Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand). This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.

- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.

- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage. [http://researchspace.auckland.ac.nz/feedback](http://researchspace.auckland.ac.nz/feedback)

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library [Thesis Consent Form](#)
Characterising GIGANTEA interactors: the 
*Arabidopsis* BELL-LIKE HOMEO_DOMAIN 3 and 
BELL-LIKE HOMEO_DOMAIN 10 proteins

Raechel Jean Milich

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

School of Biological Sciences, University of Auckland
August, 2006
ABSTRACT

The ability to detect and respond to environmental signals is fundamental in coordinating floral induction in plants to favourable conditions. An important flowering time cue is day length and it is proposed that light signals are perceived and measured by an interaction between photoreceptors and an internal pacemaker, the circadian clock. The control of flowering has been best characterised in the model plant *Arabidopsis thaliana* L. Heynh (*Arabidopsis*). The *GIGANTEA* (*GI*) gene has a complex role in both the promotion of flowering in response to photoperiod and the regulation of the circadian clock. The expression of *GI* is under circadian control and is affected by day length, light quality and temperature changes. The GI protein is also circadian regulated and is actively degraded in the dark.

The biochemical function of GI is unknown and one method to elucidate the role of this protein is to identify protein interactors. The aim of this thesis project was to characterise proteins that interacted with GI. Previously, the BELL-LIKE HOMEODOMAIN 3 (BLH3) protein was identified as a putative GI protein interactor. As part of this thesis work, yeast 2-hybrid and *in vitro* pull down assays were utilised to confirm the interaction between GI and BLH3. Sequence and phylogenetic analyses were used to further examine the BELL family of proteins. The BELL-LIKE HOMEODOMAIN 10 (BLH10) protein was found to be closely related to BLH3 and also interacted with GI. Reverse 2-hybrid assays were used to determine the regions or domains within the GI, BLH3 and BLH10 proteins required to mediate protein interactions.

Expression assays established that the BLH3 and BLH10 transcripts were present throughout plant tissues and times of development. Further analyses revealed that BLH3 and BLH10 are not directly regulated by the circadian clock. The results of GFP expression assays demonstrated that the BLH3 protein is localised to the nucleus in plant cells. Transgenic *blh3* and *blh10* mutant plants were identified and analysed for flowering and light response phenotypes. BLH3 and BLH10 do not function with GI in the photoperiodic pathway to control flowering, yet the *blh3* and *blh10* mutants do have a flowering phenotype in short day conditions. Like *gi*, the *blh3* and *blh10* mutants exhibited exaggerated hypocotyl elongation in response to red and low light conditions. These results are suggestive of a role for BLH3, BLH10 and GI in flowering and de-etiolation responses to specific light conditions in plants.
ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor Jo Putterill and co-supervisor Richard Gardner for their guidance over the years.

Thanks must go to Mr George Mason for his considerable monetary contribution to this thesis project. I feel privileged to be the first recipient of the George Mason Scholarship for Plant Sciences. Thanks also to the University of Auckland for the Doctoral Scholarship.

A big thank you to all PMBies, both past and present. Where would I be without your help? Special thanks to Keith, Rosemary and Robert for help editing this thesis.

The GFP expression assays were carried out with help from Alex Goldshmidt, Pete Murphy and Adrian Walden. Thanks also to Nga for caring for my Arabidopsis plants and helping with flowering assays.

Thanks to my friends and family, you make life just that bit better in so many different ways. Thanks Mum and Dad for your support when I really needed it! Finally, a big hug for my two boys.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>adenine</td>
</tr>
<tr>
<td>aa</td>
<td>amino acid</td>
</tr>
<tr>
<td>Ac</td>
<td>acetate</td>
</tr>
<tr>
<td>AD</td>
<td>transcription activation domain</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>BD</td>
<td>DNA binding domain</td>
</tr>
<tr>
<td>bHLH</td>
<td>basic/helix-loop-helix class of transcription factors</td>
</tr>
<tr>
<td>BIS</td>
<td>N,N’-Methylene-bis-acrylamide</td>
</tr>
<tr>
<td>BLAST</td>
<td>basic local alignment search tool</td>
</tr>
<tr>
<td>bp</td>
<td>base pair(s)</td>
</tr>
<tr>
<td>Bq/MBq</td>
<td>Bequerels, megaBequerels</td>
</tr>
<tr>
<td>C</td>
<td>cytosine</td>
</tr>
<tr>
<td>CaMV 35S</td>
<td>cauliflower mosaic virus 35S promoter sequence</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>cM</td>
<td>centimorgan(s)</td>
</tr>
<tr>
<td>Col-0</td>
<td><em>Arabidopsis thaliana</em> ecotype Columbia</td>
</tr>
<tr>
<td>CTAB</td>
<td>cetyl-trimethylammonium bromide</td>
</tr>
<tr>
<td>C-terminus</td>
<td>carboxyl terminus</td>
</tr>
<tr>
<td>CVI</td>
<td><em>Arabidopsis thaliana</em> ecotype Cape Verde Islands</td>
</tr>
<tr>
<td>d</td>
<td>day(s)</td>
</tr>
<tr>
<td>Da, kDa</td>
<td>dalton, kilodalton</td>
</tr>
<tr>
<td>DAPI</td>
<td>4′,6-diamidino-2-phenylindole</td>
</tr>
<tr>
<td>dCTP</td>
<td>2-deoxyctydine-5-triphosphate</td>
</tr>
<tr>
<td>D</td>
<td>dark</td>
</tr>
<tr>
<td>DD</td>
<td>continuous dark conditions</td>
</tr>
<tr>
<td>dex</td>
<td>dexamethasone</td>
</tr>
<tr>
<td>DMDC</td>
<td>dimethyl-dicarbonate</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNase</td>
<td>deoxyribonuclease</td>
</tr>
</tbody>
</table>
dNTP  deoxynucleotide triphosphate
DTT  1,4-dithiothreitol
dT  deoxythymidine
EDTA  ethylene diamine tetraacetic acid
EMS  ethyl methylsulfonate
EST  Expressed Sequence Tag
EtBr  ethidium bromide
FR  far-red light
g, µg, ng, pg  grams, micrograms, nanograms, picograms
G  guanine
GAL4  *Saccharomyces cerevisiae* GAL4 transcription factor
GFP  green fluorescent protein
GR  glucocorticoid receptor
GUS  *Escherichia coli* β-glucuronidase reporter gene
h  hour
HA  Influenza A haemagglutinin protein epitope
HD  homeodomain region
HEPES  N-[2-hydroxyethyl]piperazine-N’-[2-ethanesulfonic acid]
IM  infiltration media
IPTG  isopropylthio-β-D-thiogalactoside
kb  kilobase(s)
L, mL, µL  litre, millilitre, microlitre
LB  left T-DNA border
LD  long day conditions (18 h light; 6 h dark)
Ler  *Arabidopsis thaliana* ecotype Landsberg erecta
leu  leucine
L  light
LL  continuous light conditions
m, cm, nm  metre, centimetre, nanometre
MBP  Maltose Binding Protein
min  minute
mRNA  messenger RNA
miRNA  microRNA
M, mM, µM  moles per litre, millimoles per litre, micromoles per litre
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOPS</td>
<td>3-[N-morpholino]propanesulfonic acid</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>mya</td>
<td>million years ago</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Centre for Biotechnology Information</td>
</tr>
<tr>
<td>nptII</td>
<td>neomycin phosphotransferase II gene</td>
</tr>
<tr>
<td>ocs 3'</td>
<td>octopine synthase 3' terminator sequence</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>Pa, kPa</td>
<td>Pascal, kiloPascal</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>PEG4000</td>
<td>polyethylene glycol, MW 4000</td>
</tr>
<tr>
<td>pers. comm.</td>
<td>personal communication</td>
</tr>
<tr>
<td>pfu</td>
<td>plaque forming unit</td>
</tr>
<tr>
<td>pnos</td>
<td>nopaline synthase promoter region</td>
</tr>
<tr>
<td>psi</td>
<td>pounds per square inch</td>
</tr>
<tr>
<td>QTL</td>
<td>quantitative trait loci</td>
</tr>
<tr>
<td>RB</td>
<td>right T-DNA border</td>
</tr>
<tr>
<td>Re</td>
<td>constant red light conditions</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>s</td>
<td>second</td>
</tr>
<tr>
<td>SD</td>
<td>short day conditions (8 h light; 16 h dark)</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SEM</td>
<td>standard error of the mean</td>
</tr>
<tr>
<td>SSC</td>
<td>standard saline citrate</td>
</tr>
<tr>
<td>T</td>
<td>thymine</td>
</tr>
<tr>
<td>TAE</td>
<td>tris-acetate-EDTA</td>
</tr>
<tr>
<td>Taq</td>
<td><em>Thermus aquaticus</em></td>
</tr>
<tr>
<td>TBE</td>
<td>tris-borate-EDTA</td>
</tr>
<tr>
<td>TE</td>
<td>tris-EDTA</td>
</tr>
<tr>
<td>TEMED</td>
<td>N,N,N',N'-Tetramethyl-ethylendiamine</td>
</tr>
<tr>
<td>Tris</td>
<td>tris(hydroxymethyl)-aminomethane</td>
</tr>
<tr>
<td>trp</td>
<td>tryptophan</td>
</tr>
<tr>
<td>T-DNA</td>
<td>transfer DNA</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>U</td>
<td>unit</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>UV</td>
<td>ultra violet light</td>
</tr>
<tr>
<td>V</td>
<td>volt(s)</td>
</tr>
<tr>
<td>v/v</td>
<td>volume/volume</td>
</tr>
<tr>
<td>w/v</td>
<td>weight/volume</td>
</tr>
<tr>
<td>×g</td>
<td>times the force of gravity</td>
</tr>
<tr>
<td>~</td>
<td>approximate</td>
</tr>
<tr>
<td>#</td>
<td>number</td>
</tr>
<tr>
<td>ºC</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>95% CI</td>
<td>95% confidence intervals</td>
</tr>
</tbody>
</table>
# Table of Contents

