current status of GM plants in broad terms. Despite public opposition and political difficulties in some regions of the world, particularly in Europe, the area of land cultivated with transgenic crops continues to increase. Figures from the International Service for the Acquisition of Agri-biotech Applications (ISAAA) clearly show the importance of transgenic crops in both developed and developing countries (Box 12.1). It is estimated that in 2001 over 52 million hectares, or 130 million acres, of land was planted with transgenic crops (an increase of 8.4 million hectares, or 19%, on figures for the year 2000), compared with only 1.7 million hectares in 1996. The USA has the greatest amount of land dedicated to transgenic crops, but several developing

BOX 12.1 Synopsis of commercial transgenic crop utilisation in 2001

Transgenic crop cultivation by country

Four countries are responsible for 99% of the area cultivated with transgenic crops. The percentage figure is the proportion of the total area devoted to transgenic crops grown in that particular country.

Country	Area*	%
USA	35.7	68
Argentina	11.8	22
Canada	3.2	6
China	1.5	3

* Area is in millions of hectares.

Transgenic crop cultivation by crop

Four crops account for nearly 100% of the commercially grown transgenic crops. The percentage figure is the proportion of the total area dedicated to transgenic crops made up of that particular crop.

Crop	Area*	%
Soybean	33.3	63
Corn	9.8	19
Cotton	6.8	13
Canola	2.7	5

* Area is in millions of hectares.

Transgenic crop cultivation by trait

Herbicide resistance is the trait found in by far the largest area of transgenic crops, accounting for 77%. With insect resistance, these two traits account for effectively all the area cultivated with transgenic crops.

Trait	Area*	%
Herbicide resistance	40.6	77
Insect resistance (<i>Bt</i>)	7.8	15
Insect resistance (<i>Bt</i>)+	4.2	8
Herbicide resistance		

* Area is in millions of hectares.

BOX 12.1 *Continued*

Transgenic crop cultivation by proportion of total area cultivated

The proportion of the total area for each crop under cultivation which is accounted for by transgenic crops.

Crop	%
Soybean	46 (68) ^a
Cotton	20 (69)
Canola	11
Corn	7 (26)

^a Figures in brackets are for the USA.
* Area is in millions of hectares.

The figures in this Box can give a misleading picture of the current status of GM crops. Although GM crops are grown on a large scale in countries like the USA, and for some crops make up the majority of the acreage, other countries refuse to allow GM crops to be grown or the products of GM to be imported. This difference in the status of GM crops is particularly pertinent if the situation in the European Union (EU) is considered. No EU member countries appear in these figures (although some, such as Spain, Germany and France, have grown small areas of GM crops), and there has been a moratorium on licensing new GM products in the EU since 1998. GM crops and products made from GM ingredients face considerable public antagonism in many EU countries.

countries (Argentina and China) grow a significant area. Over 5.5 million farmers (compared with 3.5 million in 2000) grew transgenic crops, many of these being in developing countries.

Who has benefited from these first-generation GM crops?

An analysis of the figures in Box 12.1 shows that a large area is planted with GM crops. However, just two traits, herbicide resistance and insect resistance (singly or in combination), and four crops, grown predominantly in North America, account for the vast majority of this area. The agribiotech multinationals, like Monsanto and Syngenta, drove the development of these crops. The main beneficiary, other than the agribiotech multinationals, appears to have been the farmer, who has generally enjoyed increased yields and revenues

from growing GM crops, despite a sometimes considerable premium on GM seed. Monsanto puts this increase in revenue at \$US16.46 per acre from *Bt* GM maize in areas where the European corn borer is a problem (see Chapter 6). This increased revenue is, in many cases, associated with a decrease in the use of herbicides and/or insecticides (although the figures are disputed). This reduction in the use of potentially harmful chemicals was trumpeted as one of the real benefits of these first-generation GM crops. However, it is interesting to note that surveys in the USA have shown that environmental benefits are not a driver for the adoption of GM technology by farmers. In many cases, but not all, there does seem to be a real reduction in the use of herbicides or pesticides associated with growing GM crops. For example, the introduction of herbicide-tolerant cotton resulted in increased yields but no decrease in herbicide usage, whereas the adoption of herbicide-resistant soybeans led to significant reductions in herbicide use and only small increases in yield (see Chapter 5). The herbicides used on GM crops are also considered to be more environmentally friendly than are some of the alternatives. There does therefore appear to be a real environmental benefit associated with these GM crops. Other environmental benefits from growing GM crops might also be apparent. For instance, more environmentally friendly cultivation methods (non-tilling) can be used that reduce soil erosion. Biodiversity may also be increased as weeds can be allowed to grow postemergence if herbicide-resistant crops are being grown. However, agricultural management has to respond to the use of these crops if negative effects on the environment are not to materialise. Some farmers in Canada have encountered problems with multiply herbicide-resistant volunteer canola for instance, and the altered herbicide application regime used with some GM crops can also bring about its own problems (increased spread from later sprayings, when the crop is taller, for instance).

What is noticeable about these first-generation crops, though, is the absence of engineered traits that are not primarily for agronomic benefit. The agribiotech multinationals invested a large amount in the development of these crops with the expectation that the investment would eventually be repaid and produce a profit. The technology was relatively simple and there was obvious commercial gain to be had. However, other traits that could considerably improve the general well being of farmers and consumers in both the developing and developed worlds (some of which are discussed later in this chapter) are conspicuous by their absence. This commercially driven development has implications for the future of GM crops in general. Distrust of the large multinationals, who control both the GM seed and the herbicide for instance, contributes significantly to the distrust of GM technology, particularly in Europe. The perception, at least, is that GM crops have failed to live up to the considerable claims that were made for them. Many people therefore view future developments in GM crops, ones that could have significant benefits for the people of the world and the environment, with suspicion. They are seen by many as being unnecessary (at the very least) and commercially driven.

What will drive the development of the future generations of GM crops?

Given the considerable opposition to GM crops in some quarters, it is interesting to briefly consider whether GM crops will continue to develop in the same way.

Commercial concerns will undoubtedly continue to drive the development of some GM crops, but what is being seen now is the development of GM crops with traits other than agronomic. We have already looked at the example of 'Golden Rice' (Chapter 10), which is a pioneering example of a GM crop developed for the developing world and without the backing of a commercial company. Like many of the new developments in GM crops it was primarily developed in academia with backing from public funding and charitable institutions. Later in this chapter we will look at some future developments in GM crops that will have potential benefits for the environment and the people of the developing world.

However, until the benefits of these 'new generation' GM crops become apparent, the technology faces some considerable hurdles to widespread acceptance.

In this next section we will look at how GM crops are regulated and some of the concerns surrounding GM crops. To some extent this regulation reflects the concerns of the general public; as we will see, European legislation, in particular, is being modified to try to allay some of those concerns.

Concerns about GM crops

There is a considerable degree of public concern in many parts of the world about the real and perceived issues associated with GM crops. Some of these are based on fears about the potential impact of the technology on food safety, human health and the environment. At the other end of the spectrum, there is moral or ethical opposition to GM crops on the basis that they are wrong in principle. These public concerns can have a massive impact on plant biotechnology and its adoption. In this section we will look at some of the issues of public concern about GM crops, and examine what plant biotechnologists are doing to allay those public fears.

Antibiotic resistance genes

Previous chapters have looked at the use of antibiotic resistance genes as selectable markers for transformed plants. Plant transformation is a very low-frequency event and some means of selecting the transformed cells from the untransformed cells is required. The use of antibiotics as selective agents was already well established as a fundamental tool of molecular biology and cloning in other organisms, particularly in *Escherichia coli*. Their use in plant

biotechnology was therefore logical. However, the use of antibiotic marker genes has proved to be one of the hurdles to the widespread acceptance of GM crops.

In 1996, Novartis sought approval for a maize variety that carried an ampicillin resistance gene. The UK Advisory Committee on Novel Foods and Processes (ACNFP) blocked this approval for a considerable length of time, but eventually the maize line was approved for cultivation in France. However, concerns about GM crop safety and lack of market acceptance meant that this GM maize was never widely cultivated.

This episode was in some ways unusual, as ampicillin resistance genes are not widely found in the current generation of GM crops. It did, however, help to set the tone for much of the discussion about the safety of antibiotic resistance genes in GM crops. Ampicillin is an antibiotic of the penicillin family that is widely used to treat a variety of infections in humans. The presence of a resistance gene in a genetically modified organism (GMO) released into the environment was perhaps bound to raise fears about creating antibiotic-resistant bacteria, and particularly human pathogens.

However, there are several scientific arguments that indicate that this scenario is unlikely and here we will look at just a few:

1. The antibiotic resistance genes used in creating GM crops were originally isolated from bacteria. Resistance to antibiotics, which can confer a selective advantage, occurs naturally in bacteria, and is often carried on plasmids that can be readily transferred from one strain to another. The development of antibiotic-resistant strains of bacteria is therefore a real problem in hospitals, where antibiotics are used routinely and create a selective pressure for the development of resistance. Against this background, the transfer of antibiotic resistance genes from plants to bacteria (for which there is no known mechanism) will not significantly alter the pool of antibiotic resistance genes in the environment. The transfer of intact, functional, antibiotic resistance genes to gut flora from ingested plant material is also highly unlikely, though not impossible. There is certainly little evidence for the transfer of DNA from food to gut flora, and it is difficult to see why a transgene would behave differently from the DNA in which it is integrated. Even if resistance genes were transferred, no real selective advantage would accrue and the resistant bacterium would be unlikely to survive.

2. Many of the antibiotic resistance genes commonly found in GM crops (such as *nptII*) confer resistance to antibiotics that are not used to treat disease in humans, their use having been superseded by less toxic and/or more effective alternatives.

Despite these explanations of the relative safety of antibiotic resistance genes in GM crops, plant biotechnologists are seeking ways to obviate the need for their use or presence in GM crops. Various alternative selectable marker genes have been developed (see Chapter 4) and technologies that remove selectable markers, so called 'clean-gene' technology, are also being developed (see Chapter 4).

Herbicide resistance and 'super-weeds'

It is possible to use selectable marker genes that confer resistance to herbicides as an alternative to antibiotic resistance genes (see Chapter 4). However, one problem associated with the use of herbicide resistance markers is the potential for the creation of 'super-weeds' (see Chapter 5). Gene transfer of herbicide resistance genes, predominantly via cross-pollination, to weedy relatives of GM crops could create such 'super-weeds'. This problem is unlikely to occur if the herbicide resistance gene was only used as a selectable marker during regeneration from tissue culture. However, the creation of GM crops engineered specifically to express herbicide resistance, as a desirable agronomic trait (see Chapter 5) to simplify crop production, is more problematic.

Gene transfer and the creation of herbicide-resistant weeds is not a theoretical consideration. The usefulness of some herbicides has already been reduced by the presence of naturally occurring herbicide-resistant weedy relatives of several major crops. The transfer of herbicide resistance genes from GM crops could exacerbate this problem. Practically every major crop species has weedy relatives that could be cross-pollinated by a GM crop.

The transfer of herbicide resistance to weeds is already an issue, though at present it is more a question for agricultural management than widespread public concern. The transfer of herbicide resistance genes to weedy relatives may result in the weed becoming resistant to one or more herbicides, but it will still be susceptible to other herbicides. This may, of course, be of major concern to the farmer (or others), as the herbicide of choice may be rendered ineffective against some weeds. However herbicide resistance genes confer no selective advantage on weeds that are not subject to treatment with the herbicide, and therefore the trait is unlikely to spread throughout the population.

Plant biotechnologists can also decrease the use of herbicide resistance genes as selectable marker genes by again either employing alternatives or by using clean-gene technology. A particularly attractive alternative is to engineer the chloroplast genome, which is discussed later in this chapter. Chloroplasts are, in most, but not all cases, inherited maternally, so negating the chance of gene transfer by cross-pollination.

Gene containment

Preventing the transfer of foreign genes from GM crops to other plants is a wider environmental issue, a concern that does not only apply to herbicide resistance genes. A great variety of foreign genes are being introduced into GM crops, but the environmental impact of these genes is currently difficult to predict. In view of these concerns and to prevent what has been described as 'genetic pollution', several strategies for preventing foreign-gene transfer are being developed.

Gene transfer usually occurs through pollen, although GM crops, if a wild relative pollinated them, could also serve as a female parent for hybrid seeds. The dispersal of seeds from GM crops amongst weedy relatives could also produce mixed populations, with introgression of, for example, a herbicide resistance gene resulting in herbicide-resistant weeds. The potential for foreign genes from GM crops to be transferred to weedy relatives depends on a great many variables. If gene transfer is through pollen, the amount produced, the longevity, the dispersal mechanism and distance as well as the proximity of weedy relatives that are sexually compatible with the GM crop all influence the frequency of transfer. Dispersal by seeds obviously depends on whether the GM crop is allowed to set seed, whether the seeds are dispersed amongst or close to weedy relatives and the viability and longevity of those seeds.

Techniques for gene containment

A wide variety of techniques (see Table 12.1) are capable of preventing or reducing gene transfer. Some of the techniques listed in Table 12.1 are, at present, theoretical or have only been demonstrated in laboratory situations. It is also interesting to note that at least one of the techniques, so-called 'terminator technology' (Box 12.2), has itself been the cause of much opposition to GM crops as seed can not be collected and saved by the farmer for planting in subsequent years.