Abstract ii  
Acknowledgements iii  
Abbreviations iv  

## Chapter 1 Introduction 1

1.1 Overview 1  
1.2 The transition to flowering 2  
1.3 The control of flowering in *Arabidopsis* 3  
1.3.1 *Arabidopsis* as a model plant 3  
1.3.2 Environmental signals that control flowering in *Arabidopsis* 4  
1.3.2.1 Photoperiod 5  
1.3.2.2 Light quality and quantity 5  
1.3.2.3 Vernalisation 5  
1.3.2.4 Growth temperature 6  
1.4 The genetic control of flowering time in *Arabidopsis* 6  
1.4.1 The effect of light on flowering 8  
1.4.2 The light quality pathway 8  
1.4.3 The photoperiodic pathway and the circadian clock 9  
1.4.3.1 The *Arabidopsis* circadian clock 10  
1.4.3.2 The central clock oscillator 11  
1.4.3.3 Light input and the entrainment of the clock 12  
1.4.3.4 Entrainment of the clock to alternative environmental signals 15  
1.4.3.5 The photoperiodic pathway 16  
1.4.3.6 'Florigen' and the photoperiodic signal 19  
1.4.4 The ambient temperature pathway 20  
1.4.5 The gibberellin pathway 20  
1.4.6 Pathways that converge at FLC 21  
1.4.6.1 The autonomous pathway 22  
1.4.6.2 The vernalisation pathway 22
| 1.4.7 | Floral pathway integrators | 22 |
| 1.5 | The GIGANTEA gene | 25 |
| 1.5.1 | The pleiotropic effects of the gi mutation | 25 |
| 1.5.2 | Isolation and characterisation of GI | 27 |
| 1.5.3 | GI expression and the circadian clock | 28 |
| 1.5.4 | The GI protein | 30 |
| 1.5.5 | The characterisation of GI in other plant species | 31 |
| 1.5.6 | Identifying GI protein interactors | 35 |
| 1.5.6.1 | The yeast 2-hybrid assay | 35 |
| 1.5.6.2 | The TALE homeodomain proteins | 36 |
| 1.5.6.3 | BLH3 can be further classified as a member of the BELL family | 37 |
| 1.6 | Aims of this thesis | 40 |

### CHAPTER 2 MATERIALS AND METHODS 41

| 2.1 | General | 41 |
| 2.1.1 | Antibiotics | 41 |
| 2.1.2 | Buffers and solutions | 42 |
| 2.1.3 | Plasmids | 43 |
| 2.1.4 | Oligonucleotides | 45 |
| 2.1.5 | Chemicals and laboratory consumables | 48 |
| 2.2 | Bacteria | 48 |
| 2.2.1 | Bacterial strains | 48 |
| 2.2.2 | Bacterial media | 49 |
| 2.2.3 | Transformation of bacteria | 49 |
| 2.2.3.1 | Transformation by electroporation | 49 |
| 2.2.3.2 | Transformation of Agrobacterium | 50 |
| 2.2.4 | Long term storage of bacteria | 50 |
| 2.2.5 | Screening a cDNA library | 50 |
| 2.3 | Yeast | 51 |
| 2.3.1 | Yeast strain | 51 |
| 2.3.2 | Yeast media | 51 |
2.3.3 Transformation of yeast
2.3.4 Yeast mating and yeast 2-hybrid
2.3.5 Long term storage of yeast

2.4 Arabidopsis

2.4.1 Growth of Arabidopsis
  2.4.1.1 Plant growth media
  2.4.1.2 General growth conditions
  2.4.1.3 Harvesting Arabidopsis seed
  2.4.1.4 Growing seedlings on agar plates
  2.4.1.5 Analysis of flowering time
  2.4.1.6 Hypocotyl elongation assays

2.4.2 Transformation of Arabidopsis
  2.4.2.1 Growth of plants for transformation
  2.4.2.2 Agrobacterium-mediated floral dip transformation

2.4.3 Cross-pollination of Arabidopsis

2.4.4 Analysis of transgenic plants
  2.4.4.1 Screening for transformants
  2.4.4.2 Determination of T-DNA copy number

2.4.5 Transient expression of GFP using biolistics

2.4.6 GUS staining and visualisation

2.5 Molecular techniques

2.5.1 Plasmid DNA isolation
  2.5.1.1 Isolation of E. coli plasmid DNA
  2.5.1.2 Purification of plasmid DNA for sequencing
  2.5.1.3 Isolation of plasmid DNA from Agrobacterium
  2.5.1.4 Isolation of plasmid DNA from yeast

2.5.2 Isolation of nucleic acids from Arabidopsis
  2.5.2.1 Small scale extraction of genomic DNA
  2.5.2.2 Miniprep scale extraction of genomic DNA
  2.5.2.3 Doyle and Doyle extraction of genomic DNA
  2.5.2.4 Isolation of RNA from Arabidopsis
  2.5.2.5 Purification of RNA

2.5.3 Manipulation and examination of nucleic acids
  2.5.3.1 Enzymes and Markers
2.5.3.2 Agarose gel electrophoresis
2.5.3.3 Polyacrylamide gel electrophoresis (PAGE)
2.5.3.4 Dephosphorylation of vector DNA
2.5.3.5 Ligation of DNA
2.5.3.6 Restriction enzyme digestion
2.5.3.7 Quantification of nucleic acids
2.5.3.8 DNA sequencing

2.5.4 Polymerase chain reaction (PCR)
2.5.4.1 PCR from DNA
2.5.4.2 Reverse Transcriptase (RT)-PCR
2.5.4.3 Genome walking to isolate flanking DNA

2.5.5 Purification of DNA
2.5.5.1 Purification by phenol:chloroform extraction
2.5.5.2 Extraction of DNA from agarose gels
2.5.5.3 Purification of PCR products

2.5.6 Hybridisation analysis
2.5.6.1 Transfer of nucleic acids
2.5.6.2 Preparation of radio-labelled DNA probes
2.5.6.3 Hybridisation and quantification of gene expression

2.5.7 Protein work
2.5.7.1 Protein extraction from total plant tissue
2.5.7.2 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)
2.5.7.3 Western transfer and immunodetection
2.5.7.4 Synthesis of radio-labelled protein
2.5.7.5 Purification of MBP fusion proteins
2.5.7.6 In vitro protein binding assay

2.6 Computer analysis
2.6.1 General computing
2.6.2 Sequence analysis
2.6.3 Phylogenetic analyses
2.6.4 Microscopy, phosphorimages and photographs

2.7 Compliance
CHAPTER 3  THE BLH3 HOMOEODOMAIN PROTEIN INTERACTS WITH GI  75

3.1  Introduction  75
   3.1.1  Previous work  75

3.2  Testing the GI and BLH3 protein interaction  76
   3.2.1  Identification and repair of a pAS:GI error  76
   3.2.2  Confirmation of the GI::BLH3 protein interaction  78
   3.2.3  Investigation of the GI::BLH3 interaction by \textit{in vitro} protein interaction assay  79
      3.2.3.1  Creating MBP fusion constructs  79
      3.2.3.2  Expression of MBP fusion proteins  81
      3.2.3.3  Expression of radio-labelled proteins \textit{in vitro}  82
      3.2.3.4  Protein interaction assays  85

3.3  A study of the BELL proteins  87
   3.3.1  Introduction  87
   3.3.2  Identification of BELL-like proteins  87
   3.3.3  Analysis of BELL protein structure  88
   3.3.4  A phylogenetic analysis of the BELL family  90
      3.3.4.1  BLH3 and BLH10 group together  93
   3.3.5  Genomic organisation of the BELL family  93

3.4  Analysis of the BLH3 and BLH10 genes  96
   3.4.1  Identification of the BLH3 cDNA  96
   3.4.2  Identification of the BLH10 cDNA  97
      3.4.2.1  Which form of the BLH10 cDNA is expressed in plants?  97
      3.4.2.2  Screening a cDNA library  99
      3.4.2.3  Characterising the BLH10 cDNA  101
   3.4.3  Which ATG is the translation start site?  101
   3.4.4  Analysis of the sequences upstream of BLH3 and BLH10  102

3.5  Discussion  103
   3.5.1  The BELL family  104
      3.5.1.1  The expansion of the BELL family  104
      3.5.1.2  Function of the BELL proteins  105
   3.5.2  Comparison of the BLH3 and BLH10 cDNAs  107
CHAPTER 4  ANALYSIS OF THE GI AND BLH3 PROTEINS USING A REVERSE 2-HYBRID SYSTEM  111

4.1  Introduction  111
   4.1.1  BELL protein interaction domains  111
   4.1.2  Which region of GI is important for mediating protein interactions?  112
   4.1.3  Does GI interact with other BELL proteins?  112

4.2  Creation of the BLH3 deletion constructs  113
   4.2.1  The construction of pACT:Dom1 and pACT:Dom4  114
   4.2.2  The construction of pAS:BLH3  114
   4.2.3  Summary  115

4.3  The creation of GI protein deletions  116
   4.3.1  Analysis of the structure of GI proteins  116
   4.3.2  The GI deletion constructs  118
      4.3.2.1  Construction of the pAS:Del5 deletion  119
      4.3.2.2  Construction of the pAS:Del6 deletion  120
      4.3.2.3  Construction of the pAS:Del7 deletion  121
   4.3.3  Summary  121

4.4  Analysis of the BLH3 and GI deletions  122
   4.4.1  The BELL and homeodomain regions are involved in the BLH3::GI interaction  122
   4.4.2  BLH3 homodimerises in yeast 2-hybrid  124
   4.4.3  Deletions of GI interact with BLH3  124
   4.4.4  Proteins expressed by the deletion constructs in yeast  125

4.5  Does GI interact with other BELL proteins?  125
   4.5.1  Generation of two BLH10 deletion constructs  126
   4.5.2  The generation of alternative BELL constructs  128
4.5.3 Testing BELL interactions with yeast 2-hybrid 129

4.6 Discussion 131

4.6.1 The BELL and homeodomain regions are important for mediating BELL::GI interactions 131

4.6.2 GI interacts with BLH10 in yeast 2-hybrid 132

4.6.3 Two regions of GI can mediate protein interaction 133

4.6.4 Protein interactions within the BELL family 133

4.6.4.1 BELL::BELL interactions are a feature of the BELL family 133

4.6.4.2 Functions of BELL protein interactions in plants 134

4.6.5 Summary 136

CHAPTER 5 ANALYSIS OF BLH3 AND BLH10 EXPRESSION 137

5.1 Introduction 137

5.2 When and where are BLH3 and BLH10 expressed in plants? 138

5.2.1 Time-course of BLH3 and BLH10 expression 139

5.2.2 Developmental regulation of BLH3 and BLH10 expression 141

5.2.2.1 BLH3 and BLH10 are detected by RT-PCR in all plant tissues tested 141

5.2.2.2 Optimising the Northern hybridisation protocol to detect BLH3 and BLH10 141

5.2.2.3 Northern hybridisation analysis of BLH3 and BLH10 142

5.2.3 BLH3 and BLH10 expression during the day re-examined by Northern analysis 144

5.2.4 Analysis of BLH3 expression in circadian clock mutants 145

5.2.5 BLH3 expression in response to red light 147

5.3 Localisation of BLH3 using the GFP reporter gene 148

5.3.1 The GFP reporter constructs 149

5.3.2 Localising BLH3 using transient expression assays 149

5.3.2.1 Optimising GFP expression assays 149

5.3.2.2 GFP::BLH3 is nuclear localised 150
5.3.3 Stable expression of GFP fusion proteins in transgenic Arabidopsis

5.3.3.1 Introduction of expression vectors into Agrobacterium 153
5.3.3.2 Transformation of Arabidopsis and the selection of transformants 154
5.3.3.3 Examination of GFP expression in transgenic plants 154
5.3.3.4 Why was GFP not detected in transgenic plants? 155

5.4 Discussion 156

5.4.1 Comparison of the expression of GI, BLH3 and BLH10 in plants 156
5.4.2 BLH3 expression in circadian clock mutants 158
5.4.3 Summary 159

CHAPTER 6 CHARACTERISATION OF BLH3 AND BLH10 161

6.1 Introduction 161

6.2 Over-expression of BLH3 and BLH10 in transgenic plants 162

6.2.1 Generation of binary vectors for use in plant transformation 162
6.2.1.1 Construction of p35S:BLH3 162
6.2.1.2 Construction of p35S:BLH10-R 162
6.2.2 Introduction of p35S:BLH3 and p35S:BLH10-R into Arabidopsis 163
6.2.2.1 Transformation of Agrobacterium with two binary constructs 163
6.2.2.2 Agrobacterium-mediated transformation of Arabidopsis 164
6.2.3 Analysis of transgenic plants containing p35S:BLH3 165
6.2.3.1 Verifying the presence of the p35S:BLH3 insertion in transgenic lines 165
6.2.3.2 Assessing the phenotype of transgenic plants 166
6.2.3.3 Confirmation of T-DNA insertion copy number 167
6.2.3.4 Analysis of BLH3 expression in transgenic plants 168

6.3 Identification of blh3 and blh10 mutant plants 169

6.3.1 Screening of Wisconsin lines 169
6.3.2 Screening of Syngenta lines 170

6.3.2.1 Back-crossing blh3 and blh10 mutants 170

6.3.2.2 Characterising the T-DNA insertions in the blh3 and blh10 mutants 171

6.3.2.3 Confirmation of T-DNA number in the blh3 and blh10 mutants 173

6.3.2.4 Analysis of expression in mutant plants 173

6.3.3 Identification of blh3 and blh10 point mutants 174

6.3.4 The generation of double and triple mutants 175

6.3.4.1 blh3gi-2 double mutants 175

6.3.4.2 blh3blh10 double mutants 176

6.3.4.3 blh3blh10gi-2 and blh10gi-2 mutants 176

6.3.4.4 Verifying the T-DNA number in double mutants 176

6.4 Analysis of mutant phenotypes 178

6.4.1 Flowering time assays 178

6.4.1.1 Flowering time of back-crossed blh3 and blh10 lines 179

6.4.1.2 Plastochron index 181

6.4.1.3 Flowering time in two SD light conditions 182

6.4.1.4 Flowering time of double and triple mutants 182

6.4.2 Light response assays 185

6.4.2.1 Analysis of hypocotyl elongation in SD 185

6.4.2.2 Analysis of hypocotyl elongation in red light 187

6.4.2.3 An alternative red light hypocotyl assay 190

6.4.2.4 Hypocotyl elongation in plants containing point mutations 190

6.4.3 Over-expression of GI in the blh3 mutant background 191

6.4.3.1 Generation of 35S:GI transgenic plant lines 192

6.4.3.2 Analysis of flowering time in 35S:GI transgenic plants 192

6.4.3.3 Hypocotyl elongation in 35S:GI lines 193

6.4.4 LHY expression in the blh3 and blh10 mutants 193

6.5 Discussion 194

6.5.1 Over-expression of BLH3 in transgenic plants 194

6.5.2 Over-expression of BLH10-R in transgenic plants 195

6.5.3 What is the role of BLH3 and BLH10 in flowering? 195
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.5.4</td>
<td>What is the role of <em>BLH3</em> and <em>BLH10</em> in hypocotyl elongation?</td>
<td>197</td>
</tr>
<tr>
<td>6.5.5</td>
<td>Improving light response assays</td>
<td>199</td>
</tr>
<tr>
<td>6.5.6</td>
<td>Summary</td>
<td>200</td>
</tr>
</tbody>
</table>

**CHAPTER 7  CONCLUDING DISCUSSION**  

7.1  Introduction  203

7.2  Remarks on GI and BELL interactions  204

  7.2.1  The BELL protein interaction network  204
  7.2.2  Protein interactions and BELL function  205
  7.2.3  GI and BELL protein interactions  206
    7.2.3.1  Protein sequences that mediate interactions  206
    7.2.3.2  What can be learnt from BELL::GI interactions?  207

7.3  Perspectives and future work  208

  7.3.1  *BLH3* and *BLH10* function in plants  208
  7.3.2  A complex role for *GI* in plants  209
    7.3.2.1  The light response of *GI*  209
    7.3.2.2  The circadian clock and *GI* function  210
    7.3.2.3  Identifying *GI* targets  211
    7.3.2.4  Where is *GI* required to promote flowering?  211
  7.3.3  Application of knowledge to other plants  212

7.4  Overview  213

**APPENDIX I  ANALYSIS OF THE BELL FAMILY PROTEINS**  215

Appendix 1.1  Pairwise comparison of amino acid sequence of predicted *Arabidopsis* BELL proteins  215
Appendix 1.2  The sequences of the BELL family proteins used in phylogenetic analyses  216
Appendix 1.3  Sequence alignment of the BELL family  224
Appendix 1.4  Analysis of the genomic structure of the *Arabidopsis* BELL family and flanking genes  237
Appendix 1.5 The *BLH3* cDNA sequence and translation 241
Appendix 1.6 The assembly and translation of the *BLH10* cDNA sequence 243

**APPENDIX II**  
**CONSTRUCTS USED FOR YEAST 2-HYBRID ASSAYS** 247

Appendix 2.1 pAS:GI plasmid map 247
Appendix 2.2 Restriction enzyme maps of *BLH3* deletion constructs 247
    pACT:Dom2 and pACT:Dom3 248
Appendix 2.3 Restriction enzyme maps of GI deletion constructs 247
    pAS:Del1, pAS:Del2 and pAS:Del4 248

**APPENDIX III**  
**SEQUENCES OF THE GI PROTEINS** 249

**APPENDIX IV**  
**GFP EXPRESSION ANALYSES** 252

Appendix 4.1 Transient GFP expression in *Arabidopsis* leaf tissue 252
Appendix 4.2 GFP expression in transgenic *Arabidopsis* plants 253

**APPENDIX V**  
**GENERATION OF ANTIBODIES TO BLH3** 254

**APPENDIX VI**  
**SUMMARY OF BLH3 AND BLH10 MUTANTS** 256

**APPENDIX VII**  
**AN INDUCIBLE GI EXPRESSION SYSTEM** 257

Appendix 7.1 Introduction 257
Appendix 7.2 The ethanol inducible gene expression system 258
Appendix 7.3 The dexamethasone inducible gene expression system 262
Appendix 7.4 Summary 267

**REFERENCES** 271
CHAPTER ONE: INTRODUCTION

1.1 Overview

Co-ordinating flowering with the most favourable season is important for the survival of a plant species. Environmental and endogenous information is processed and integrated by the plant to time flowering to conditions that are most beneficial for pollen and fruit production. Synchronising flowering within a species also facilitates cross-pollination, thus ensuring genetic diversity is maintained. The cruciferous plant *Arabidopsis thaliana* L. Heynh (*Ababidopsis*) has been employed as a model system for the study of the environmental and genetic regulation of reproductive development (reviewed by Koornneef *et al.*, 1998a, Simpson *et al.*, 1999; Mouradov *et al.*, 2002; Jack, 2004; Komeda, 2004). The current understanding of the genetic control of flowering is due to work with *Arabidopsis*. It has become apparent that there are distinct signalling pathways and interacting networks that either induce or repress flowering in response to internal and environmental stimuli. The basic framework of flowering regulation in *Arabidopsis* appears to be conserved in the plant species examined so far (reviewed by Mouradov *et al.*, 2002; Izawa *et al.*, 2003; Hayama and Coupland, 2004). This implies that *Arabidopsis* is a useful model plant and that the knowledge of flowering control in this plant may be applied to other plant species.