Big business

The threat of terminator technology being introduced led to protests in many parts of the world, and was seen as yet another example of big business imposing its wishes on farmers and consumers alike. It is informative to look at the example of terminator technology as it illustrates some interesting points about the plant biotechnology industry and public perception (Box 12.2).

Terminator technology has not been introduced commercially, although most of the big plant-biotechnology firms are thought to have the technology. Terminator technology has the potential to make GM crops safer by reducing gene transfer to weedy relatives. The publicity has, however, focused on its potential use to prevent farmers from saving a portion of their grain for use as seed in subsequent years, thus tying them to an expensive source of fresh seed each year. Concerns that terminator technology had already been introduced led to GM cotton crops being burnt in India amid wide-scale unrest.

For terminator technology to become a problem for farmers in developing countries (and even in the USA, where 20–30% of soybean farmers reuse seed) they would already have to be growing seed, with terminator technology engineered in, from a commercial company like Monsanto. Otherwise seed can continue to be kept as normal. There is concern, though, that future development will mean that the highest yielding varieties (which governments or banks will promote) will come with technology protection schemes (TPS)

Table 12.1 Techniques for gene containment

Technique	Advantages	Disadvantages
Chloroplast transformation ^a	Maternally inherited. High level of transgene expression	Expressed proteins may not be properly processed
Male sterility ^b (disrupt anther development)	No pollen formed	Crops must be propagated by pollination from a non-GM crop or by other means. Male sterile plant could serve as female parent
Terminator technology ^c	Seeds are sterile. Inducible	Silencing of transgene results in viable seeds and introgression. Terminator transgenes and desirable trait must remain linked
Apomixis ^d (seeds produced without fertilisation)	Can be used to fix heterosis	Complex trait, not usual in crop plants. Plant may not be 100% apomictic
Cleistogamy ^e (fertilisation occurs within unopened flower)	Self-pollination prevents outcrossing	Complex trait, not usual in crop plants. Introgression still occurs
Transgenic mitigation ^f	Introduced trait is advantageous or neutral for the GM crop but deleterious for weeds	Does not prevent gene transfer. May cause extinction of weedy relatives, thus reducing biodiversity

^a See Chapter 3 and Box 12.6.
^b Male sterility can be engineered by interfering with the development of the tapetum, which is part of the anther and is important in pollen development. Tapetum-specific promoters have been used to drive the expression of ribonucleases in the tapetum, resulting in targeted destruction of the tapetum and no pollen development.
^c See Box 12.2.
^d See apomixis section in this chapter.
^e In cleistogamy, self-pollination and fertilisation occur within the unopened flower. The genes responsible for the trait have yet to identified and characterised.
^f The deleterious traits need not be drastic for transgenic mitigation to work, as competition between weeds is intense. An example of transgenic mitigation is preventing seed-pod shatter (preventing the release of seeds), which is deleterious to weeds, but can actually be an advantage in crops.
Data from Daniell, H (2002), with permission.

attached. Farmers will, therefore, be either forced to pay for seed every year or continue to grow the low-yield varieties.

Apomixis

In May 1998, in what may well come to be heralded as the beginning of a new era for GM crops, leading researchers in apomixis declared that the technology needs to remain available to developing countries. The Bellagio Apomixis

BOX 12.2 Terminator technology terminated?

Although by no means the only patented technology that has been labelled as 'Terminator Technology' by its opponents (for instance, Astra Zeneca were issued US Patent 5, 808, 034 in September 1998 to cover their own terminator-type technology, and currently over 30 patents covering similar technology have been issued), it is the technology developed by the USDA and the Delta and Pine Land Co. that has attracted most publicity. Much of this publicity, the vast majority of which was unfavourable to say the least, stemmed from the involvement of Monsanto in the saga following their attempts to acquire Delta and Pine Land Co. and rights to the technology.

Terminator technology is described in US Patent 5, 723, 765 (Control of Plant Gene Expression), issued on March 3rd, 1998, to the USDA's Agricultural Research Service (ARS) and Delta and Pine Land Co., a cotton and soybean breeder. The patent covers technology that was referred to as the 'Technology Protection System' (TPS). Three genes (two from bacteria and one from plants) are introduced into self-pollinating plants in TPS. The key to TPS technology is treatment of the seeds, before sale, with a compound (the inducer) that activates what has been called a 'molecular switch', which initiates a chain of events, culminating with the introduced plant gene being activated shortly before the seed matures fully. The plant gene prevents the seeds from completing their development (after all the marketable elements of the seed, such as oil, have developed) and becoming fertile.

Various modes of action have been proposed for TPS, but one possible method can be illustrated as follows. A repressor gene normally prevents activation of the plant 'terminator' gene (therefore allowing the seed company to use the seed as normal), but after treatment with the inducer (such as tetracycline) the repressor is itself repressed. This results in the second gene (a recombinase, like the Cre enzyme that was also looked at in Chapter 4, for instance) being activated, which leads to a piece of DNA (a 'blocker' flanked by *lox* sites), that prevents expression of the plant gene, being removed. The terminator gene (a ribosome inhibitor protein, for example) is then under the control of a promoter that is only activated very late in seed development. The advantage of TPS is that, in the absence of treatment with the inducer, the seeds can be used as normal. It is only after treatment with the inducer that the seeds produced are destined to be sterile.

TPS was developed to protect the interests of seed companies and, it is claimed, to protect the environment by preventing gene spread (Table 12.1). The seed company's investment in developing and breeding new crops would be protected by preventing farmers from saving seed from one year to the next, thereby avoiding paying for seed (or selling seed illegally). The developers claim that this protection of the interests of seed companies will, in the long term, benefit agriculture because it encourages the development of new strains and varieties.

The future of terminator technology is uncertain. Following widespread public concern and adverse publicity the then chairman of Monsanto, Robert B. Shapiro, wrote to the president of the Rockefeller Foundation, Gordon Conway, in 1999 informing him that Monsanto had no intention of marketing seeds containing terminator technology. Astra Zeneca followed suit shortly after, and no terminator technology seeds have been marketed, though this may, in part, be due to difficul-

BOX 12.2 Continued

ties with the technology. Companies like Monsanto are at pains to point out that terminator technology was not one of their main development interests, but it is interesting to note that many companies continued to work on the technology long after 1999.

Declaration recognised that apomixis offers too great a benefit for it to be controlled by a small number of concerns (which inevitably would be western-based biotechnology companies protecting their investment with patents). The needs of commercial concerns (that will undoubtedly play a major role in technology development) are not ignored, but the emphasis is put on novel approaches to patenting, licensing and technology development to ensure 'open-access' to the technology.

So what is apomixis, and why is it such an important technology? In plants, seeds are usually produced as a result of fertilisation in sexual reproduction. In some plants, however, seeds can be produced without fertilisation, a process known as 'apomixis'. Natural apomixis occurs in a limited number of plant species (about 400 in 40 families), but is not found in any major crop plant. Plants from apomictic seeds are genetically identical to the mother plant and are therefore clones. Any desirable features of the mother plant are retained in subsequent generations.

Apomixis, if it could be introduced into cereal crops like rice, wheat and maize, would provide major advantages to both plant breeding/biotechnology companies and farmers, particularly those in the developing world. Much of the seed sown by farmers is high-performance hybrid seed, produced from parents that have different, desirable characteristics. Hybrid seed is expensive, though, because of the continuous effort needed to produce seed from the parents, as hybrid plants do not breed true and the desirable characteristics are lost. If apomixis were introduced into the hybrid, it could be cultivated indefinitely, as the offspring would be genetically identical and therefore retain the desirable traits. This is known as 'fixing heterosis' (where the desirable characteristics of both parents are combined in one plant).

Plant biotechnology companies are interested in genetically engineering apomixis into crop plants because it can be used to considerably speed up the development of new varieties. The advantage to farmers is that seed, even from high-performance hybrids, can be saved and reused, without the loss of desirable characteristics. This frees the farmer from having to buy expensive seed every year, and would make expensive, high-performance, hybrids available to farmers in the developing world. The advantage of apomixis, the freeing of farmers from dependence on expensive seed is, of course, one of the potential hurdles to its development, as seed companies cannot protect their investment.

Several ways round this problem have been suggested, such as coupling apomixis and terminator technology, so that farmers who could afford to buy new seed each year are supplied with seed with the terminator technology activated. Seed for farmers in the developing world would not have the terminator technology activated and seed could therefore be reused. The problem, according to some from the plant biotechnology industry, may not be all that great, as farmers in the developed world (the natural market for hybrid seed) prefer to grow the latest hybrids anyway, and would not save and reuse on a large scale.

The regulation of GM crops and products

We have looked at a few of the issues surrounding GM crops and seen that some of the concerns are about the safety of such crops, both to human health and the environment. It is perhaps appropriate to look at some of the regulations that apply to all genetically modified organisms (GMOs).

The future of GM crops is, to a large extent, determined by the regulatory framework that is applied to their growth and processing. As a minimum expectation, regulations should ensure that the GM crops and products derived from them are safe (both to the environment and to human/animal health). However regulations pertaining to GM crops, particularly in the European Union (EU), cover additional aspects and reflect a range of public concerns. These additional regulations cover aspects such as the labelling of GM foods and GM-derived products so that consumers can exercise choice over whether to eat GM ingredients. Regulations are also designed to ensure that the public is reassured and confident that any GM crops or products are safe and that they are not being consumed without their knowledge.

The European Union (EU)

Brief introduction to EU regulatory processes

In EU countries, national laws implement various Council Directives and Regulations governing the regulation of GM crops and products. These include the key legislative elements Directive 90/220/EEC on the deliberate release into the environment of genetically modified organisms (which is being replaced by Directive 2001/18/EC in late 2002) and Regulation (EC) 258/97 concerning novel foods and novel food ingredients.

Directive 90/220/EEC (and subsequent amendments) requires an extensive environmental risk assessment (ERA) to be carried out before any authorisation to place the GM crop on the market under the Directive can be given.

The company responsible for the GM crop (the notifier) submits a technical dossier and a summary (the notification) to the competent authority of a Member State. The technical dossier is examined, with particular reference to the ERA, to ensure compliance with the Directive. If the outcome is

favourable then the process progresses, with the dossier being forwarded to the European Commission. The competent authorities of all Member States evaluate the notification and raise any objections. If no objections are received, the product can be placed on the market under the consent of the Member State where the original notification was made (see Box 12.3). If objections are raised, the proposal is subjected to further scrutiny. A Commission proposal is submitted to the Regulatory Committee of Member States and is subjected to a vote. This proposal takes advice from the Scientific Committee on Plants. If a qualified majority of Member States support the proposal, then it is accepted. If no qualified majority is achieved then the proposal is referred to the Council of Ministers of the Environment, where it can be accepted, rejected or sent back to the Commission. The regulatory process is illustrated in Figure 12.1.

In October 2002, Directive 90/220/EEC is being replaced by Directive 2001/18/EC, which considerably tightens the regulations pertaining to the

BOX 12.3 How European legislation works at a national level

We can look in more detail at how GMOs are regulated in the EU by examining their workings in one particular country. We will look at their application in the UK.

In the UK, the Genetically Modified Organisms (Deliberate Release) Regulations 1992 (amended 1995 and 1997), drawn up under several Acts of Parliament (The Environmental Protection Act 1990, The European Communities Act 1972 and the Health and Safety at Work Act 1974), implement the relevant pieces of EU legislation, particularly EU Directive 90/220/EEC. These regulations apply to all GMOs, but by far the most interest has been shown in genetically modified plants.

A step-by-step approach is adopted in the regulation and release of GMOs. Initial developments take place in a contained environment (such as a laboratory or greenhouse). The safety of the GMO is assessed in this contained environment before small-scale release into the environment is permitted. The scale of these releases is gradually increased if safety assessments indicate it is safe to do so.

The Health and Safety Executive regulate the initial, contained use of GMOs in the UK. Releases into the environment are subject to regulation under the terms of the EU Directive 90/220/EEC. Release into the environment can be divided into two categories: small-scale (research) trials; and release for commercial purposes. A detailed environmental risk assessment (ERA), prepared by the applicant, underpins all releases of GMOs into the environment, irrespective of the scale of release. This ERA is subjected to scientific scrutiny, and release is only allowed if any risks are considered to be very low-level. Low-level risks generally mean that the GMO poses no greater risk than a non-GM equivalent.

Approval for a research trial

Applicants must submit a detailed dossier, which includes an ERA, to the Joint Regulatory Authority (JRA), who advise the Secretary of State at DEFRA (Department for Environment, Food and Rural Affairs). The JRA reviews the dossier, which is also forwarded to experts in other government departments and other organisations, such as English Nature. The Advisory Committee on Releases to

BOX 12.3 Continued

the Environment (ACRE) also reviews all applications for release into the environment. ACRE, a committee of independent scientists, advises on both the human health and environmental aspects of the proposal. If ACRE and the other assessors (the Advisory Committee on Novel Food and Processes, for instance, if the GMO is to be used as food) are satisfied that the proposal represents a very low risk, then the release of the GMO into the environment, on a small scale, can proceed. The release approval may be subject to conditions, and scientists from the Central Science Laboratory closely monitor these initial releases. If the JRA, ACRE and relevant ministers are satisfied that no problems are connected with the release of the GMO, the scale of the release can be increased.

Approval for marketing releases

Approvals for large-scale releases, aimed at commercialisation, have to be sought at a European level. Some of the process is similar to that required for approval of a small-scale release. The applicant must submit another dossier to the Member State, which contains—in addition to all the previous information—details on the diversity of sites where it may be used, ecosystems that might be affected and proposals for labelling. The Member State reviews the proposal—if satisfied that the proposal represents a very low risk, the dossier is forwarded to the European Commission for consideration by the Member States as outlined above in the introduction to EU regulations.

Farm-scale evaluations (FSEs)

In a unique programme to attempt to put the evaluation of GMOs on a sound scientific basis, the UK decided that no GM crops would be commercially grown in the UK until the results of a series of FSEs were known. Only if these FSEs show that there are no significant risks to the environment will the commercial planting of GM crops be allowed.

release and marketing of GMOs. In particular, the provisions of the ERA have been strengthened to cover direct or indirect and immediate or delayed effects of GMOs. Postmarketing monitoring of unanticipated effects of GMOs will also be introduced. In moves designed, in part, to allay public fears about some of the dangers of GMOs, antibiotic resistance marker genes that pose a risk to human and animal life will be phased out, GMO products will be clearly labelled and will be traceable and there will be public consultation over proposed releases.

It is hoped that this new legislation will increase public confidence in GM technology and its regulation, so that the EU moratorium on licensing new biotechnology products will be ended.

The EU regulatory framework for GM foods

GM food and GM-derived food ingredients are regulated (under Regulation (EC) 258/97) in much the same way, although there are some slight

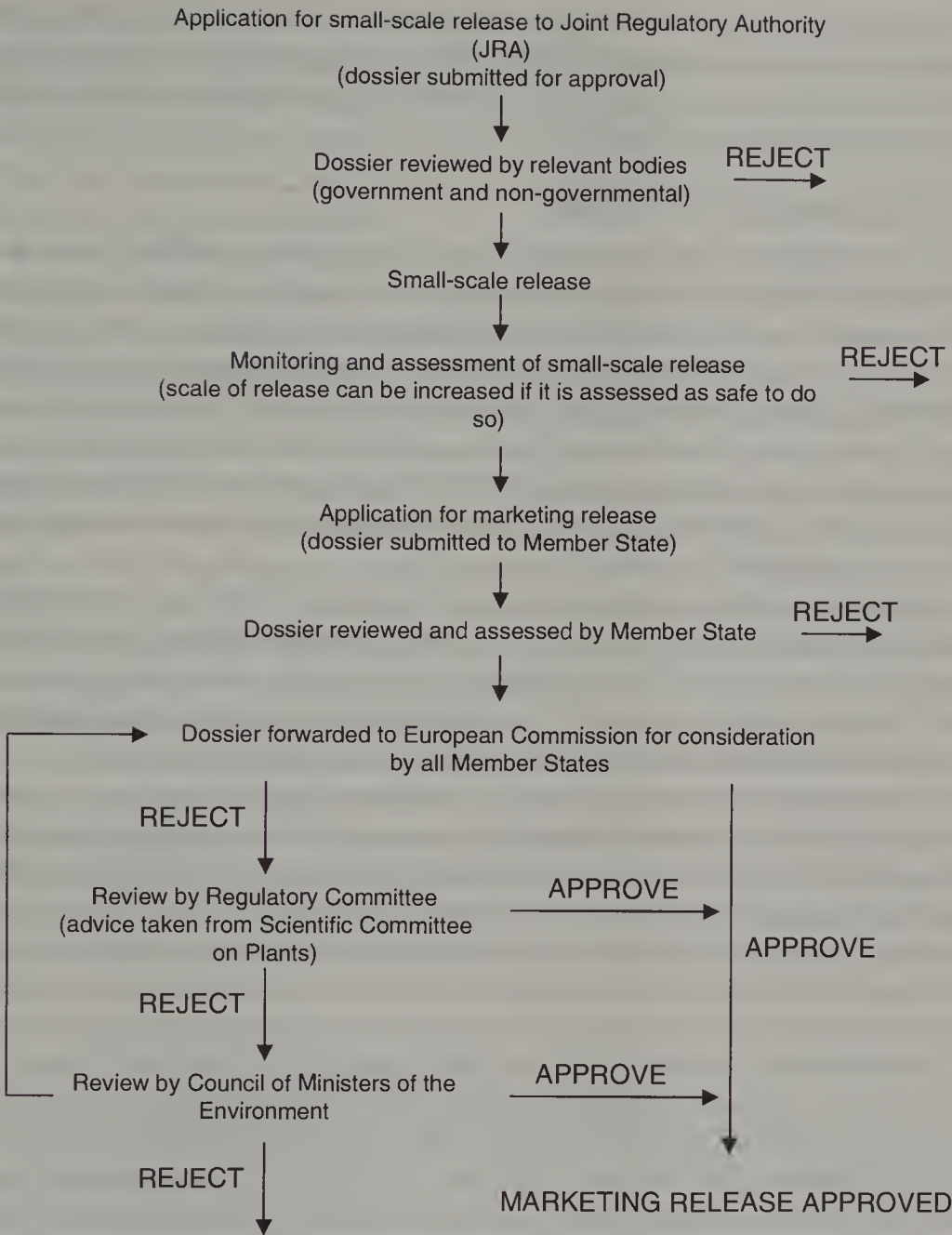


Figure 12.1 A flow chart summarising the regulatory process for the release of GM crops in to the environment in the EU.

differences. Notification includes information on the proposed labelling of the GM food so that consumers can make informed choices about eating GM food or not.

One point worth noting is that foods derived from GMOs (but no longer containing GMOs) which are substantially equivalent to existing foods can be regulated by a simplified procedure (the concept of substantial equivalence is discussed later in this chapter in Box 12.5). In such cases, the European Commission is simply notified (with scientific justification as to why the product is

substantially equivalent, or with an opinion to the same effect from the competent authority of a Member State) when such a product is placed on the market. The notification is passed to Member States and the product can then be marketed throughout the EU.

Concerns about GM food safety

The 'Pusztai affair', and its effect on public confidence in GM foods, was looked at in Chapter 6 (Box 6.4). Whatever the scientific limitations of the work carried out by Pusztai and his colleagues, the revelations certainly contributed to the general anti-GM food feeling in the UK. Several scientists have supported Pusztai's findings, and, as was pointed out in Chapter 6, the area of GM-food safety clearly needs further investigation. One of the contributory factors to the public reaction to the Pusztai report was undoubtedly the loss of public confidence in food safety following the bovine spongiform encephalopathy (BSE) outbreak.

The public feeling, particularly in the UK, against GM foods was badly underestimated by some, and the use of the full weight of the UK academic hierarchy to investigate Pusztai's findings only served to create the impression that Pusztai was being 'silenced' by government, academia and big business. Thus, a bad situation for plant biotechnology was made even worse.

It is important to remember there is no evidence that GM foods are any less safe than non-GM foods. However, the issue of GM food safety has now become much broader. How do plant biotechnologists assess the safety of GM foods, and how does this data get fed into a meaningful debate involving all interested parties? Does the public trust either the information or, indeed, the scientists? If GM foods are to be accepted in Europe, it is incumbent upon plant biotechnologists and governments to engage in a real dialogue with the public, and not simply criticise or ignore public attitudes in a post-BSE Europe.

Part of this dialogue must deal with the very real scientific issues surrounding concepts like substantial equivalence. What sort of compositional analysis, if any, is sufficient, or more than sufficient, to prove substantial equivalence? What alternative methods (like transcriptomics and metabolomics) could be used in analysing GM foods? Is substantial equivalence a scientifically valid and, more importantly, publicly acceptable concept? Part of the dialogue also needs to emphasise the real benefits (e.g. improved nutritional or health values) that GM crops and food could bring to the consumer, rather than to the agribiotech multinationals.

The USA

In the USA, regulations pertaining to GM crops are implemented by three federal agencies.

The US Department of Agriculture (USDA) and, in particular, the Animal and Plant Health Inspection Service (APHIS) under the authority of the Fed-

eral Plant Pest Act determines whether a transgenic crop is likely to become an agricultural or environmental pest. The USDA regulates the transport, import and testing of transgenic plants. In many cases, field tests of GM crops can be conducted if APHIS are simply notified and six criteria regarding safety of the GM crop are met (Box 12.4) and if the applicant meets certain standards to ensure biological confinement.

If these criteria and standards are not met, a more involved process is required to obtain a permit for release from APHIS. The field tests and the subsequent reviews by federal authorities can take many years to ensure the safety of the GM crop to human health and the environment.

GM crops that are to be commercialised can be granted ‘non-regulated’ status by APHIS upon petition. This process requires that extensive data dealing with the nature of the modification and its effects on the crop and the wider ecosystem are presented to APHIS. If non-regulated status is granted, the product and any offspring can be released in the USA without APHIS review. APHIS can prevent the continued use of a GM crop if it is thought to present a risk of becoming a pest. However, the USDA regulatory process has been criticised recently (in a series of National Academy of Sciences reports) for not being rigorous, inclusive and transparent.

The Environmental Protection Agency (EPA) ensures the safety of pesticides, including those produced biologically. Under the authority of the Federal Insecticide, Fungicide, and Rodenticide Act, the EPA regulates the testing, distribution, marketing and use of plants that have been genetically engineered to be pest-resistant. The EPA also sets tolerances for herbicide residues on herbicide-resistant crops. The EPA issues permits for

BOX 12.4 The six criteria for notification in the USA The six criteria that must be met if APHIS is to be notified of environmental release can be summarised as: 1. The plant species is maize, cotton, potato, soybean, tobacco or tomato (or any other plant species that has been so determined). 2. The transgene(s) must be stably integrated. 3. The function of the transgene(s) must be known and expression must not be detrimental to plant health. 4. The transgene does not: (a) result in the production of an ‘infectious entity’ like a virus; (b) encode substances known, or likely, to be toxic to non-target organisms likely to feed or live on the plant. 5. Introduced sequences that are derived from plant viruses must not pose the risk of the creation of new plant viruses. 6. The plant must not contain certain genetic material derived from an animal or human pathogen.

field-testing and registrations for the commercialisation of GM crops that are pest-resistant.

The Food and Drug Administration (FDA), under authority of the Federal Food, Drug and Cosmetic Act, regulates foods and animal feed derived from transgenic plants. Genetically engineered foods have to meet the same safety standards as all other foods. However, many GM crops do not need premarket approval from the FDA because they do not contain substances that are significantly different (they are deemed to be substantially equivalent, see Box 12.5) from those already present in other foods.

One or more of the federal agencies may regulate a particular GM plant, depending on the nature of the modification and its intended use. Thus, herbicide-tolerant crops would be regulated by the USDA (to ensure the GM crop would not become a pest), the EPA (the new use of a herbicide) and the FDA (to ensure the crop was safe to eat). If no food or feed use is intended, the FDA need not be involved in the regulatory process.

As well as federal regulation, some states require additional regulation of GM crops.

The acceptance and regulation of GM crops has been a source of considerable tension between the EU and the USA. The EU moratorium has, for instance, cost the USA over \$200 million annually in corn sales to EU Member States. The new regulations and the lifting of the moratorium should open up the EU market to (properly assessed) GM imports from the USA. However, the USA claims that some aspects of EU regulation put imports from the USA at an unfair disadvantage.

BOX 12.5 **Substantial equivalence and the precautionary principle in risk assessment**

Few issues have generated such a wealth of conflicting data and opinion as has the use of substantial equivalence in GM crop risk assessment. It is not our intention to review the mass of information that surrounds these issues, but simply to make the reader aware of these concepts and some of the issues associated with their use in risk assessments.

Substantial equivalence

The ‘substantial equivalence’ dogma, at least in the context of GM crops, stems from an Organisation for Economic Co-operation and Development (OECD) document (*Safety evaluation of foods derived by modern biotechnology: concepts and principles*), which reported the findings of a group of experts in the safety of food biotechnology. It maintained that the most effective way of assessing the safety of novel foods and food products was to consider whether the novel food product was substantially equivalent to an existing analogous food product (if indeed one existed). Substantial equivalence is based on observation of phenotypic

BOX 12.5 Continued

and agronomic characteristics and a compositional analysis of key components. This, of course, makes perfect sense on one level. After all, if it looks like a potato, grows like a potato and tastes like a potato, chances are it's a potato (even if it is a GM potato) and has, therefore, been safely consumed for several hundred years. Substantial equivalence and its application to GM foods has, though, been criticised by a number of scientists. Firstly 'substantial equivalence' has not been strictly defined, and therefore remains open to interpretation. The use of compositional analysis in determining substantial equivalence has also been criticised as limited and imprecise by some scientists. In the USA in particular (where regulations focus on the product and not the technology used for production), if a GM food is considered to be substantially equivalent then it is exempt, in many cases (self-cloning or if the gene product has previously been approved, for instance), from the necessity of safety evaluation. Like any technology (and conventional breeding is certainly not immune from unexpected results), plant transformation can lead to unexpected results or unforeseen changes in the product. These potentially deleterious effects may not be detected before marketing due to the lack of comprehensive safety testing. It has also been argued that GM crops must be substantially different from their predecessors, by virtue of the way they are produced and/or by virtue of the transgene that is introduced, if they are to be patented.

Substantial equivalence, properly applied, may provide a sound framework for safety assessment (it was never intended that the concept of substantial equivalence should be used as a substitute for a safety assessment). However, its use to justify marketing GM foods and products without comprehensive testing has generated a great deal of public resentment, particularly in Europe.