This thesis describes the isolation and characterisation of proteins that interact with the protein encoded by the flowering time gene *GIGANTEA* (*GI*), with the aim of providing further insight into the biological function of *GI* and elucidating how *GI* fits into the genetic networks that regulate flowering and plant development. This chapter will review the current knowledge of the control of flowering in *Arabidopsis* covering the induction of flowering and the regulation of flowering by environmental and genetic elements. Following this, the previous identification and isolation of a putative *GI* protein interactor BELL-LIKE HOMEODOMAIN 3 (*BLH3*) (Snowden and Putterill, this laboratory) will be described to put the contents of this thesis in perspective.
1.2 THE TRANSITION TO FLOWERING

Following germination, the majority of plants undergo a period of vegetative growth, termed the juvenile phase. To become competent to flower, the vegetative meristem must progress from the juvenile to adult phase. The reproductive phase occurs in adult vegetative plants in response to environmental and internal stimuli, which induce the shoot apical meristem to undergo the transition from vegetative to floral growth (Figure 1.1) (reviewed by Baurle and Dean, 2006). Some plant species are largely responsive to internal cues, such as the number of vegetative nodes or plant size. In contrast, plants may respond to environmental variables to coordinate flowering to favourable conditions. Environmental signals include photoperiod (day length), light quantity, light quality (spectral composition), vernalisation and nutrient availability.

Figure 1.1 Plant developmental transitions

The first major transition is germination and occurs between embryonic and postembryonic development. The second transition, progression from the juvenile state into the adult state, occurs during the vegetative phase. The third transition is the floral transition, where the plant progresses from the adult vegetative state to the reproductive state. Key signals that regulate the floral transition are shown. From Baurle and Dean (2006).
Of the environmental signals, vernalisation is an important flowering time regulator in many crop species. The prolonged exposure to low temperatures prevents winter annuals from flowering until the arrival of favourable conditions in spring. In biennial plants such as *Daucus carota* (carrot) and *Allium cepa* (onion) vernalisation is necessary for second year flowering. A major difference between annual and biennial cultivars of the *Brassica napus* and *Brassica rapa* species is due to variation in flowering time in response to vernalisation (Osborn *et al.*, 1997).

Plants may variously respond to short day and long day photoperiods to induce flowering. Grafting experiments have determined that day length signals perceived in the leaves can travel to induce flowering at the shoot apex (Bernier *et al.*, 1993). The graft-transmissible day length signal, dubbed florigen, can function interchangeably between short day (SD) and long day (LD) responsive plants and different species (reviewed in Zeevaart, 1976). Although the identity of ‘florigen’ has been elusive, substantial new evidence reveals that key flowering time genes may participate in this signalling mechanism (Chapter 1.4.3.6). For example, a single floral inducer *SINGLE-FLOWER TRUSS* (*SFT*), isolated from day length-neutral tomato, can replace inductive photoperiods and induce flowering in both LD (*Arabidopsis*) and SD (*Nicotiana*) plants (Lifschez *et al.*, 2006). The spectral quality of light has also been shown to be important in floral regulation. Studies with the short day responsive plant *Pharbitis nil* have demonstrated when a normally inductive short day/long night photoperiod is followed by illumination with far red light at the end of the day, flowering is inhibited (Fredericq, 1964).

The environmental cues involved in the regulation of flowering in *Arabidopsis* are described in Chapter 1.3.2 and the genetic pathways that respond to these signals are discussed in detail in Chapter 1.4.

### 1.3 THE CONTROL OF FLOWERING IN *ARABIDOPSIS*

#### 1.3.1 *ARABIDOPSIS* AS A MODEL PLANT

*Arabidopsis* has been chosen as a model system to study many aspects of plant biology (reviewed by Meyerowitz and Somerville, 1994; Meinke *et al.*, 1998). Due to the small plant size, the ability to grow under fluorescent lights, short generation time (approximately 6-8 weeks), the ability to self pollinate and high yield of seed per plant, it
is possible to study large populations and many generations in a relatively short period of time. This is advantageous in a laboratory environment where time and space are limited resources. In addition, the genetic transformation of *Arabidopsis* is relatively straightforward and efficient. The cross-fertilisation and subsequent breeding of plants is simple when required. These factors contribute to the rate and ease with which transgenic plant lines can be generated.

At approximately 125 Mb, the *Arabidopsis* genome is relatively small. Importantly, in 2000 the *Arabidopsis* genome became the first plant genome to be completely sequenced and assembled (Arabidopsis Genome Initiative, 2000). It is predicted that the genome is made up of only ~25 900 genes and the process of decoding the *Arabidopsis* genome is now well underway. Gene knockouts, cDNA collections, microarrays and comparative sequence analysis have been employed to determine the function of *Arabidopsis* genes (reviewed by Borevitz and Ecker, 2004). Extensive community resources are available that aid in many molecular genetic experiments, including a centralised database (www.arabidopsis.org), Expressed Sequence Tag projects, expression databases containing whole genome microarray results and stock centres with seed representing of hundreds of thousands of mutagenised *Arabidopsis* lines.

### 1.3.2 Environmental Signals that Control Flowering in *Arabidopsis*

*Arabidopsis* is an annual plant that has a vegetative phase during which rosette leaves are produced and a reproductive phase which results in inflorescence production. The transition from vegetative growth to flowering occurs when the plant receives suitable cues from the environment and when the plant has reached a certain age at which it is competent to flower (reviewed by Thomas and Vince-Prue, 1997).

The transition to flowering in response to environmental cues is a complex process. It is regulated by numerous signals, such as day length (photoperiod), cold temperature (vernalisation), light quantity and quality. Environmental variables can interact to delay or hasten flowering (reviewed by Martínez-Zapater et al., 1994). The effect of these environmental stimuli on the transition to flowering in *Arabidopsis* is detailed below.

#### 1.3.2.1 Photoperiod
Early flowering ecotypes of *Arabidopsis*, such as Columbia (Col-0) and Landsberg erecta (Ler), originate from temperate regions and exhibit a facultative long day response. These ecotypes flower rapidly in long day photoperiods (LD; 16 h light, 8 h dark), resulting in the promotion of flowering in the spring and summer months. Plants grown in short day photoperiods (SD; 8 h light, 16 h dark) exhibit delayed flowering (Koornneef *et al.*, 1991; Martínez-Zapater *et al.*, 1994), producing a greater number of rosette leaves than those grown in LD conditions, reflecting the repression of flowering in the winter months. However, there is significant variation in flowering-time in response to day length between genetically diverse ecotypes. For example, ecotypes originating from the Isle of Skye, UK (SY-0) and Tenela, Finland (TE-0) do not flower in response to day length (Karlsson *et al.*, 1993). Early day length insensitive flowering in natural populations from the Cape Verde Islands (CVI) and Frankfurt, Germany (FR-2) are associated with allelic variations at photoreceptor loci (Alonso-Blanco *et al.*, 1998; El-Assal *et al.*, 2001; Balasubramanian *et al.*, 2006a).

### 1.3.2.2 Light Quality and Quantity

The transition to flowering is also affected by the spectral quality of light. Flowering time was accelerated in all *Arabidopsis* ecotypes examined by a combination of blue (400-530 nm) and far-red (>700 nm) light (Meijer, 1959), whereas red light has an inhibitory effect on floral induction (reviewed by Martínez-Zapater *et al.*, 1994). Studies in LD plants indicate that both the duration and quality of light treatments can affect flowering (Goto *et al.*, 1991). Light enriched with FR radiation is reflected from green leaves and the perception of this FR light by neighbouring plants provides a signal for potential shading. Exposure of plants to FR enriched light conditions can induce flowering in what is known as a shade avoidance response (Smith and Whitelam, 1997; Cerdan and Chory, 2003).

Flowering in *Arabidopsis* is more sensitive to changes in light duration (photoperiod) than to changes in light intensity, although it has been shown that higher photon flux density can reduce flowering time in some late-flowering *Arabidopsis* mutants (Bagnall, 1993).

### 1.3.2.3 Vernalisation

Prolonged exposure to low temperature is an important regulator of the transition to flowering in many naturally occurring ecotypes of *Arabidopsis*. Vegetative plants can be induced to flower more rapidly after exposure to cold temperatures (below 10 °C) for extended periods. This process of vernalisation is quantitative, as longer cold treatments
are associated with more rapid flowering (reviewed by Henderson and Dean, 2004; Sung and Amasino, 2005). The response to vernalisation can depend on many other variables, including photoperiod conditions, genotype, developmental stage of plant, and the temperature and length of the cold treatment. For example, an ecotype from the Canary Islands (CAN-0) requires vernalisation to flower. In comparison, the CVI ecotype flowers in response to an interaction between vernalisation and photoperiod (Karlsson et al., 1993).

1.3.2.4 GROWTH TEMPERATURE

An increase in growing temperature affects flowering time in Arabidopsis. Studies by Westerman and Lawrence (1971) and Araki and Komeda (1993b) found that higher ambient temperatures significantly reduced flowering time in the ecotypes studied. An increase in night temperature is sufficient to induce earlier flowering in Arabidopsis plants entrained in 12 h light/12 h dark cycles (Thingnaes et al., 2003). The opposite is also true; a decrease in ambient temperature can delay flowering time, even in already late flowering mutant plants (Blazquez et al., 2003).

1.4 THE GENETIC CONTROL OF FLOWERING TIME IN ARABIDOPSIS

The signals that regulate floral development act on genetic pathways to control the time to flowering in Arabidopsis. Subsequently, these pathways regulate floral pathway integrator genes to signal the transition from vegetative to floral state. The integrator genes in turn activate downstream floral meristem identity genes to initiate floral development. Study of mutations and variations in flowering time amongst different ecotypes has lead to the identification of over 80 genes that form the genetic framework for the regulation of flowering in Arabidopsis. Genes involved in the transition to flowering have been placed into model genetic pathways by analysis of the physiology of mutant plants and their response to environmental signals, as well as the phenotype of double and triple mutants (Figure 1.2) (Koornneef et al., 1998b; reviewed by Boss et al., 2004; Jack, 2004; Komeda, 2004). In general, late flowering mutants are thought to carry mutations in genes that are normally involved in promoting flowering, and early flowering mutants are thought to carry mutations in genes that act to suppress flowering (Martínez-Zapater et al., 1994).
Chapter 1

The photoperiodic, light quality, ambient temperature and gibberellin pathways all act to promote flowering in response to environmental and internal cues (Chapter 1.4.1-1.4.5). Pathways that enable flowering are those involved in the vernalisation responsive and autonomous pathways (Chapter 1.4.6). These pathways converge at the floral pathway integrators (Chapter 1.4.7), which function to induce the transition to flowering (Figure 1.2) (reviewed by Boss et al., 2004; Corbesier and Coupland, 2004; Putterill et al., 2004). The key genes that define each pathway and their functions and interactions are described in the sections that follow.

**Figure 1.2** Pathways that control flowering in *Arabidopsis*

The photoperiodic, gibberellin, ambient temperature and light quality pathways promote flowering by initiating floral pathway integrators. The photoperiodic pathway promotes flowering in response to long day lengths. Photoreceptors perceive light and detect day length along with entrainment factors and the circadian clock. Flowering in reaction to light quality can occur dependently and independently to the photoperiodic pathway. The autonomous and vernalisation pathways enable flowering, primarily by removal of the floral repressor *FLC*. In turn, floral pathway integrators activate floral meristem identity genes which initiate floral development. MicroRNAs (miRNAs) (~21-25 bp) function as regulators of gene expression by targeting complementary mRNA and play a novel role in the regulation of flowering time and flower development (reviewed in Jones-Rhoades et al., 2006; Mallory and Vaucheret, 2006). Repressive effects are indicated by and promotive effects are indicated by . Adapted from Boss et al. (2004).

1.4.1 THE AFFECT OF LIGHT ON FLOWERING
Light can act to accelerate or delay flowering in *Arabidopsis*; the opposing function of this signal depends on the quality (or wavelength) of the light. Photoreceptors perceive red (R), far red (FR) and blue light. Red light represses flowering through the action of the light receptor *PHYTOchrome B* (*PHYB*). FR light and blue light, perceived by the *PHYTOchrome A* (*PHYA*), *CRYPTOchrome 1* (*CRY1*) and *CRYPTOchrome 2* (*CRY2*) photoreceptors respectively, promotes flowering (Figure 1.3) (reviewed by Spalding and Folta, 2004). However, recent evidence suggests complex interaction between photoreceptor pathways can affect flowering. *PHYA* and *CRY2* are also thought to act partially redundantly to promote flowering by inhibiting the *PHYB* mediated promotion of flowering (Mockler *et al.*, 1999; 2003) (Figure 1.3).

In a significant series of experiments, Valverde *et al.* (2004) demonstrated that the inhibition of flowering by red light and the promotion of flowering by FR and blue light occurred at the level of the CONSTANS (CO) protein. CO was stabilised in FR and blue light in a process mediated by *PHYA* and *CRY2* (Figure 1.3). This is likely to explain the accelerated flowering seen in plants grown in these light conditions. Additionally, the inhibitory effect of red light was found to be due to the degradation of the CO protein in these conditions (Valverde *et al.*, 2004). The function of CO stability in the photoperiodic pathway to control flowering is discussed in more detail in Chapter 1.4.3.5.

![Figure 1.3 The action of photoreceptors in regulating flowering](image)

The coloured lines represent signal transduction pathways from the photoreceptors. Promotive effects are indicated by and repressive effects by. Adapted from Mockler *et al.* (2003) and Spalding and Folta (2004).

### 1.4.2 THE LIGHT QUALITY PATHWAY

An increase in shaded light is defined as a decrease in the ratio of red to far red (R:FR) light, often due to the reflection of light from green leaf tissue in crowded plant populations. The detection of this quality of light triggers a shade avoidance response in
plants, which ultimately results in the acceleration of flowering. This is thought to be mediated partly by \emph{PHYB}, via the \emph{PHOTTOCHROME AND FLOWERING TIME 1 (PFT1)} gene (Halliday \emph{et al.}, 1994; Cerdan and Chory, 2003). The \textit{phyB phyD} and \textit{phyB phyE} double mutants displayed a reduced response to low R:FR light compared to the \textit{phyB} mutant alone, indicating that \textit{PHYD} and \textit{PHYE} are also involved in the detection of this quality of light (Devlin \emph{et al.}, 1998; 1999). Supporting evidence presented by Franklin \emph{et al.} (2003) demonstrated that \textit{PHYB}, \textit{PHYD} and \textit{PHYE} act redundantly in response to low R:FR light. The shade avoidance response is proposed to be under the control of an internal circadian clock, as plants respond to a low R:FR light pulse most effectively at the end of the day (Salter \emph{et al.}, 2003).

To identify novel genes involved in the shade avoidance response, microarray technology was employed to detect global changes in transcription in response to low R:FR light (Devlin \emph{et al.}, 2003). Over 300 shade responsive genes were identified, some from floral induction pathways; these included \textit{GI} and \textit{FLOWERING LOCUS T (FT)}. This observation may provide a link between the detection of light quality and flowering time.

\subsection*{1.4.3 The Photoperiodic Pathway and the Circadian Clock}

Plants are able to detect and respond to differing day lengths in a developmental process defined as the photoperiodic control of flowering. This allows plants to adapt to seasonal changes and therefore reproduce at the most favourable time of the year (reviewed by Thomas and Vince-Prue, 1997; Yanovsky and Kay, 2003; Corbesier and Coupland, 2004). The circadian clock is an internal time keeper that enables biological organisms to synchronise developmental processes with external cues such as light/dark and temperature cycles. Evidence presented by Dodd \emph{et al.} (2005) show that plants which sustain a circadian clock that best matches environmental cues have an increased rate of survival and superior productivity and thus have a distinct advantage over competitors. The period length of the circadian clock varies between \textit{Arabidopsis} ecotypes and this has been shown to correlate with day length in the different environments in which the ecotypes evolved (Michael \emph{et al.}, 2003a). The detection of different day lengths is a key process in the control of photoperiodic flowering. Recently, strong physiological and genetic evidence suggests that the circadian clock plays a role in the measurement of day length.
Physiological experiments show that flowering is promoted in LD plants when extended dark periods are interrupted with light applications (light breaks) in a process that mimics long days by creating a short night. The response of the plant to light breaks changes through the daily cycle, as only applications of light at particular times of the day affect flowering. It is proposed that the circadian clock is the timing mechanism that drives this rhythmic sensitivity to light. The external coincidence model of photoperiodism hypothesises that the coincidence of an external signal (light) with an internal light-responsive phase of a circadian rhythm will act to promote flowering in inductive conditions (Thomas and Vince-Prue, 1997). In this model the role of light is twofold; first, it acts on the circadian clock to reset rhythms, which in turn controls the output of circadian regulated genes; secondly, light acts on a circadian regulated gene/molecule at a particular point of the daily cycle to promote flowering. The measurement of day length and subsequent promotion of flowering occurs only when light and the regulatory gene/molecule overlap at a specific point in the circadian cycle (reviewed by Yanovsky and Kay, 2003). Significant evidence supporting the function of the external coincidence model in the control of photoperiodic flowering in *Arabidopsis* has been obtained with work on a key floral promoter CO. This is described in Chapter 1.4.3.5.

### 1.4.3.1 THE ARABIDOPSIS CIRCADIAN CLOCK

A simplified model of the circadian clock can be described by three components (Figure 1.4). First, **input** pathways, which act to synchronise the circadian clock to the environment. Second, the **central oscillator**, comprised of components that are central to regulating rhythmic behaviour. All oscillators studied to date rely on feedback systems consisting of positive and negative regulatory elements, which act to induce and block transcription of clock-genes respectively (reviewed by Dunlap, 1999). Finally, **output** pathways are thought to connect the 'time' generated by the clock to the organism, thus coordinating the appropriate behavioural response.

Many publications have described the isolation and characterisation of genes involved in clock-controlled processes in the model plant *Arabidopsis* (reviewed by Mouradov *et al.*, 2002; Millar, 2003; Salome and McClung, 2005a). Recently, large-scale gene expression analyses and screening for altered circadian gene expression have proved useful for the identification of more clock-regulated and clock-controlling genes (Harmer *et al.*, 2000; Schaffer *et al.* 2001; Michael and McClung, 2003; Onai *et al.*, 2004; Hazen *et al.*, 2005;
Darrah et al., 2006). The identification and investigation of novel genes will improve the understanding of how the circadian clock is regulated in Arabidopsis.

**Figure 1.4** Simplified model of the Arabidopsis circadian clock

Light input to the clock is perceived by photoreceptors. The central clock oscillator consists of a feedback loop represented by the black circular arrows, entrainment factors are indicated by the grey arrow. The output of the oscillator includes clock regulated genes. Adapted from Yanovsky and Kay (2003).

### 1.4.3.2 The central clock oscillator

The putative central clock oscillator is proposed to consist of a negative feedback loop comprised of three genes, CIRCADIAN CLOCK ASSOCIATED 1 (CCA1), LATE ELONGATED HYOCOTYL (LHY) and TIMING OF CAB EXPRESSION 1 (TOC1) (Figure 1.4). CCA1 and LHY are partially redundant clock regulated genes with a peak of expression in the morning. Both encode similar MYB proteins and act to down regulate each others expression when over expressed (Schaffer et al., 1998; Wang and Tobin, 1998; Green and Tobin, 1999; Mizoguchi et al., 2002). The transcription of CCA1 and LHY is induced by TOC1 during the night, reaching a peak at dawn. CCA1 and LHY act directly to repress TOC1 by binding an element in the TOC1 promoter. Subsequently, CCA1 and LHY levels decline due to the decrease of TOC1 (Alabadi et al., 2001).

The expression of TOC1 is also circadian controlled, with a peak in the evening that is opposite to that of CCA1/LHY (Strayer et al., 2000; Matsushika et al., 2000). Both increasing and decreasing the levels of TOC1 can disrupt circadian rhythms in various light
conditions, leading to the proposal that TOC1 integrates light signalling to the clock (Mas et al., 2003a). The involvement of TOC1 in a regulatory feedback loop is highlighted by the fact that over-expression of TOC1 alters CCA1 and LHY expression. Likewise, over-expression of CCA1 disrupts the rhythmic expression of TOC1 (Makino et al., 2002; Matsushika et al., 2002). The toc1-1 mutant is early flowering and has shortened circadian rhythms of 21 h. The early flowering phenotype is overcome by growing plants under shortened 21 h day lengths, which match the endogenous rhythm of the mutant. Importantly, these results reveal that TOC1 regulation of circadian period is important for the measurement of day length (Somers et al., 1998b).