The precautionary principle

Another approach to guide risk assessment is being adopted in the EU, the 'precautionary principle'. Like the concept of substantial equivalence, it is not a new idea (it was first recognised in the 1970s) and has become a general principle of international environmental law. It engenders what is considered by some to be a fundamentally different approach to risk assessment from the concept of substantial equivalence.

The precautionary principle is invoked in the event of a potential risk, even if this risk cannot be fully demonstrated or quantified, or its effects determined. Appropriate measures, designed to anticipate and prevent harm, should then be taken. The precautionary principle is considered by many to be appropriate when considering GMOs, because of the complexity and novelty associated with GM technology. Basically, in using the principle it is accepted that it is impossible to precisely predict every possible outcome from GM technology, even if at a superficial level the end product appears no different from its 'natural' counterpart.

The precautionary principle has, like substantial equivalence, also been widely criticised in some quarters, being described as unscientific and antitechnology.

However, if GM crops are to become widely accepted, particularly in Europe, a transparent regulatory process, which embraces the precautionary principle, despite any shortcomings, may well prove to be the only acceptable way of dealing with GM crop regulation.

Future developments in the science of plant biotechnology

In previous chapters we have looked at some of the developments taking place in plant biotechnology. Some of these developments are essentially market-driven, whilst some are aimed at either making the technology safer and more acceptable (clean-gene technology, or chloroplast transformation, which is discussed in Box 12.6, for instance) or bringing the benefits of plant biotechnology to people in developing countries.

These developments cannot take place in isolation, because social, political, economic, ethical and moral pressures all influence the development of plant biotechnology. In this last section we will look at just a few of the developments taking place in plant biotechnology.

'Greener' genetic engineering

Although many of the products of plant biotechnology on the market today can offer real environmental and health benefits (from reduced pesticide use or provitamin A rice, for instance), considerable opposition to GM crops remains. However, new approaches to GM crop production are becoming apparent and might help to make the technology more acceptable. These new approaches, which make use of data from genome sequencing projects together with our increased understanding of the fundamental biology underlying agronomically important traits, rely on plant genes to produce the desired effects. Gene transfer from widely differing organisms and the creation of

BOX 12.6 Chloroplast engineering

Engineering of the chloroplast genome provides one of the most exciting developments in modern plant biotechnology, and holds out the promise of being a clean, efficient and environmentally friendly way of genetically engineering plants.

Chloroplasts are normally inherited maternally, in the same way that organelles such as mitochondria are in animals. This means that in the majority of cases, chloroplasts are not found in pollen (the pollen of some plants does contain active chloroplasts, although the chloroplast DNA is usually lost during maturation of the pollen. Some plants also exhibit biparental inheritance of chloroplasts).

This predominantly maternal inheritance has two beneficial effects. First, gene transfer to weedy relatives, which occurs principally through pollen, is reduced. Second, pollen from plants in which the chloroplast genome has been transformed does not represent a toxic risk to non-target species (Chapter 6 highlighted the Monarch butterfly controversy), as the transgene products are not found in pollen. This has proved to be the case even when the transgene product normally accumulates to a very high level.

Chloroplasts are also ideal targets for multigene engineering. In the earlier chapters of this book we looked at several examples of what might be termed first-

BOX 12.6 Continued

generation GM crops. These are, in relative terms, the products of fairly simple, single-gene transformations. Insect- and herbicide-resistant GM crops are prime examples of single-gene transformation GM crops. However, many traits that might be desirable in a GM crop are the products of complex biochemical pathways. Introducing the biosynthetic capability into GM crops might require multiple transgenes to be introduced. This can now be done by introducing multiple transgenes into the nuclear genome, but this is a time-consuming and difficult process. The production of provitamin A rice ('Golden Rice') was discussed in Chapter 10. This was a remarkable achievement, not only because of the potential health benefits to consumers in the developing world, but because it took 7 years to introduce just three genes into rice. Most agronomic traits are, in fact, quantitative and controlled in a polygenic fashion. An efficient method for introducing multiple transgenes therefore becomes a necessity if developments in GM crops are to keep pace with needs.

Chloroplast genes are, like prokaryotic genes, often found in operons, where one promoter controls the expression of a group of genes, polycistronic mRNA being produced. There is therefore the possibility of introducing a group of foreign genes, as an operon, under the control of a single promoter. This has the added advantage that all the genes in the operon are expressed at similar levels. Expression of the *Bacillus thuringiensis cry2Aa2* operon (insecticidal Cry proteins were discussed in Chapter 6) in chloroplasts results in extremely high levels of the transgene product accumulating as stable crystals. This is thought to be due partly to the presence of a chaperonin gene on the *cry2Aa2* operon, which mediates the folding of the Cry protein.

High levels of transgene expression and foreign protein accumulation are relatively common with chloroplast transformation, due to each chloroplast having multiple copies (>50) of the genome and each cell having a number (50 or more, although this figure varies) chloroplasts. Thus, tremendous amplification is achieved if each copy of the genome in each chloroplast carries the transgene.

Gene silencing does not appear to be a problem encountered with chloroplast transformation, and homologous recombination allows transgenes to be inserted precisely into the chloroplast genome. Antibiotic resistance genes need not be used as selectable markers or they can be easily excised. Toxicity of the foreign protein appears to be less of a problem, in many cases, than with nuclear transformation. Chloroplasts apparently seem to offer one of the best hopes for finding an acceptable system for foreign gene expression in plants.

'Frankenfoods' is avoided, and the emphasis shifts towards modification of plant growth, development and physiology to bring about the desired changes.

This new approach could bring benefits to the developed and developing world alike. Plant genes that control or influence agronomically important traits such as time of flowering, dwarfing, bolting and disease resistance have been identified and manipulated to produce agronomic benefits. These benefits could be applied to a range of crops, benefiting producers and consumers from a range of economic backgrounds.

For instance, rice takes only a little more than 6 months to produce a crop in some areas. Introduction of a gene from *Arabidopsis*, called 'LEAFY' (which promotes flowering), into rice accelerates flowering just enough to allow two crops per year, with only a small reduction in yield per crop. Similar approaches have been taken to speed up the reproductive development of citrus trees. The constitutive expression of either 'LEAFY', or another *Arabidopsis* gene known as 'APETALA', shortens the juvenile phase of citrus trees, so that they crop as early as their first year, as opposed to having a delay of more than 6 years.

The development of dwarf strains of cereals like wheat and rice led to the 'Green Revolution' of the 1960s and 1970s, which transformed countries like India into net wheat exporters. So great an achievement was the development of these dwarf strains that Dr N. E. Borlaug was awarded the Noble Peace Prize for his pioneering work in this area. Dwarf strains are generally higher yielding as less of the photoassimilate goes into growth and more goes into the grain. They also have the advantage of being less prone to damage from wind and rain. Dwarf strains can, however, take a considerable time to produce by conventional breeding, and in many cases breeding in dwarfing characteristics has led to the loss of other, desirable characteristics. Scientists are now beginning to manipulate the mechanism (which involves the plant hormone gibberellin) responsible for stem growth directly, using GM technology. Both a dominant-mutant form of the *Arabidopsis* *GAI* (gibberellin insensitive—a gibberellin derepressible repressor of plant growth) gene and the wild-type gene, when introduced into rice, confer reduced responses to gibberellins and result in dwarf rice plants. New, dwarf cereal strains could, therefore, be produced relatively quickly using this approach. The ability to directly manipulate 'elite' varieties also means that the other desirable characteristics might not be lost.

These examples are interesting, not only because of the hope they hold out for increasing crop yield and productivity, but because they also highlight how information and resources from fundamental investigations conducted in model plant species, like *Arabidopsis*, can directly benefit crop plant biotechnology.

These examples also illustrate how plant biotechnology is beginning to address a criticism sometimes levelled against it, that it does not allocate resources to projects aimed at increasing plant yield that would benefit farmers and consumers in developing countries. Other approaches to the problem of plant growth and yield, such as improving the photosynthetic efficiency of plants, are also being adopted. These, and other, developments in plant biotechnology are listed in Table 12.2.

Transgenic plants are also being developed that have an improved nutritional status and/or health value. Some of these, like provitamin A rice, are quite specifically aimed at the developing world, others offer potential benefits to people in both the developing and developed worlds. The new

Table 12.2 Recent and future developments in GM crops

Substance/trait	Gene	Benefit	Plant
<i>Health benefits</i>			
Provitamin A	Phytoene synthase; phytoene desaturase; lycopene cyclase	Vitamin A supplement	Rice ^a
Provitamin A	Phytoene desaturase	Vitamin A supplement	Tomato ^b
Iron	Ferritin	Iron supplement	Rice ^c
Vitamin E	γ -Tocopherol methyltransferase	Vitamin E supplement	Canola ^d
Flavanols	Chalcone isomerase	Antioxidants, reduce risk of heart disease and cancer	Tomato ^e
Fructans	Sucrose:sucrose fructosyl	Low-calorie alternatives to sucrose	Sugar beet ^f
<i>cis</i> -Stearates	Modified acyl–acyl carrier protein thioesterase	Lower risk of heart disease than <i>trans</i> -isomers	Canola ^g
Isoflavones	Isoflavone synthase	Reduces osteoporosis, reduces blood cholesterol	<i>Arabidopsis</i> ^h
Methionine	Antisense threonine synthase	Increased methionine levels	Potato ⁱ
Seed albumin	Seed albumin gene <i>AmA1</i>	AmA1 protein is non-allergenic and rich in all essential amino acids. Increased nutritional value	Potato ^j
Lycopene	S-Adenosylmethionine decarboxylase	Increased lycopene levels in ripe fruit	Tomato ^k
<i>Yield</i>			
Photosynthesis	Phosphoenolpyruvate carboxylase	Improved photosynthetic efficiency, better yield	Rice ^l
Photosynthesis	Fructose-1,2-/sedoheptulose-1,7-bisphosphate	Improved photosynthetic efficiency, better yield	Tobacco ^m
Shade avoidance	Phytochrome A	Inhibits shade response. Plants can be planted at higher densities giving an increased yield per area	Tobacco ⁿ

Table 12.2 Continued

Substance/trait	Gene	Benefit	Plant
Photosynthesis and lifespan	Phytochrome B	Higher photosynthetic performance and longer lifespan	Potato ^o
Dwarfing	Gibberellic acid oxidase	Inhibits GA accumulation and stem growth	Lettuce ^p
Senescence	<i>ipt</i>	Senescence specific expression of <i>ipt</i> delays senescence, prolonging photosynthesis and increasing yield	Tobacco ^q
<i>Others</i>			
Spider silk	Synthetic <i>spidroin1</i> gene	Useful biomaterial for some medical applications	Tobacco and potato ^r
Heavy metals	mutated GUS gene	Mutation assay (which restores activity) used to detect toxins in environment	<i>Arabidopsis</i> ^s
Epoxy fatty acids	Cytochrome p450	Vernolic acid-enriched seed oils can be used to replace petroleum-derived plasticisers in polymer manufacture and in adhesives and paints	Tobacco and soybean ^t
Tolerance to low iron availability	Nicotianamine aminotransferase	Low iron availability is a major abiotic stress	Rice ^u
TNT (trinitrotoluene)	Nitroreductase	Phytoremediation. Reduces the contamination of land by explosives' residues	Tobacco ^v
Lignin	cinnamyl alcohol dehydrogenase/caffeate 5-hydroxy-ferulate O-methyltransferase	Altered lignification. More environmentally benign paper-making processes	Poplar ^w

^a Potrykus, I. (2001); ^b Romer, S., *et al.* (2000); ^c Goto, F., *et al.* (1999); ^d Shintani, D. and DellaPenna, D. (1998); ^e Muir, S. R., *et al.* (2001); ^f Sevenier, R., *et al.* (1998); ^g Faciotti, M. T., *et al.* (1999); ^h Jung, W., *et al.* (2000); ⁱ Zeh, M., *et al.* (2001); ^j Chakraborty, S., *et al.* (2000); ^k Mehta, R. A., *et al.* (2002); ^l Ku, M. S. B., *et al.* (1999); ^m Miyagawa, Y., *et al.* (2001); ⁿ Robson and Smith (1997); ^o Thiele, A., *et al.* (1999); ^p Gan, S. and Amasino, R. M. (1995); ^q Niki, T., *et al.* (2001); ^r Scheller, J., *et al.* (2001); ^s Kovalchuk, O., *et al.* (2001); ^t Cahoon, E. B., *et al.* (2001); ^u French, C. E., *et al.* (2001); ^v French, C. E., *et al.* (1999) and Hannink, N., *et al.* (2001); ^w Pilate, G., *et al.* (2002).

generation of transgenic crops also includes plants with obvious environmental benefits. Transgenic plants can be used, for instance, to detect and treat heavy metal contamination. Transgenic plants can also be used to remove toxic explosives that contaminate some land. Transgenic plants can be used as factories to produce biodegradable, renewable and CO₂-neutral alternatives to oil-based products, for instance (see Chapter 11), which, as oil becomes a rarer and more expensive resource, will allow the dwindling resources to be conserved and only used in applications where no alternative is feasible (Table 12.2).

No plant biotechnologist claims that plant biotechnology and, more specifically GM crops, is the answer to all the world's problems. There are undoubtedly other issues, such as wealth distribution, that require urgent attention.

What these examples show, however, is that plant biotechnology can make a difference to life expectancy and the quality of life, and the environment, in both the developed and developing worlds. Some of these developments would simply not be achievable through conventional plant breeding. Plant biotechnology should be seen for what it is: not a panacea, but part of an integrated package of developments aimed at achieving the goal of sustainable growth and development, and food security, throughout the world.