Modelling of the Arabidopsis circadian clock by Locke et al. (2005) suggested that the current circadian clock model was unable to explain the experimental data. The revised model is comprised of an interlocked feedback loop network, which is proposed to contain a fourth gene between CCA1/LHY and TOC1. EARLY FLOWERING 4 (ELF4) is a circadian regulated gene that is proposed to act as part of the feedback loop to control the circadian clock (Doyle et al., 2002) (Figure 1.4). The light induced expression of ELF4 is similar to TOC1 and peaks in the evening. In parallel with TOC1, ELF4 is required for the expression of CCA1/LHY and conversely CCA1/LHY negatively regulate ELF4 (Kikis et al., 2005). Another clock candidate is a gene encoding a new MYB protein LUX ARRhythmo (LUX) (also known as PHYTOCLOCK1 [PCL1]). Significantly, CCA1 and LHY bind to the evening element in the LUX promoter, suggesting these genes regulate LUX as they do clock oscillator component TOC1 (Hazen et al., 2005; Onai and Ishiura, 2005).

The interlocked feedback loop model includes a second feedback loop that is proposed to function via GI (Locke et al., 2005). Although GI has traditionally been placed in the photoperiodic pathway that controls flowering time (Chapter 1.4.3.5), GI expression is both clock-controlled and acts as part of a feedback mechanism that affects circadian rhythms (Fowler et al., 1999; Park et al., 1999). The function of GI in the circadian clock is discussed in more detail later (Chapter 1.5.3).

1.4.3.3 LIGHT INPUT AND THE ENTRAINMENT OF THE CLOCK

In order to synchronise rhythmic behaviour with light cues, plants must detect and quantify light signals. This light information is used to adjust (or entrain) the internal circadian clock to keep pace with the outside world. Some of the molecular components involved in
light perception and signalling to the clock have been characterised and it is thought that control of the pace of the circadian clock (period) is entrained by light via multiple photoreceptors (Figure 1.4) (reviewed by Fankhauser and Staiger, 2002; Millar, 2003; Salome and McClung, 2005a).

The Arabidopsis photoreceptor genes which perceive light quality and quantity also regulate light input to the circadian clock. Plant cryptochromes are evolutionarily distinct from the cryptochromes that are central regulators of the animal circadian clock (reviewed by Cashmore et al., 1999). The cry1 cry2 double mutants show robust circadian cycling, indicating that these genes do not function in the central oscillator like cryptochromes in animal circadian systems (Devlin and Kay, 2000). Detailed studies have established that CRY1, CRY2 and PHYA are involved in transmitting blue-light to the circadian clock (Somers et al., 1998a; Devlin and Kay, 2000; Yanovsky et al., 2001). Red light signalling to the clock is mediated redundantly by PHYA and PHYB, with a role for both CRY1 and CRY2 (Somers et al., 2000, Devlin and Kay, 2000; Salome et al., 2002). PHYA may be involved in resetting the circadian clock, as circadian leaf movement can be re-phased by exposure to far red light in Arabidopsis wild-type, but not in phyA mutants (Yanovsky et al., 2000a). Further analyses of triple and quadruple mutants are suggestive of a role of other photoreceptors such as PHYD and PHYE in light signalling to the clock (Devlin and Kay, 2000; Yanovsky et al., 2000b; Mazzella et al., 2001).

The PHYA and PHYB phytochromes interact with members of the basic/helix-loop-helix (bHLH) transcription factor family. The first bHLH protein shown to interact with phytochromes was PHYTOCHROME INTERACTING FACTOR 3 (PIF3) (Ni et al., 1999). It was also established that PIF3 binds the G box motifs within the promoters of light-regulated genes such as CCA1 and interacts with TOC1 in vitro (Martínez-Garcia et al., 2000; Makino et al., 2002). This provides a possible link between light-responsive phytochromes and the circadian clock. Although PIF3 is not essential for light signalling, a number of evolutionarily related bHLH proteins have been shown to interact with phytochromes, consequently redundancy in function between members of this family is possible (reviewed by Salome and McClung, 2005a). The bHLH transcription factor PIF4 binds preferentially to PHYB (Huq and Quail, 2002). A PIF4 protein that contained an altered protein binding domain was unable to bind PHYB, nor complement the pif4 photomorphogenic mutant when expressed in plants. This is indicative of a functional role of the PIF4:PHYB protein interaction in light signalling (Khanna et al., 2004). A number
Chapter 1

of the PHYTOCHROME INTERACTING FACTORS (PIFs) are circadian regulated and it is thought that these proteins act largely as positive regulators of phytochrome signalling (reviewed by Duek and Fankhauser, 2005; Spalding and Folta, 2004).

The ZEITLUPE (ZTL) and LOV KELCH PROTEIN 2 (LKP2) genes are proposed to function as additional circadian photoreceptors (Figure 1.4). These genes belong to a small family of proteins that contain three conserved domains; the F-box, the Kelch protein-interaction domain and the light sensitive flavin-binding LOV domain (Kiyosue and Wada, 2000; Nelson et al., 2000; Somers et al., 2000; Schultz et al., 2001). Although the LOV domain is found in other photoreceptors in Arabidopsis and Neurospora, it has yet to be determined if the ZTL and LKP2 proteins act as photoreceptors in plants. Altered expression of ZTL and LKP2 can repress flowering and cause arrhythmic expression of clock controlled genes (Schultz et al., 2001; Somers et al., 2004; Kim et al., 2005a), mutant phenotypes which are common to genes that function in the circadian clock. A ztl mutant which contains a single amino acid change within the LOV domain causes an altered circadian period without affecting flowering time (Kevei et al., 2006). This separates the circadian and developmental roles of ZTL and suggests that the LOV domain is important in circadian function. Protein interactions between ZTL and the PHYB and CRY1 photoreceptors have been described by Jarillo et al. (2001), providing an additional link between light input and the circadian clock. In a regulatory feedback mechanism, the ZTL protein is positively regulated in the light and negatively regulated in the dark in a process mediated by the circadian clock (Kim et al., 2003). ZTL is required for the dark-dependent degradation of the clock component TOC1 (Mas et al., 2003b), adding a further level of feedback regulation to this circadian system.

Mutations in the EARLY FLOWERING 3 (ELF3) gene cause arrhythmia and early flowering, therefore it is hypothesised that ELF3 also functions in the control of the circadian clock. The circadian phenotype of elf3 is dependent on light, signifying ELF3 may regulate light signalling to the clock (Zagotta et al., 1996; Hicks et al., 1996; 2001; McWatters et al., 2000). Analysis of elf3 ztl double mutants verifies that ELF3 and ZTL function via different mechanisms to regulate the circadian clock (Kim et al., 2005a). As ELF3 is implicated in resetting the clock (Covington et al., 2001), it has been proposed that ELF3 acts as gate keeper of the clock, preventing arrhythmicity when the circadian system is perturbed by pulses of light (reviewed by Salome and McClung, 2005a).
1.4.3.4 ENTRAINMENT OF THE CLOCK TO ALTERNATIVE ENVIRONMENTAL SIGNALS

The plant circadian clock can be entrained to the environment by temperature signals. Michael et al. (2003b) identified two independent circadian clocks in *Arabidopsis*, which were able to be distinguished by their differing sensitivity to temperature cues. Natural variation in the temperature compensation of the clock is evident in different *Arabidopsis* ecotypes. Quantitative trait loci (QTL) were mapped using this variation by Edwards et al. (2005). Key clock component ZTL was identified, which is likely to function in both the light and temperature input pathways to regulate the circadian clock (Figure 1.4). Gould et al. (2006) have found that TOC1 expression increases in parallel with increasing temperature. As ZTL is crucial for the degradation of the oscillator component TOC1 in the dark, it is tempting to suggest that TOC1 may be a point of convergence of both light and temperature inputs to the clock. In addition, two TOC1 related genes, *PSEUDO-RESPONSE REGULATOR 7* (PRR7) and PRR9 act partially redundantly to reset the clock in response to temperature cues (Salome and McClung, 2005b).

The flowering regulators GI and *FLOWERING LOCUS C* (FLC) were also identified as candidates for the temperature compensation QTL (Edwards et al., 2005). Both FLC and vernalisation can affect the periodicity of the circadian clock (Salathia et al., 2006). Comparison of the fri flc and FRI FLC plants suggests FLC acts at 27°C to control the circadian period in response to temperature (Edwards et al., 2006). Over 1000 FLC responsive genes were identified, including LUX, a candidate circadian oscillator gene (Edwards et al., 2006) (Chapter 1.4.3.2). GI was also identified as an essential component of the temperature compensation mechanism to the circadian clock (Gould et al., 2006). The role of GI and the circadian clock will be discussed later (Chapter 1.5.3).

Circadian clock-controlled and clock-regulating genes such as *CCA1*, *LHY* and GI have been identified in the stress-inducible plant *Mesembryanthemum crystallinum* (Common Ice Plant). The affect of environmental signals on the circadian clock were also studied in this plant. The expression of central oscillator genes *CCA1* and *LHY* cycled robustly, even after severe salt stress. This is suggestive that the circadian clock compensated against salt stress, indicating that plant circadian systems can respond to abiotic stresses as well as light and temperature cues (Boxall et al., 2005).
1.4.3.5 THE PHOTOPERIODIC PATHWAY

Mutations that result in a late-flowering phenotype in plants grown under LD conditions have been traditionally placed in the photoperiodic pathway. These genes are thought to promote flowering predominantly in LD conditions and include *GI*, *CO* and *FLOWERING LOCUS T (FT)* (Koornneef *et al.*, 1991; reviewed by Corbesier and Coupland, 2004; Putterill *et al.*, 2004). The focus of this thesis research is *GI*, a circadian regulated gene that acts to promote flowering (Fowler *et al.*, 1999). The role of *GI* will be discussed in detail in Chapter 1.5. Genes that carry mutations that cause day length-insensitive flowering have also been placed in the photoperiodic pathway. Mutations in a number of these genes affect the circadian clock, including *ELF3*, *ELF4* and *CCA1*. The role of these genes was discussed in Chapter 1.4.3.2 and Chapter 1.4.3.3.

The photoreceptor genes which act in light perception and signalling play a significant role in day length measurement and the control of photoperiodic flowering (reviewed by Spalding and Folta, 2004). The photoreceptor *PHYA* is generally thought to promote flowering, and has been implicated in sensing extended light periods in day length extension experiments. Non-inductive short days (8 h) of white light are extended with 8 h of low fluence non-photosynthetically active radiation and in these extended LD conditions *phyA* mutants flower later than wild-type (Johnson *et al.*, 1994; Mazzella *et al.*, 2001). *PHYB* has a complicated role in regulating flowering; over-expressors of *PHYB* and the *PHYB* mutant long hypocotyl 3 (*hy3*) both exhibit early flowering, indicating that plants are sensitive to the balance of *PHYB* (Somers *et al.*, 1991; Bagnall *et al.*, 1995). Expression analyses have shown that *PHYB* in the leaf mesophyll represses expression of the central floral promoter *FT*, thereby inhibiting flowering (Endo *et al.*, 2005). The recent isolation of *phyC* mutants have provided evidence that this phytochrome also regulates flowering in response to day length and may act to promote flowering in the absence of *PHYA* (Monte *et al.*, 2003).

The classical late flowering mutant *fha* is allelic to *CRY2* and has traditionally been placed in the photoperiodic pathway. *CRY2* proteins are stabilised in LD and decrease during the light in SD, indicating this photoreceptor is directly affected by photoperiod (Mockler *et al.*, 2003). A screen of plants that flowered early in non-inductive SD uncovered a novel *CRY2* allele (*CRY2-Cvi*), which contained a single amino acid substitution. The early flowering phenotype was associated with reduced down regulation of *CRY2* in SD in
response to light. These results are suggestive that light regulation of \textit{CRY2} in different photoperiods is important for the perception of day length and control of flowering. \textit{CRY2} is likely to act upstream of photoperiodic regulators \textit{CO} and \textit{GI}, as the promotion of flowering by \textit{CRY2-Cvi} requires the products of these genes (El-Assal \textit{et al.}, 2001; 2003).

A central regulator of flowering is the \textit{CO} gene, which encodes a putative zinc finger transcription factor containing a conserved domain thought to mediate protein-protein interactions (Putterill \textit{et al.}, 1995; Robson \textit{et al.}, 2001). \textit{CO} is a positive regulator of flowering; over expression results in early day length-insensitive flowering and the loss of \textit{CO} function causes late flowering in LD (Simon \textit{et al.}, 1996; Samach \textit{et al.}, 2000). \textit{CO} is likely to function downstream of the circadian clock and the \textit{CRY2} photoreceptor to control flowering as the over expression of \textit{CO} can complement the \textit{gi-2, fha-1} and \textit{lhy} late flowering phenotypes (Suarez-Lopez \textit{et al.}, 2001).

In order to measure and respond to day length changes, light is first detected by photoreceptors. The length of the light or dark period is then measured by an internal time-keeper, the circadian clock. The current hypothesis for day length perception and the photoperiodic control of flowering is based on the external coincidence model (Chapter 1.4.3). This model proposes that the timing of \textit{CO} expression to coincide with light is essential for the accurate measurement of day length and regulation of photoperiodic flowering (Figure 1.5) (reviewed by Putterill \textit{et al.}, 2004).

Strong evidence supporting this model is provided at the level of \textit{CO} in \textit{Arabidopsis}. The \textit{CO} transcript is regulated by the circadian clock, with a peak 16-20 h after dawn that is disrupted in flowering time and clock mutants. It has been demonstrated that the phase of \textit{CO} expression is crucial in floral promotion and is dependent on the circadian clock (Roden \textit{et al.}, 2002). \textit{CO} expression occurs chiefly in the dark in non-inductive SD conditions (Suarez-Lopez \textit{et al.}, 2001), hence is not able to promote flowering. In contrast, the coincidence of \textit{CO} expression with the light in LD is the penultimate step in the perception of day length and the promotion of flowering (Yanovsky and Kay, 2002) (Figure 1.5). The recent discovery that the CO protein is stabilised in the light by a mechanism mediated by photoreceptors \textit{PHYA} and \textit{CRY2} and degraded in the dark via the 26S proteasome supports the external coincidence model for photoperiodism (Valverde \textit{et al.}, 2004).
A novel family of four WD-domain proteins SUPPRESSOR OF PHYA 1-4 (SPA1-4) were previously identified as suppressors of photomorphogenesis in darkness (Laubinger et al., 2004). It is proposed that this family of proteins regulates the dark specific degradation of CO to control flowering (Laubinger et al., 2006). The spa1spa3spa4 triple mutant flowers early in SD and this is associated with increased levels of the CO protein. The SPA and CO proteins interact \textit{in vitro} and \textit{in vivo} and \textit{SPA} transcripts increase during the night phase, at the time when CO is actively degraded.

\textit{FLAVIN-BINDING KELCH REPEAT F-BOX 1 (FKF1)} is a circadian clock controlled gene that encodes a putative novel blue light photoreceptor that is proposed to function in the photoperiodic pathway upstream of \textit{CO} (Nelson et al., 2000; Imaizumi et al., 2003). \textit{FKF1} belongs to a small family which includes the putative clock photoreceptors \textit{ZTL} and \textit{LKP2} genes (Chapter 1.4.3.3). Recently, it has been shown that FKF1 interacts with and destabilises a transcription factor CYCLING DOF FACTOR 1 (CDF1). \textit{CDF1} is a floral repressor; over expression results in low levels of \textit{CO} and consequently late flowering plants. This repression of flowering is proposed to occur via direct down regulation of \textit{CO} as \textit{CDF1} binds to the \textit{CO} promoter (Imaizumi et al., 2005).
FT was initially placed in the photoperiodic pathway as plants carrying mutations in the FT gene are insensitive to day length (Koornneef et al., 1991). The CO transcript is unchanged in the ft mutant and FT is up-regulated in response to CO, signifying that CO functions upstream of FT (Samach et al., 2000; Suarez-Lopez et al., 2001). The over expression of FT causes extremely early flowering and is able to suppress the late flowering co phenotype (Kardailsky et al., 1999). Although FT is the major output of CO, FT is proposed to function as a floral pathway integrator, acting independently of CO and the photoperiodic pathway to affect flowering (see Chapter 1.4.7). This is supported by evidence that FT is eventually up-regulated in co mutants and 35S:CO ft plants still flower earlier than wild type despite the lack of functional FT (Kobayashi et al., 1999; Yoo et al., 2005). Investigation of the role of FT in Populus trees has surprisingly shown that FT mediates flowering, growth cessation and bud set in response to day length (Böhlenius et al., 2006). It is possible that FT has a more general role in day length perception and response than was previously thought.

1.4.3.6 ‘FLORIGEN’ AND THE PHOTOPERIODIC SIGNAL

As perception of day length occurs in the leaves and floral development is induced at the SAM, flowering must be initiated by a long distance signal. The biochemical nature of the signalling molecules involved has long been elusive. Recently, the flowering time genes CO and FT have been implicated in this long distance signalling mechanism (reviewed in Baurle and Dean, 2006). Leaf-specific expression of CO is sufficient to promote flowering and this signal is graft transmissible (Ayre and Turgeon, 2004). Furthermore, a combination of grafting and mis-expression experiments have demonstrated that CO acts in the phloem to promote flowering and that this function is partly dependent on FT (An et al., 2004). FT expression is up-regulated via CO in the leaves in response to day length and in contrast to CO, FT is able to act non-cell autonomously to activate flowering. Furthermore, induction of the FT transcript in a single leaf is able to trigger flowering, even in non-inductive SD conditions (Huang et al., 2005).

The bZIP transcription factor FD is present at the shoot apex before flowering and is required for floral promotion by FT. The FT signal travels through the vascular tissue to the shoot apex where interaction with FD occurs to induce transcription of floral meristem identity genes (Abe et al., 2005; Wigge et al., 2005). These results led to the hypothesis that FT is a component of the mobile flowering signal. Although FT fulfils many of the
criteria of the hypothetical ‘florigen’ molecule, it can not be resolved from these results if the $FT$ RNA or protein is the long distance mobile signal that activates flowering or if intermediaries are involved.

1.4.4 **THE AMBIENT TEMPERATURE PATHWAY**

The relationship between flowering time and temperature is complex, although in general plants flower more rapidly in response to higher ambient temperatures (Chapter 1.3.2.4). Balasubramanian et al. (2006b) have demonstrated that an apparently small increase in temperature, from 23°C to 27°C, can induce flowering in *Arabidopsis* plants grown in SD conditions as effectively as inductive LD photoperiods. Interestingly, the phytochromes which mediate light perception and the regulation of flowering time are also affected by changes in temperature. Analysis of the *phyA cry2* double mutant suggests that *CRY2* activity is important at 16°C, whereas *PHYA* activity is lost at this lower temperature. Conversely *PHYA* acts with *CRY2* to control flowering at 23 °C (Blazquez et al., 2003). In an independent study of phytochrome mutants, Halliday and Whitelam (2003) found that *PHYB* and *PHYE* were important for the regulation of flowering at 22°C and 16°C respectively.

The analysis of expression profiles of the *Arabidopsis* transcriptome in response to photoperiod and temperature cues indicate that the pathways activated by these signals are not equivalent (Balasubramanian et al., 2006b). The temperature dependent mechanism for flowering control does not require $CO$ and is proposed to converge at the floral integrator $FT$ (Blazquez et al., 2003; Halliday et al., 2003, Balasubramanian et al., 2006b).

1.4.5 **THE GIBBERELLIN PATHWAY**

It has been hypothesised that the gibberellin responsive pathway is important for the induction of flowering in the absence of the promotional effects of long day photoperiods (Reeves and Coupland, 2001). Plants that are insensitive to gibberellic acid (GA), or are deficient in GA biosynthesis, flower late in SD. In the most severe case, *gal* mutants that are defective in an early step of GA biosynthesis are unable to flower in SD (Wilson et al., 1992). Analyses of double mutants have demonstrated that the GA responsive pathway acts independently of both the photoperiodic and vernalisation pathways (reviewed by Boss et al., 2004).
It has recently become apparent that the regulation of gene expression by miRNAs has a role in floral induction (reviewed by Mallory and Vaucheret, 2006). The over expression of microRNA159 (miR159) resulted in a late flowering phenotype in SD. This was associated with a reduction in LFY expression. Furthermore, miR159 levels were regulated by gibberellin pathway genes GA and GAI, signifying that miR159 functions downstream of these gibberellin pathway genes to affect flowering (Achard et al., 2004).