Summary

In this final chapter we have looked at GM crops in a wider context, considering their acceptance, regulation and future prospects. It is not meant to be an exhaustive survey, but examples have been given to stimulate debate and illustrate how GM crops are no different from many other new technologies in that they will be challenged. What we have seen is that GM crops do have great potential for improving the quality of life of many people, but fully realising this potential depends on acceptance. The lack of public acceptance of GM crops, particularly in Europe, has proved to be a major stumbling block in the development of GM crops. For much of the developed world this lack of acceptance effectively means lack of consumer choice and dwindling export markets for GM producers. Some countries from the developing world, like China, have begun to develop an independent strategy to the development of GM crops, relying on local solutions to local problems. The consequences for the developing world, where countries may not have the resources available for an independent approach like that being adopted in China, could be much worse.

Attempts are being made to persuade the public that GM crops are safe and offer previously unknown opportunities for a better life and a better environment. However, studies show that the public in Europe is still sceptical about the involvement of companies like Monsanto in GM crop production and is not convinced by some of the claims made for GM crops.

The challenge for plant biotechnology is to convince public opinion that GM crops offer real advantages, that can not be realistically achieved any other way, and that the technology is safe. Only then will the world market be fully open to GM crops and products and the full potential realised.

Further reading

Benefits, risks, acceptance and safety of GM crops and GM foods

- Dale, P. J., Clarke, B. and Fontes, E. M. G. (2002). Potential for the environmental impact of transgenic crops. *Nature Biotechnology*, 20, 567–74.
- Daniell, H. (2002). Molecular strategies for gene containment in GM crops. *Nature Biotechnology*, 20, 581–6.
- Fletcher, L. (2001). GM crops are no panacea for poverty. *Nature Biotechnology*, 19, 797–8.
- Fu, X., Sudhakar, D., Peng, J., Richards, D. E., Christou, P. and Harberd, N. P. (2001). Expression of *Arabidopsis* GAI in transgenic rice represses multiple gibberellin responses. *Plant Cell*, 13, 1791–802.
- Gaskell, G., Bauer, M. W., Durant, J. and Allum, N. C. (1999). Worlds apart? The reception of genetically modified foods in Europe and the US. *Science*, 285, 384–7.
- Johnson, B. and Hope, A. (2000). GM crops and equivocal environmental benefits. *Nature Biotechnology*, 18, 242.
- Junk, B. and Gaugitsch, H. (2001). Assessing the environmental impacts of transgenic plants. *Trends in Biotechnology*, 19, 371–2.
- Kleter, G. A., van der Krieken, W. M., Kok, E. J., Bosch, D., Jordi, W. and Gilissen L. J. W. J. (2001). Regulation and exploitation of genetically modified crops. *Nature Biotechnology*, 19, 1105–10.
- Kuiper, H. A., Kleter, G. A., Noteborn, P. J. M. and Kok, E. J. (2001). Assessment of the food safety issues related to genetically modified foods. *Plant Journal*, 27, 503–28.
- Marris, C. (2001). Public views on GMOs: deconstructing the myths. *EMBO Reports*, 21, 545–8.
- Moffat, A. S. (1998). Toting up the early harvest of transgenic plants. *Science*, 282, 2176–8.
- Smyth, S., Khachatourians, G. G. and Phillips, P. W. B. (2002). Liabilities and economics of transgenic crops. *Nature Biotechnology*, 20, 537–41.
- Wolfenbarger, L. L. and Phifer, P. R. (2000). The ecological risks and benefits of genetically engineered plants. *Science*, 290, 2088–93.
- The Advisory Committee on Releases to the Environment (ACRE)—web-link 12.1: <http://www.defra.gov.uk/environment/acre/index.htm>
- AgBioForum—web-link 12.2: <http://www.abioforum.org>
- AgBiotechNet—web-link 12.3: <http://www.agbiotechnet.com>
- AgBioWorld—web-link 12.4: <http://www.agbioworld.org>
- Agricultural Biotechnology—web-link 12.5: <http://www.usda.gov/agencies/biotech>
- Agriculture and Environment Biotechnology Commission—web-link 12.6: <http://www.aebc.gov.uk>
- Biotechnology Industry Organisation—web-link 12.7: <http://www.bio.org>
- The Department for Environment, Food and Rural Affairs (DEFRA)—web-link 12.8: <http://www.defra.gov.uk>
- Council for Agricultural Science and Technology—web-link 12.9: <http://www.cast-science.org>
- Environmental Protection Agency (EPA)—web-link 12.10: <http://www.epa.gov>

- Essential Biosafety—web-link 12.11: <http://www.essentialbiosafety.info/main.php>
- European Federation of Biotechnology—web-link 12.12: <http://www.efbweb.org>
- European Union On-Line—web-link 12.13: <http://europa.eu.int>
- Food and Drug Administration (FDA)—web-link 12.14: <http://www.fda.gov>
- Genetically modified organisms: the regulatory process (DEFRA report)—web-link 12.15: <http://www.defra.gov.uk/environment/fse/leaflet/pdf/gmcrop2.pdf>
- Information Systems for Biotechnology—web-link 12.16: <http://www.isb.vt.edu>
- Institute of Science in Society—web-link 12.17: <http://www.i-sis.org.uk>
- National Agricultural Biotechnology Council—web-link 12.18: <http://www.cals.cornell.edu/extension/nabc>
- NCBE GM Food Information—web-link 12.19: <http://www.ncbe.reading.ac.uk/NCBE/GMFOOD/menu.html>
- New Developments in Crop Biotechnology and their Possible Implications for Food Product Safety—web-link 12.20: <http://www.rikilt.wageningen-ur.nl/Publications/Tekstrapport2000%20004.htm>
- New Scientist* GM Food Special Report—web-link 12.21: <http://www.newscientist.com/hottopics/gm>
- The Plant Journal*. Plant GM Technology—web-link 12.22: <http://www.blacksci.co.uk/~cgilib/jnlpage.asp?tpj&File=tpj&Page=plantGM>
- The United States Department of Agriculture—web-link 12.23: <http://www.usda.gov>
- The United States Mission to the European Union—web-link 12.24: <http://www.useu.be>
- The USDA Animal and Plant Health Inspection Service (APHIS)—web-link 12.25: <http://www.aphis.usda.gov/ppq/biotech>
- The USDA Economic Research Service (ERS)—web-link 12.26: <http://www.ers.usda.gov>

Future developments

- Cahoon, E. B., Ripp, K. G., Hall, S. E. and McGonigle, B. (2002). Transgenic production of epoxy fatty acids by expression of a cytochrome p450 enzyme from *Euphorbia lagascae* seed. *Plant Physiology*, 128, 615–24.
- Chakraborty, S., Chakraborty, N. and Datta, A. (2000). Increased nutritive value of transgenic potato by expressing a nonallergenic albumin gene from *Amaranthus hypochondriacus*. *Proceedings of the National Academy of Sciences, USA*, 97, 3724–9.
- Daniell, H. and Dhingra, A. (2002). Multigene engineering: dawn of an exciting new era in biotechnology. *Current Opinion in Biotechnology*, 13, 136–41.
- Daniell, H., Wiebe, P. O. and Fernandez-San Milan, A. (2001). Antibiotic-free chloroplast engineering—an environmentally friendly approach. *Trends in Plant Science*, 6, 237–9.
- Facciotti, M. T., Bertain, P. B. and Yuan, L. (1999). Improved stearate phenotype in transgenic canola expressing a modified acyl-acyl carrier protein thioesterase. *Nature Biotechnology*, 17, 593–7.
- French, C. E., Rosser, S. J., Davies, G. J., Nicklin, S. and Bruce, N. C. (1999). Biodegradation of explosives by transgenic plants expressing pentaerythritol tetranitrate reductase. *Nature Biotechnology*, 17, 491–4.
- Gan, S. and Amasino, R. M. (1995). Inhibition of leaf senescence by autoregulated production of cytokinin. *Science*, 270, 1986–8.
- Goto, F., Yoshihara, T., Shigemoto, N., Toki, S. and Takaiwa, F. (1999). Iron fortification of rice seed by the soybean ferritin gene. *Nature Biotechnology*, 17, 282–6.

- Hannink, N., Rosser, S. J., French, C. E., Basran, A., Murray, J. A. H., Nicklin, S. and Bruce, N. C. (2001). Phytodetoxification of TNT by transgenic plants expressing a bacterial nitroreductase. *Nature Biotechnology*, 19, 1168–72.
- He, Z. H., Zhu, Q., Dabi, T., Li, D. B., Weigel, D. and Lamb, C. (2000). Transformation of rice with the *Arabidopsis* floral regulator *LEAFY* causes early heading. *Transgenic Research*, 9, 223–7.
- Jung, W., Yu, O., Lau, S.-M. C., O’Keefe, D. P., Odell, J., Fader, G. and McGonigle, B. (2000). Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. *Nature Biotechnology*, 18, 208–12.
- Kovalchuk, O., Titov, V., Hohn, B. and Kovalchuk, I. (2001). A sensitive transgenic plant system to detect toxic inorganic compounds in the environment. *Nature Biotechnology*, 19, 568–72.
- Ku, M. S. B., Agarie, S., Nomura, M., Fukayama, H., Tsuchida, H., Ono, K., Hirose, S., Toki, S., Miyao, M. and Matsuoka, M. (1999). High-level expression of maize phosphoenolpyruvate carboxylase in transgenic rice plants. *Nature Biotechnology*, 17, 76–80.
- Mehta, R. A., Cassol, T., Li, N., Ali, N., Handa, A. K. and Mattoo, A. K. (2002). Engineered polyamine accumulation in tomato enhances phytonutrient content, juice quality, and vine life. *Nature Biotechnology*, 20, 613–18.
- Miyagawa, Y., Tamoi, M. and Shigeoka, S. (2001). Overexpression of a cyanobacterial fructose-1,6-/sedoheptulose-1,7-bisphosphate in tobacco enhances photosynthesis and growth. *Nature Biotechnology*, 19, 965–9.
- Moffat, A. S. (2000). Can genetically modified crops go greener? *Science*, 290, 253.
- Muir, S. R., Collins, G. J., Robinson, S., Hughes, S., Bovy, A., De Vos, C. H. R., van Tunen, A. J. and Verhoeven, M. E. (2001). Overexpression of petunia chalcone isomerase in tomato results in fruit containing increased levels of flavanols. *Nature Biotechnology*, 19, 470–4.
- Niki, T., Nishijima, T., Nakayama, M., Hisamatsu, T., Oyama-Okubo, N., Yamazaki, H., Hedden, P., Lange, T., Mander, L. N. and Koshioka, M. (2001). Production of dwarf lettuce by overexpressing a pumpkin gibberellin 20-oxidase gene. *Plant Physiology*, 126, 965–72.
- Pena, L., Martin-Trillo, M., Juarez, J., Pina, J. A., Navarro, L. and Martinez-Zapater, J. M. (2001). Constitutive expression of *Arabidopsis* *LEAFY* or *APETALA1* genes in citrus reduces their generation time. *Nature Biotechnology*, 19, 263–7.
- Pilate, G., Guiney, E., Holt, K., Petit-Conil, M., Lapierre, C., Leple, J.-C., Pollet, B., Mila, I., Webster, E. A., Marstorp, H. G., Hopkins, D. W., Jouanin, L., Boerjan, W., Schuch, W., Cornu, D. and Halpin, C. (2002). Field and pulping performances of transgenic trees with altered lignification. *Nature Biotechnology*, 20, 607–12.
- Potrykus, I. (2001). The ‘Golden Rice’ tale. *In Vitro Cellular and Developmental Biology—Plant*, 37, 93–100.
- Robson, P. R. H. and Smith H. (1997). Fundamental and biotechnological applications of phytochrome tremogenes. *Plant cell and Environment* 20, 831–839.
- Romer, S., Fraser, P. D., Kiano, J. W., Shipton, C. A., Misawa, N., Schuch, W. and Bramley, P. M. (2000). Elevation of the provitamin A content of transgenic tomato plants. *Nature Biotechnology*, 18, 666–9.
- Scheller, J., Guhrs, K.-H., Grosse, F. and Conrad, U. (2001). Production of spider silk proteins in tobacco and potato. *Nature Biotechnology*, 19, 573–7.

- Sevenier, R., Hall, R. D., van der Meer, I. M., Hakkert, H. J. C., van Tunen, A. J. and Kops, A. J. (1998). High level fructan accumulation in a transgenic sugar beet. *Nature Biotechnology*, **16**, 843–6.
- Shintani, D. and DellaPenna, D. (1998). Elevating the vitamin E content of plants through metabolic engineering. *Science*, **282**, 2098–100.
- Takahashi, M., Nakanishi, H., Kawasaki, S., Nishizawa, N. K. and Mori, S. (2001). Enhanced tolerance of rice to low iron availability in alkaline soils using barley nicotianamine aminotransferase genes. *Nature Biotechnology*, **19**, 466–9.
- Thiele, A., Herold, M., Lenk, I., Quail, P. and Gatz, C. (1999). Heterologous expression of *Arabidopsis* phytochrome B in transgenic potato influences photosynthetic performance and tuber development. *Plant Physiology*, **120**, 73–81.
- Zeh, M., Casazza, A. P., Kreft, O., Roessner, U., Bieberich, K., Willmitzer, L., Hoefgen, R. and Hesse, H. (2001). Antisense inhibition of threonine synthase leads to high methionine content in transgenic potato plants. *Plant Physiology*, **127**, 792–802.

Index

A

- ABA, *see* abscisic acid
- abscisic acid
- regulation of gene expression 23–4, 87
 - response elements 23–4, 87
 - in tissue culture 40, 42, 48–9
- ACC oxidase 232–3, 239–42
- ACC synthase 239, 240
- ACC 239
- ACCase, *see* acetyl CoA carboxylase
- acclimation, cold 219–22
- acetohydroxyacid synthase (AHAS), *see* acetolactate synthase (ALS)
- acetolactate synthase (ALS) 89, 106, 108, 110, 124, 125, 129
- acetosyringone 61
- acetyl CoA – carboxylase (ACCase) 108, 271
- ACNFP, *see* Advisory Committee on Novel Food and Processes
- ACRE, *see* Advisory Committee on Release into the Environment
- ADP-glucose 265; *see also* starch synthesis
- Advisory Committee on Novel Foods and Processes 310, 318
- Advisory Committee on Releases to the Environment 318
- AFLP, *see* amplification fragment length polymorphism
- agar 39, 45
- AgrEvo 120, 124, 139
- Agrobacterium rhizogenes* 67
- Agrobacterium tumefaciens* 55–8
- CP4 gene 116–17, 119, 124
 - plant disease 160, 162
 - plant transformation 55–66, 173, 175
 - strains 65
 - super-virulent 65
 - T-DNA 56–62
 - Ti-plasmid 56–9
 - Vir proteins 58–62
- Agrobacterium*-mediated transformation 55–65
- agrochemical industry 104, 120, 125, 130
- albumin, engineering endosperm proteins 251
- Alcaligenes eutrophus* 262, 276, 278–9
- alfalfa
- defensin gene 174
 - economics of antibody production 300
 - oxidative stress 226–8
 - somatic embryogenesis 48–9
 - vaccine production 295
 - yield 281
- alfalfa mosaic *alfamovirus* (AIMV)
- coat protein mediated resistance 188–90
 - heteroencapsidation 199
- alternative splicing 15–17, 21
- amino acid 9–11, 219; *see also* specific named amino acids
- essential 105–6
 - aromatic 110, 112–14
 - branched chain, 106, 110, 123
- aminocyclopropane, *see* ACC
- aminomethylphosphonic acid (AMPA) 120
- ammonia 120
- AMPA, *see* aminomethylphosphonic acid
- amplification fragment length polymorphism (AFLP)
- in plant breeding 159
- α -amylase inhibitor 148, 150–1
- anatomical defence 163
- Animal and Plant Health Inspection Service 320–1
- anthocyanins, and flower colour 243, 244–5
- anthranilate 114
- antibiotic resistance
- concerns about selectable marker genes 309–10
 - as selectable marker 87–90
- antibodies/plantibodies 261
- custom-made antibodies 291–3
 - medically related proteins 289–91
 - production systems 279–80
- anticodon 10–13

- antimicrobial proteins 163–4
 - disease resistance 173
 - in transgenic plants 174–5
 - see also* defensin proteins, PR proteins
- antioxidant 223, 225, 226–7
- antisense RNA 193–7, 266, 274
- APETALA 326
- APHIS, *see* Animal and Plant Health Inspection Service
- apomixis 313, 315–16
- apple trees, disease resistance and transgenic plants 173, 175
- Arabidopsis*
 - genome 25–30
 - transformation *in planta* 66
- arabis mosaic *nepovirus* (ArMV), coat protein mediated resistance 190–2
- arachidonic acid 275
- arginine vasopressin 298
- aroA* gene 116
- arogenate 114
- Arthrobacter globiformis* 212–13
- ascorbate 223, 224, 225, 226–7, 228
 - monodehydroascorbate 223, 226
 - dehydroascorbate 223, 226
 - peroxidase 226, 228
- ascorbic acid, *see* ascorbate, vitamin C
- asters yellow 160; *see also* phytoplasma, plant diseases
- asulam 109
- atrazine 108, 110, 122, 123–4
- autoimmune disease 301
- auxin; *see also* indoleacetic acid
 - classification of 41
 - herbicide action 109, 114
 - in crown gall disease 56
 - regulation of gene expression 23, 87
 - response elements 23
 - in tissue culture 40–2
 - in tobacco transformation 65
 - T-DNA biosynthetic genes 58–9, 240
- Aventis 120–1, 124, 130, 138–9
- avidin, production in genetically engineered plants 285, 287
- avirulence 164; *see also* gene to gene interactions, HR response

B

- Bacillus amyloliquifaciens* 269
- Bacillus anthracis* 301
- Bacillus subtilis* 262, 268
- Bacillus thuringiensis* 132–4, 137; *see also* Bt
- bacterial pathogens of plants 159–62; *see also* specific names
- bait test 186–7
- bar* gene 89, 120–1
- barley
 - diseases of 160–2
 - PR proteins in transgenic plants 172–3
 - RIP 173
 - α -thionin 174
- barley yellow dwarf *luteovirus* (BYDV) 162
 - coat protein mediated resistance 189
- Basta 110, 111, 120, 124
- beet necrotic yellow vein *furovirus* (BNYVV)
 - bait test 186–7
 - coat protein mediated resistance 189
 - molecular techniques for detection 187
 - rhizomania 185
 - rhizomania resistant crops 187
 - structure 184
- Bellagio Apomixis Declaration 313
- betaine aldehyde dehydrogenase 212–14
 - selectable marker for chloroplast transformation 73, 90
- betaine, *see* glycine betaine
- betalains, and flower colour 244
- Bialaphos 109, 120–1; *see also* glufosinate, phosphinothricin
- binary vectors 64, 79–101
 - basic features 79–82
 - evolution of 95–6
 - families of 93
 - optimisation of components 94–7
- biolistics 55, 66–72
 - gene rearrangement 69
 - transformation of rice 69–72
 - see also* direct gene transfer, particle bombardment
- biopesticide 137
- biopharmaceuticals 296–7
- bioplastics 275–9
- biotrophs 159
- Bipolaris maydis*, and leaf blight 158, 161
- blight 158, 160–2
- Bollgard 138, 141–2
- Borlaug, Norman E. 246, 326
- Botrytis* spp 173
- Brassica napus* (canola, oilseed rape) 282–4; *see also* oilseed rape
- Brazil nuts 251
- bromoxynil 108, 111, 121, 124
- Bt cotton 137–8, 140–2, 145; *see also* Bollgard, cotton
- Bt crops 137–42, 306; *see also* individual crops
- Bt genes 137–42; *see also* cry genes
- Bt maize 138–40, 143–4, 147, 151; *see also* maize
- bulk enzymes 279

C

- cocoa swollen shoot *badnavirus* (CSVV) 162
- Calgene 124, 232, 237
- callus 42–3, 51
- Calvin cycle 254–5
 - manipulation of enzymes 256–7
- cancer
 - anticancer antibodies 291
 - anticancer biopharmaceuticals 297
- canola, *see* oilseed rape
- carbohydrate
 - molecular farming 261–70
 - in tissue culture 39
- β -carotene 223, 227, 241, 243, 247–9
- carotenoid
 - biosynthesis 106, 108–9, 227, 241, 243, 244, 247–8
 - flower colour 244
 - fruit 241–2
 - Golden Rice 247
- carrot 223; *see also* somatic embryogenesis
- cassava African mosaic *bigeminivirus* 162
- CAT, *see* chloramphenicol acetyltransferase
- catalase 223, 225
- cauliflower mosaic *caulimovirus* (CaMV)
 - classification and structure 180
 - 35S promoter 20, 70, 82, 83, 85–7, 191
 - structure/activity relationship 83
 - virus resistance 190
- CBF transcription factor 220–2
- celery 209
- cell culture 35–53
- cell suspension culture 43–4, 51
- cellulases, production in genetically engineered plants 287
- chalcone synthase 244
- chicory 121, 267
- chilling 205, 206, 213, 214, 217, 228; *see also* cold stress
- chitinase 164
 - PR proteins 166
 - in transgenic plants 172–4
- chloramphenicol acetyltransferase 83, 92
- chloroplast 105–6
 - genetic engineering 311, 313, 324–5
 - genome 1–4
 - proteins produced in chloroplasts 301
 - starch 263, 277
 - transformation of 73–4
 - transgenic proteins expression levels 300
 - transit peptide 116–18
 - see also* photosynthesis
- chlorophyll 106, 232–3
- chlorsulphuron 110, 124
- choline 212–14
 - dehydrogenase (CDH) 212–14
 - monooxygenase (CMO) 212–14
 - oxidase (COD) 212–14
 - O-sulphate 211
- chorismate 113–14
- chromatin 1, 4, 12, 14; *see also* gene expression
- chromoplast 106
- Ciba Geigy 124
- cis*-elements 14, 16, 21–4
- citrus tristeza *closterovirus* (CTV)
 - coat protein mediated resistance 189
 - cross protection 185
- Cladosporium fulvum* 176
- clean-gene technology 97–9, 310–11, 324
- coat protein mediated resistance (CPMR) 188–90
 - arabis mosaic *nepovirus* case study 190–2
 - multiple pathogen derived resistance strategies 194
 - risk of transcapsidation 199
 - virus-like particles (VLPs) 191
 - see also* genome structure
- codon 9–13
- codon usage 11
- co-integrative vectors 93
- cold stress 218–19; *see also* chilling
- Coleoptera 132, 133, 136, 146, 150–1
- co-linearity 31–2
- Colorado beetle 132, 138
- comovirus
 - movement proteins 184
 - structure and translation strategy 182–3
- compatible solute 209, 210–11, 215, 218; *see also* osmolyte, osmoprotectant
- concerns about GM crops
 - antibiotic resistance 309–10
 - Golden Rice 250
 - herbicide resistance 139, 140–2, 144, 152–3, 311
 - pest resistance 126–9
- Copy Nature strategy, and insect resistance 145–52
- COR regulon 219–20, 221–2
- corn rootworm 138, 140
- corn, *see* maize
- Corynebacterium* spp 160
- co-suppression 76, 193–4, 196, 236–7
- cotton
 - bioplastic production 262, 279
 - herbicide resistant 117, 124, 126, 128
 - insect damage 132, 148
 - see also* Bt cotton, Bollgard, bollworm, leafworm

cowpea mosaic *comovirus* (CPMV)
 capsid structure 184, 191
 genome structure and translation strategy 182–3
 movement proteins 184
 cowpea trypsin inhibitor (CpTI) 146–50
 Cre-lox 99
 crown gall disease 56
 CRT/DRE 220, 221–2
 cry gene 132–47
 cry1Aa 133, 134, 135–6
 cry1Ab 135, 138, 139, 144, 147
 cry1Ac 133, 135, 138, 139, 141–2, 147
 cry1F 133, 138, 139–40, 142
 cry2A 133, 135, 142
 cry3A 135, 137, 138
 cry3Bb 138, 140
 cry9C 133, 138, 139
 Cry protein 132–47; *see also cry gene*, and individual *cry* genes
 Crystalline protein (ICP), *see Cry protein*
 cucumber mosaic *cucomovirus* (CMV)
 coat protein mediated resistance 189, 194
 commercialisation of resistance 198
 post transcriptional gene silencing 198
 risk studies of transgenic plants 199–200
 satellite RNA 185, 194
 cucumber, PR proteins in transgenic plants 166, 173
 C-value paradox 25
 cyclodextrin 262–6
 cytochrome P450 121–2
 cytokinin(s)
 classification 41
 in plant tissue culture 40–3
 T-DNA biosynthetic genes 58–9
 cytoplasmic male sterility (CMS) 158
 CMS-T 158

D

dark reaction 252, 254–5
 manipulation 257
 De Kalb 138–9
 defective interfering RNAs 194
 defensins
 defense proteins 163
 elicitor response 164
 PR proteins 166
 use in transgenic plants 174
 DEFRA, *see Department for Environment, Food and Rural Affairs*
 dental caries 291
 deoxyribonucleic acid 1
 genomic 1

promiscuous 4
 structure of 1–6
 transcription of, *see transcription*
 Department for Environment, Food and Rural Affairs 317
 detoxification 107, 111, 119, 120–2
 DHFR, *see dihydrofolate reductase*
 Dicamba 41, 109
 Dicer RNase 197
 2,4-dichlorophenoxyacetic acid (2,4-D) 41, 109, 111, 124; *see also auxin*
 dihydrofolate reductase (DHFR) 115
 dihydropteroate synthesis 109
 Diptera 132, 133, 136
 direct gene transfer 55, 66–75
 see also biolistics, electroporation, protoplasts, silicon carbide fibres
 disease resistance pathway 162–70
 DNA, *see deoxyribonucleic acid*
 drought 207, 208, 213, 214, 216–17, 218, 221, 228
 DuPont 124
 dwarfing
 cereal 245–7, 325–6, 328
 potato 255
 tobacco 255
 tomato 242

E

E.coli, *aroA* gene 116–17
 economics of molecular farming 279–81, 298–301
 edible vaccines 293–6
 electroporation 68, 72–4
 elicitors
 inducible 164
 endogenous 164–5
 exogenous 165–6
 embryo culture 45
 embryogenic callus 69
 endoplasmic reticulum 273, 275
 modification of fatty acids 271
 modification of proteins 281, 285
 endosperm, cereal 247–8, 252, 263, 269
 endotoxin- δ , *see Cry protein*
 enhancers
 35S promoter 82–3
 Ti-plasmid 57, 58
 enolpyruvylshikimate 3-phosphate (EPSP) 112–14
 enolpyruvylshikimate 3-phosphate synthase (EPSPS) 109, 112–14, 116–17, 119
 environmental impact
 herbicide-resistant plants 123, 125–9

Bt crops 143–4
 insect resistance to Bt 140–2
 Copy Nature strategy 146
 Environmental Protection Agency (EPA) 139,
 141–2, 321–2
 Environmental Risk Assessment (ERA)
 316–19
 EPA, *see* Environmental Protection Agency
 EPSP, *see* enolpyruvylshikimate 3-phosphate
 EPSPS, *see* enolpyruvylshikimate 3-phosphate
 synthase
 ERA, *see* Environmental Risk Assessment
 erucic acid 273–4
Erwinia spp 160
 Fireblight 175
 target for antimicrobial proteins 174
 ethylene 232–4, 237–41
 in tissue culture 40, 42
 European corn borer (ECB) 132, 138–9
 exon 6–7

F

fatty acid, synthesis 271–5
 FCA 17
 FDA, *see* Food and Drug Administration
 field trials
 herbicide-resistant crops 103–4, 123–4
 insect-resistant crops 137, 149
 Fireblight 160
 transgenic resistance in fruit trees 175
 flavonoids 244
 FlavrSavr tomato 232, 237
 flax 124
 flowering, acceleration of 325–6
 flowers
 ornamentals and flower colour 240, 242,
 244–5
 FLP-*frt* 99
 Food and Drug Administration (FDA) 322
 free radical, *see* reactive oxygen species (ROS)
 scavenger 215, 226
 freezing stress 206, 215, 216, 218–20, 222,
 228
 fructan 209, 211, 216, 266–70; *see also*
 oligofructan, polyfructan
 fruit ripening 232–43
 functional genomics 27
 fungal pathogens of plants 159–62
 fungicides, use in USA 157
 Furovirus, *see* beet necrotic yellow vein
furovirus, wheat soil borne mosaic
furovirus
Fusarium spp 161
 mycotoxins 174

G

gai/GAI 245–7, 326
 gelling agents 39–40
 gene
 amplification 114–15
 duplication 114–15
 structure 6–7
 gene containment 311–13
 gene expression 1–25
 chromatin 4–5, 12, 14
 codon usage 11
 control of 12, 14–20
 gene-for-gene hypothesis 166
 gene silencing 67, 73, 75–6 82, 96, 179
 antisense effects 195
 chalcone synthase 244
 virus resistance and gene silencing 201
 virus suppressor of gene silencing 198
see also transcriptional gene silencing, post
 transcriptional gene silencing
 genetic code 11
 genetically modified crops
 area cultivated 305–7
 benefits 307–9
 concerns 309–12
 antibiotic resistance 309–10
 herbicide resistance 311
see also concerns about GM crops
 future developments in 324–9
 gene transfer, horizontal 128
 genome sequencing 25–32
 genome sizes 25–6
 GFP, *see* green fluorescent protein
 gibberellin(s) 40–2, 245–7
 β -glucanase
 PR proteins 164–5
 induction by elicitors 165
 use in transgenic plants 172–3
 glucocerebrosidase 296
 β -glucuronidase 69–70, 90–2, 285, 287
 glufosinate ammonia, *see* Basta,
 phosphinothricin
 glutamate 120–1
 glutamine 109, 110, 120–1, 217
 glutamine synthetase (GS) 120–1
 glutathione 223, 225, 226, 228
 conjugation 122
 peroxidase 226–7, 228
 reductase 226–7, 228
 synthetase 228, 229
 -S-transferase 122, 228
 glycine, group of herbicides 109, 110
 glycine betaine 209, 212–14
 glycosylation 280, 291
 glyoxylate, breakdown product of glyphosate
 119

glyphosate
 mode of action 109, 110, 112–13
 oxidoreductase (GOX) 89, 111
 resistance 113–14, 116–17, 119, 123, 124, 126–7
 structure 110, 112
 usage 127
 GNA, *see* snowdrop lectin
 Golden Rice 247–51, 309, 325
 Humanitarian Board 249
 IPR 249
 green fluorescent protein (GFP)
 reporter gene 91–2
 selectable marker 88, 90
 sequence optimisation 98
 Green Revolution 245–6, 250, 326
 guard hypothesis 170, 175
 GUS, *see* β -glucuronidase

H

heat shock 220, 221–2, 223
 heat shock protein (HSP) 221–2, 223
 heat shock element (HSE) 221–2
 heat shock factor (HSF) 221–2
 heat stress 206, 207
 tolerance 213–14, 216, 220–2
 herbicide
 broad spectrum 105, 112, 119
 resistance 104–29
 super-weeds 311
 toxicity 106, 127
see also individual herbicides
 heterologous encapsidation 199–201
 high dose/refuge management scheme 142–3
 hirudin, production in *B. napus* 282–5, 297
 histone 5
 Hoechst 120, 124
 Homoptera 132, 150–1
 hydrogen peroxide 212, 222–7
 hydroxyl radical 222, 224–6
 hypersensitive response (HR) 164–8, 171–2
 induction in transgenic plants 175–6
 protein interactions 169

I

identity preservation 274, 299
 idiotypic vaccine 291–3
 imazapyr 108, 123
 imazathapyr 110, 123

imidazolinone 106, 108, 110, 123, 124
 immunoglobulins, *see* antibodies
 inclusion bodies; *see also* Cry protein, crystal protein
 indoleacetic acid (IAA) 109, 114
 inducible disease resistance 158, 163–4, 170, 176
 insect
 larva 132, 134, 136–7
 larva midgut 134, 136–7
 resistance to Bt 140–3, 146
see also Coleoptera, Diptera, Homoptera, Lepidoptera, Orthoptera
 insecticidal crystal protein (ICP), *see* Cry protein
 Integrated Pest Management (IPM) 142
 intellectual property rights (IPR) 249
 International Rice Genome Sequencing Project 31–2
 International Service for the Acquisition of Agri-biotech Applications 305
 Intron 6–7
 IRGSP, *see* International Rice Genome Sequencing Project
 ISAAA, *see* International Service for the Acquisition of Agri-biotech Applications

J

jasmonic acid sequence elements 23
 Jerusalem artichoke 267, 270
 Joint Regulatory Authority (JRA) 317, 319
 JRA, *see* Joint Regulatory Authority

K

kanamycin 65, 87–8, 89, 95
Klebsiella ozaenae 111

L

LEAFY 326
 lectin 148, 151, 152
 legislation
 EU 316–20
 UK 317–18
 USA 320–2
 legumin box 24
 legumin mRNA 16

Lepidoptera 132, 133, 136, 139, 141, 144,
146, 148–9, 150–1
Liberty Link 120–1; *see also* phosphinothricin
light regulated gene expression 17, 23–4
Lightning herbicide 123
lignin 114, 119
linoleic acid 272–3, 275
linolenic acid 272–3, 275
lipid – biosynthesis 106, 109, 261–2, 270–6,
278
localisation, protein 19
luciferase 91, 92
lycopene 239–40, 241–3, 247–8
lymphoma 291, 293
lysine 251–2

M

macronutrients 37–8
maize
 chloroplast genome 2–4
 diseases of 158, 161, 174
 herbicide resistant 117, 121, 122, 123, 124,
 128
 insect resistance 132, 138–40, 143–4, 147,
 151
 molecular farming 262, 266, 269, 270, 281
 relative genome size 26
maltose 39, 48–9
mannitol 209, 211, 215, 216, 264–5
MARs, *see* matrix attachment regions
material transfer agreement (MTA) 249
matrix attachment regions (MARs) 96
media 36–42
Medicago falcata, *see* alfalfa
meristem culture 45
messenger ribonucleic acid (mRNA); *see also*
 ribonucleic acid
 turnover 15–17
 see also transcription, translation;
methionine 239, 251
methyl bromide 157
methyl jasmonate 21
micronutrients 37–9
microsatellites, in breeding 159
microspore culture 45–6
mitochondria, genome of 2–4
Monarch butterfly 143–4
monoculture 158
 disease resistance 168
Monsanto 112, 116, 119, 124, 130, 138–40,
 141
movement protein
 gene silencing 198
 virus encoded 181–4

 virus resistance 190
mRNA, *see* messenger ribonucleic acid
MS, *see* Murashige and Skoog
Murashige and Skoog (MS) 38
mutant 116, 117
mutation 107, 123
Mycogen 138–9

N

Na⁺/H⁺ antiport 218
NaturGard 138–9
necrotrophs 159
nematodes 157–8, 167
nepovirus genome structure 183; *see also*
 arabis mosaic nepovirus
Nicotiana tabacum, *see* tobacco
nitrate reductase, promoter analysis 22
nitrilase 111, 124
nitrile 108, 109
nopaline 56–9
Novartis 124, 138–9
nptII gene 87, 89, 95

O

Ochrobactrum anthropi 111
octopine 56–9
oil bodies 271; *also see* oleosin
oilseed rape
 herbicide resistance 117, 119, 121, 124,
 128
 molecular farming 262, 270–1, 273, 273–5,
 277, 279
 see also *Brassica napus*, canola
oleic acid 272–3, 274, 275, 299
oleosin 271, production system 282–5, 296
oligofructan 263, 270; *see also* fructan
oncogenes, T-DNA 58
D-ononitol 209, 211, 215
organogenesis 50–1
Orthoptera 132, 146, 150
Oryza sativa, *see* rice
oryzalin 109
osmolyte 209, 212, 213, 214, 222; *see also*
 compatible solute, osmoprotectant
osmoprotectant 209, 211, 213, 214, 216; *see*
 also compatible solute, osmolyte
osmotic adjustment 209, 213, 215
osmotic potential 209
osmotic stress 209, 211, 213, 215
osmotin 217
overexpression 107, 114, 116–17

P

- palmitic acid 272–3
- papaya ringspot *potyvirus* (PRSV)
 - coat protein mediated resistance 189
 - commercialised resistance 198
 - disease of citrus plants 185
- paraquat 108, 228, 229
- particle bombardment 55, 67–74
 - gene rearrangement 69
 - transformation of rice 69–72
 - see also* direct gene transfer, biolistics
- pat* gene 120–1
- pathogen associated molecular patterns (PAMPs) 168
- pathogenesis related proteins (PR proteins) 166
 - elicitor response 164
 - systemic response 170
 - transgenic plants 171–3
- pear trees, disease resistance in transgenic plants 175
- pectin methylesterase (PME) 233, 237–8
- PEG mediated transformation 72
- pest resistance 104–5, 132; *see also* insect resistance
- petroselenic acid 275, 299–300
- petunia 114, 116–17, 195
- PGRs, *see* plant growth regulators
- phenylalanine 112–14
- phosphinic acid 109–10
- phosphinothricin 110–11, 119–20, 124
- phosphoenolpyruvate (PEP) 254–5, 257
- photosynthesis 252–5, 257; *see also* dark reaction, Rubisco, phytochromes
- phytase 287–8
- phytochromes 247
 - manipulation 253–6
- phytoene synthase 232, 234, 241–3, 247–8
- phytoplasmas 159–62
 - asters yellow 160
- Phytophthora* spp 160–1
 - cryptogin elicitor 176
 - resistant potatoes 158
 - transgenic resistance 173
- picloram 41, 109, 110, 124
- pinitol 209, 211, 264–5
- pink bollworm 138, 141
- Pioneer HiBred 123, 124
- plant-derived vaccines 293–6
- Plant Genetic Systems (PGS) 120
- plant growth regulators (PGRs) 40–2
- plant pathogens 159–162
- plant regeneration, *see* regeneration
- plant tissue culture, *see* tissue culture
- plant transformation, *see* transformation
- plant viruses 159–162
- plasmid vectors
 - pBIN19 93, 95–6
 - pBluescript 81
 - pGreen 95–6
 - pROK 2 148
 - see also* binary vectors
- plasticity, developmental 35–6, 206
- plastid 106, 116, 117–18
 - plastid transformation 73–4
- pleiotropic effects of transgenes 119
- pollen 127–8, 139, 143–4, 147
- polyadenylation 6–7, 17
- polyfructan 209, 211, 216, 266–70; *see also* fructan
- polygalacturonase (PG) 232–3, 234–8
- polyhydroxyalkanoate (PHA) 275–6, 279
- polyhydroxybutyrate (PHB) 276–9
- polymerase chain reaction (PCR)
 - amplification of virus sequences 191–2
 - plant breeding 159
 - virus detection 186
- Polymyxa betae*, fungal vector for BNYVV 185
- post transcriptional gene silencing (PTGS) 76, 193
 - coat protein mediated resistance 198
 - RNA levels 193, 195, 197
- post translational modification 18–19
 - and protein production systems 280–1
- potato
 - antimicrobial proteins 174
 - biopharmaceuticals 297
 - breeding 158
 - coat protein mediated resistance 189
 - edible pharmaceutical products 280
 - herbicide resistance 124
 - insect resistance 132, 137–8, 147, 150–1, 152
 - molecular farming 262, 266, 268, 269, 279
 - phytochromes in transgenic plants 255
 - phytoplasma 160
 - risk of virus recombination in 199–201, 202
 - vaccines 295
 - virus free plants 185
 - wound inducible promoter 172
 - yield 255–6
- potato leafroll *luteovirus* (*polerovirus*)
 - coat protein mediated resistance 189, 198
 - yield decrease 185
- potato X *potexvirus* (PVX)
 - coat protein mediated resistance 189–90

cross protection 188
 post transcriptional gene silencing 196
 potato Y *potyvirus* (PVY)
 coat protein mediated resistance 188–9
 commercialised resistance 199
 post transcriptional gene silencing 196
 reduction in tuber yield 185
 risk 199
 Potrykus, Ingo 250
 precautionary principle 323
 proline 209, 211, 215–6, 222
 promoter analysis 21–5
 promoter structure 6–7, 14–15
 promoter
 35S 20, 70, 82, 83, 85–7, 191, 282, 290
 for Bt genes 139, 141, 145–7
 inducible 84–7
 abscisic acid 87
 alcohol 85–6
 auxin 87
 copper 86
 heat-shock 86–7
 senescence 87
 steroid 86
 tetracycline 85
 wound 86
 tissue specific 282
 endosperm-specific 248
 oleosin 283–4
 virus vectors-subgenomic 293, 298
 protease 148
 protease inhibitor 146, 148–51
 protein domains 134–5
 protein stability 19–20
 protein storage
 prolamins 252
 protein bodies 252
 vacuole 252
 see also albumin
 proteomics 27
 protoplasts 44–5, 51
 PEG 72
 in transformation 68, 72–4
 provitamin A 247–50; *see also* Golden Rice,
 vitamin A, β -carotene
Pseudomonas spp 160–2
 antimicrobial proteins 174
 R protein interactions 169–70
 transgenic resistance 176
 PTGS, *see* post transcriptional gene silencing
 pTOM clones 232–4
 pTOM5 241–3
 pTOM6 234–8
 pTOM13 237–41
Puccinia graminis 162
 Pusztai, A. 152–3, 320
 pyramiding 140

R

random amplified polymorphic DNA (RAPD),
 in plant breeding 159
 rape, *see* oilseed rape
 reactive oxygen species (ROS) 205–7, 222–7
 receptor, Cry protein 134, 136–7
 recombinases, site-specific 99
 refuge, pest management 141–3
 regeneration 46–52
 importance of genotype 49–50
 plant transformation 51–2
 replicase complex
 virus proteins 181
 function 184
 target for ribozymes 194
 reporter genes 88, 91–2
 resistance genes
 HR and inducible systems 165–8
 pyramid breeding 159
 transgenic plants 175
 restriction fragment length polymorphisms
 (RFLP), in plant breeding 159
Rhizobium spp 159
Rhizoctonia spp 162
 transgenic resistance 173–4
 rhizomania 185–7; *see also* BNYVV
 ribonucleic acid 1, 6–21; *see also* messenger
 ribonucleic acid
 alternative splicing 15, 17, 21
 antisense 266, 274
 processing 6–7, 15–7
 synthesis 6–8
 translation 7, 9–13
 ribosome inactivating protein (RIP) 173
 rice
 biolistics 69–72
 C3 photosynthesis 257
 chitinases and glucanases 172–3
 coat protein mediated resistance 189
 engineering nutritional qualities 251–2
 gene rearrangement 69
 insect-resistance 132, 147, 148, 150–1
 secondary products 164
 signal sequence 291–4
 virus diseases 162
 yield 281
 see also Golden Rice
 rice genome sequencing 30–2
 rice stripe *tenuivirus* (RSV)
 coat protein mediated resistance 189
 rice tungro virus complex 161–2
 rice yellow mottle *sobemovirus* (RYMV)
 coat protein mediated resistance 189
 ricolenic acid 275
 Ri-plasmid 67
 risk assessment 128–9, 144

risk; *see also* concern
 effect of satellite sequences 194
 transcapsidation and recombination
 199–201
 RNA induced silencing complex (RISC)
 197
 RNA viruses
 structure, classification and expression
 systems 181–3
 RNA, *see* ribonucleic acid
 root culture 45
 Roundup 110, 111, 112, 127
 Roundup Ready crops 117, 124, 127
 R-RS 99
 Rubisco 254
 manipulation in C3/C4 systems 255–7
 Rubisco activase 17

S

SAG promoter (senescence related) 256; *see also* promoter, inducible
 salinity 205–7; *see also* salt stress
 salt stress 209–11, 214–15, 216–17, 218
 saprophytes 160
 satellites
 cross protection 185
 RNA protection 193–4
 virus genome component 181
 scab 161–2
 resistant potato 173
 Schering 124
 secondary products as antimicrobials 163–4
 secretory antibodies, IgA 290–1
 selectable markers 65, 69, 74, 87–90
 for plastids 73
 sequence tagged sites (STS), in plant breeding
 159
Serratia marcescens, source of chitinase gene
 173
 shikimate pathway, shikimate 3-phosphate
 112–14
 shoot tip culture 45
 signal sequences
 human serum albumin 297
 targeting to ER 290–1
 signal transduction cascades, map kinases
 169–71
 silicon carbide fibres (Whiskers™) 75
 single chain antibodies 280, 289–91
 snowdrop lectin (GNA) 151–2
 somatic embryogenesis 47–51
 alfalfa 48–9
 carrot 48
 cereal 49–50
 somatotrophin
 expression levels in chloroplasts 297
 sorbitol 209, 215, 216
 Southern corn blight 158, 161
 spinach 209, 212–13, 214
 splicing 7
 squash
 coat protein mediated resistance 189
 commercialised virus resistance 198
 stable expression systems for protein
 production 281–2
 starch 263
 modified 262, 263–6
 synthesis 263–6
 starch branching enzyme (SBE) 266
 starch synthase 266
 StarLink 138, 139
 stearic acid 273–4
Streptococcus mutans
 secretory antibodies and dental carries 290
Streptomyces 160
 stress
 abiotic 205–7, 222
 biotic 206–7, 222
 oxidative 205–6, 222–9
 temperature 205–6, 208–9, 218–21
 water deficit 205–9, 211–15
 stress tolerance 205–29; *see also* stress,
 individual stress tolerances
 S-triazine 110, 124
 subgenomic promoter, TMV 182
 substantial equivalence 319–20, 322–3
 subunit vaccines, plant derived 280, 293
 sucrose, in tissue culture 36, 38, 39
 sugar alcohol 209, 211, 215
 sugar beet 124, 128, 185–6, 189, 262, 267,
 268, 270
see also beet necrotic yellow vein virus
 sulphonylureas 106, 110, 123, 124
 super weeds 126
 superoxide 222–7
 dismutase (SOD) 223–5, 227, 228
 sweet potato 185, 251
 synteny 31
 systemic acquired resistance 170–2
 systemic response 163, 170
 transgenic plants 175

T

TATA box 8, 14–15
 taxonomy, virus 179–80
 T-DNA
 border sequences 57–9, 61, 62, 63, 64
 structure 56–9

transfer 59–62
 technology property rights (TPR) 249; *see also*
 Material Transfer Agreement (MTA)
 terminator technology 312–16
 tetracycline inducible promoter 85
 tetracycline repressable promoter 85
 TGS, *see* transcriptional gene silencing
 Ti-plasmid 56–9
 tissue culture 35–53
 tobacco
 Agrobacterium-mediated transformation of
 63–5
 Calvin cycle and photosynthesis 256–7
 budworm 132, 142
 herbicide resistance 116
 hornworm 132
 insect-resistance 147, 148–9, 150–1
 molecular farming 262, 268, 269, 270, 275,
 279
 organogenesis 50
 pharmaceutical production in 290–1, 295,
 297
 phytochromes 255
 resistance to bacteria 174
 resistance to fungi 173–4, 176
 resistance to viruses 189, 194, 198
 yield 281
 tobacco etch *potyvirus* (TEV)
 post transcriptional gene silencing 196
 tobacco mosaic *tobamovirus* (TMV)
 coat protein mediated resistance 188–90
 movement protein 184
 N gene 202
 structure and expression systems 181
 translation enhancer 191
 α -tocopherol 223, 226–7
 tolerance
 virus infections 187
 tomato 124
 diseases 162
 edible vaccines and biopharmaceuticals
 280, 295–7
 insect resistance 147, 150–1
 resistance proteins 169, 176
 ripening 232–43
 virus resistance 194
 tomato spotted wilt *tosspovirus* (TSWV) 162
 coat protein mediated resistance 189
 post transcriptional gene silencing 196
 totipotency 35–6
 transcapsidation 199–201
 transcription 6–8, 12, 14–15
 transcription factor(s) 16, 220–2
 transcriptional gene silencing 76, 188
 transcriptomics 27
 transfer RNA 9–12
 transformation 55–78

transgenes
 copy number 96
 position 96–7
 features 97
 transgenic crops, *see* genetically modified crops
 transient expression, protein synthesis 285
 idiotype vaccine 291–3
 plant derive vaccines 293
 translation 6–7, 9–12, 18
 repression by uORFs 18
 trehalose 209, 215, 217, 262–5, 270
 triacylglycerol (TAG) 271–2
Trichoderma spp, source of chitinase gene
 172–3
 trichosanthin 296
 tRNA, *see* transfer RNA
 trypsin 285, 287
 trypsin inhibitor 146–51
 tryptophan 112–14
 turgor potential 208, 215
 tyrosine 112–14

U

United States Department of Agriculture
 (USDA) 314, 320–2

V

vaccines 295
 vacuolar targeting 268–9
 vacuole 122, 209, 218, 263, 265, 267–9
 vectors, binary, *see* binary vectors
Venturia spp, scab in fruit trees 173
Verticillium spp 174
 virulence genes 59–62
 virus classification and structure 180
 virus-like particles (VLPs)
 coat protein mediated resistance 188
 arabis mosaic *nepovirus* 191
 TMV resistance 188
 virus replication
 antisense 194
 effect of satellite RNA 193
 viscosity, tomato paste 235, 237–8
 vitamin A 223
 deficiency 247, 250
 see also Golden Rice, β carotene, provitamin
 A
 vitamin biosynthesis 105
 vitamin C, 223; *see also* ascorbic acid
 vitamin E, 223; *see also* α -tocopherol
 vitamins, in tissue culture 36, 39, 43

W

- water deficit, *see* stress, water deficit
- water potential 208–9
- water shell 210–11
- watermelon mosaic 2 *potyvirus* (WMV2)
 - coat protein mediated resistance 189
 - commercialisation of resistance 198
- weedkiller, *see* herbicide
- weeds 104–5, 120–1
 - herbicide resistant 126–8; *see also* super weeds
 - volunteer 127–8
- wheat 124, 128, 150–1, 205, 207, 219, 245–7
 - electroporation 74–5
 - coat protein mediated resistance 189
 - diseases 161–2
 - photosynthesis 257
 - yield 281
- wheat soil borne mosaic *furovirus* (SBWMV) 189
- Whiskers™ 75

X

- Xanthomonas* spp 160–2
 - resistance to 176

- xenobiotic 224
- xylanases 287

Y

- yield
 - antibodies 290–1
 - crop 205–7, 231, 245–7, 252–7
 - biomass 281
 - economics of biopharmaceuticals 299–300
 - hirudin 283
 - protein yield from engineered plants 281–2
- YieldGard 138–9

Z

- Zea mays*, *see* maize
- zeaxanthin 223
- Zeneca 232, 235, 237, 249
- zucchini yellow mosaic *potyvirus* (ZYMV)
 - coat protein mediated resistance 189
 - commercialisation of resistance 198
 - cross protection 185
 - risk studies 200
- zwitterion 209, 211

Plant Biotechnology

The genetic manipulation of plants

Plant Biotechnology offers an engaging introduction to this exciting area of scientific research and technology. Written for advanced undergraduates, postgraduate students and other researchers in this area taking a broad range of courses, Plant Biotechnology provides an overview of the production of GM crops, highlighting the key scientific and technical advances that underpin their development.

The text begins with a summary of current knowledge about plant genome organization and gene expression, followed by an introduction to the techniques of plant tissue culture and genetic transformation and their application to crop plants. A consideration of the design of constructs for plant genetic manipulation precedes a series of chapters covering specific targets for GM crops. These include the genetic manipulation of herbicide resistance, pest resistance and disease resistance. Strategies for engineering stress tolerance and the improvement of crop yield and quality are discussed, and the prospects for molecular farming are considered.

Key themes and strategies are developed using appropriate case studies, which place the science in its broader agricultural/commercial context.

- The first textbook to cover this highly topical subject at a level and depth suitable for undergraduates
- Case studies and keynote boxes throughout the text facilitate learning and revision
- Takes the reader from the basics of the subject (including methodology) through to the latest thinking
- Puts the science into perspective by addressing the wider issues surrounding the GM crops controversy



Companion web site at:

www.oup.com/uk/best.textbooks/biochemistry/slater

- Figures available to download
- Library of links to relevant web sites containing additional information or discussion, including those hosted by the protagonists of the GM debate
- News and updates: overviews of current issues surrounding the subject of plant biotechnology, together with updates on issues raised in the book
- Projects: a number of case-study based problems, encouraging students to consider all the issues surrounding the application of plant biotechnology in practice

The authors:

Dr Adrian Slater is the Deputy Director of the Norman Borlaug Institute for Plant Science Research at De Montfort University, Leicester

Dr Nigel Scott is a Principal Lecturer in plant molecular biology in the Norman Borlaug Institute

Dr Mark Fowler is a Senior Lecturer in plant molecular biology in the Norman Borlaug Institute

OXFORD
UNIVERSITY PRESS

www.oup.com

ISBN 0-19-925468-0



9 780199 254682