1.4.6 Pathways that converge at FLC

Genetic pathways that enable flowering do so via the control of floral repressors, which act to maintain the plant in a vegetative state. The transition to flowering occurs when the activities of the floral repressors are reduced or overcome by external and/or endogenous signals, such as vernalisation or the developmental age of the plant (reviewed by Boss et al., 2004; Henderson and Dean, 2004). The best characterised of these genes are classified within the autonomous and vernalisation pathways, which converge on a central floral repressor, FLOWERING LOCUS C (FLC) (Figure 1.2). The FLC gene encodes a MADS box transcription factor and high levels of expression are associated with late flowering in vernalisation responsive ecotypes and mutants (Michaels and Amasino, 1999; Sheldon et al., 2000; Michaels and Amasino, 2001). Experimental results revealed that FLC suppresses flowering via repression of floral integrators FT, LFY and SUPPRESSOR OF OVEREXPRESSION OF CONSTANS (SOC1) (Nilsson et al., 1998; Samach et al., 2000; Hepworth et al., 2002). FLC functions as part of a large protein complex to repress FT and SOC1 (Helliwell et al., 2006; Searle et al., 2006). The inhibition of SOC1 and FD, the target of FT, by FLC occurs in the shoot meristem, thus reducing the ability of the shoot meristem to respond to the inductive FT signal (Searle et al., 2006).

The expression of FLC is up-regulated by FRIGIDA (FRI) and allelic variation within the FRI and FLC loci are major contributors to the naturally occurring flowering variation in many Arabidopsis ecotypes (reviewed in Henderson and Dean, 2004; Sung and Amasino, 2005). Other positive regulators of FLC that have not been placed into genetic pathways include the FRIGIDA-LIKE (FRL) and VERNALISATION INDEPENDENCE (VIP) genes, which are necessary for high levels of FLC expression (Zhang et al., 2003; Michaels et al., 2004).
1.4.6.1 THE AUTONOMOUS PATHWAY

The autonomous pathway is so called as these genes do not respond to external cues such as day length or cold treatment to promote flowering. The six genes (**FCA**, **FLOWERING LOCUS D** (**FLD**), **FLOWERING LOCUS K** (**FLK**), **FPA**, **FY** and **LUMINDEPENDENS** (**LD**)) target **FLC** and result in late flowering in all photoperiods when mutated (reviewed by Henderson and Dean, 2004). The late flowering phenotype of the autonomous pathway mutants is overcome with the loss of **FLC** function, suggesting these genes regulate flowering through **FLC** (Michaels and Amasino, 2001).

1.4.6.2 THE VERNALISATION PATHWAY

Genes placed in the vernalisation pathway negatively regulate **FLC** in reaction to cold treatment. **FLC** transcript levels are reduced in response to cold and are maintained at low levels through cell division cycles until flowering. This memory of vernalisation is thought to be facilitated by epigenetic silencing of **FLC** (Bastow et al., 2004). The epigenetic changes at the **FLC** locus are proposed to occur via the function of the **VERNALISATION 1** (**VRN1**) and **VRN2** genes. In **vrn1** and **vrn2** mutants, **FLC** levels are reduced in response to cold, however suppression of **FLC** is unable to be maintained. This indicates that the **VRN** genes do not function in cold perception, but in the maintenance of the vernalised state (Gendall et al., 2001; Levy et al., 2002; reviewed in Sung and Amasino, 2005). Recently, the maintenance of dimethylation of **FLC** was shown to rely on functional LIKE HETEROCHROMATIN PROTEIN1 (Sung et al., 2006). **VERNALISATION INSENSITIVE 3** (**VIN3**) is up-regulated by cold and is required for the down regulation of **FLC** associated with the vernalisation. Intriguingly, the lack of repression of **FLC** in **vin3** mutants is linked to a lack of histone H3 modification, which is central in the epigenetic control of **FLC** in response to vernalisation (Sung and Amasino, 2004; Bastow et al., 2004).

1.4.7 FLORAL PATHWAY INTEGRATORS

Flowering pathways must converge at some point as all pathways induce floral development by activating floral meristem identity genes. Evidence suggests that genes directly downstream of **CO** may be the points of convergence and interaction of pathways that control flowering. Molecular and genetic studies have identified targets of **CO**. These include flowering-time genes and proposed floral integrators **FT** and **SUPPRESSOR OF OVEREXPRESSION OF CO 1** (**SOC1**), which are up-regulated in response to **CO** (Onouchi et al., 2000; Samach et al., 2000) (Figure 1.2). **SOC1** (allelic to **AGAMOUS**-
LIKE 20) encodes a MADS-domain gene that has a late flowering phenotype in loss of function mutants, indicating a function in the promotion of flowering.

It is hypothesised that parallel pathways via FT and SOC1 act downstream of CO to control flowering. The photoperiodic pathway is constitutively active in transgenic plants expressing 35S:CO, causing these plants to flower early. The early flowering associated with over expression of CO is only partially overcome in 35S:CO soc1 plants, implying CO can act independently of SOC1 to promote flowering (Yoo et al., 2005). Furthermore, the double mutants ft soc1 and fwa soc1 cause a greater delay in the flowering of 35S:CO transgenic plants than do single mutants alone (Onouchi et al., 2000; Samach et al., 2000). Expression analyses detected a further level of complexity between FT and SOC1, as SOC1 expression is suppressed in the ft mutant, even in the 35S:CO background (Schmid et al., 2003; Yoo et al., 2005). These results imply that CO regulates SOC1 via FT.

The central floral repressor FLC can repress the effects of the photoperiodic pathway. High levels of FLC down-regulate the photoreceptor CRY2, which is essential for the promotion of CO in LD (El-Assal et al., 2003). FLC binds directly to regulatory elements within the SOC1 and FT promoters and this interaction mediates the repression of these genes in plants (Hepworth et al., 2002; Rouse et al., 2002; Helliwell et al., 2006; Searle et al., 2006). Over expression of SOC1 can complement the late flowering phenotype of FRI FLC, signifying that SOC1 also functions downstream of the autonomous pathway (Borner et al., 2000; Lee et al., 2000). Full activation of FT and SOC1 by CO is not observed in fca mutants, indicating functional FCA is required for regulation of these genes (Samach et al., 2000). FCA is involved in regulating the autonomous pathway via the floral repressor FLC; therefore FT and SOC1 are points of convergence of both the autonomous and photoperiodic pathways (Figure 1.2). The integration of the GA-responsive pathway also occurs at SOC1, but not at FT. The expression of SOC1 is positively regulated by GAs and soc1 mutants are less responsive to GA treatment (Borner et al., 2000; Moon et al., 2003). Therefore, SOC1 is a point of convergence of at least three genetic pathways.

The vernalisation and photoperiodic pathways also converge at FT. The VRN1 gene is implicated in the regulation of flowering in response to cold treatments and is proposed to act in the vernalisation pathway (Gendall et al., 2001). In plants over expressing VRN1, FT was up-regulated and this was associated with early flowering, irrespective of cold treatment (Levy et al., 2002). In contrast, TERMINAL FLOWER 2 (TFL2) encodes a
putative floral repressor. Analysis of the expression of floral integrator genes demonstrates that \( FT \) is up-regulated in early flowering \( tfl2 \) mutants, but no change in \( CO \) or \( S OCI \) levels was detected (Kotake et al., 2003). \( TFL2 \) is able to repress \( FT \) expression in leaf tissue, even in the presence of \( CO \) (Takada and Goto, 2003). It is likely that the \( FT \) gene is the site of convergence of opposing promotive and repressive elements in the control of flowering.

The interactions between flowering-time genes and floral meristem identity genes have been highlighted by analyses of \( LFY \) expression. \( LFY \) encodes a transcription factor that is induced in the shoot apical meristem (SAM) to initiate flowering (Weigel et al., 1992; Wagner et al., 1999). Analyses of the effect of over-expression of \( LFY \) and the effect of late-flowering mutations on \( LFY \) expression indicate that \( LFY \) is the target for genes from distinct flowering-time pathways (Nilsson et al., 1998; Simon et al., 1996) (Figure 1.2). \( LFY \) is up-regulated by the photoperiodic pathway in response to \( CO \) expression (Simon et al., 1996; Blázquez and Weigel, 2000; Samach et al., 2000). Using a reporter construct consisting of the \( LFY \) promoter fused to a \( \beta\)-glucuronidase reporter gene, Blázquez and Weigel (1999) demonstrated that the \( LFY \) promoter is also activated by gibberellins. Analysis of a series of deletions of the \( LFY \) promoter demonstrated that the points of convergence of the photoperiodic and GA-responsive pathways are at separate elements of the \( LFY \) promoter (Blázquez and Weigel, 2000).

Although the above experiments show that the photoperiodic pathway can converge at \( LFY \), flowering in response to day length is also be induced via a pathway parallel to \( LFY \). When the photoperiodic pathway is constitutively activated by over-expression of \( FT \), the \( LFY \) reporter construct is not activated, therefore \( LFY \) transcription is not required for the promotion of flowering by \( FT \). \( LFY \) is also up-regulated in early flowering \( phyB \) mutants and this effect is independent of \( CO \) and \( FT \) (Blázquez and Weigel, 2000). Furthermore, \( 35S:LFY \) is unable to correct the late-flowering phenotypes of \( ft \), \( fe \) and \( fwa \), mutants, suggesting these genes act to promote flowering independent of \( LFY \) (Nilsson et al., 1998). This is supported by analysis of \( ft \) \( lfy \) and \( fwa \) \( lfy \) double mutants, which suggest that \( FT \) and \( FWA \) act independently, possibly redundantly, to \( LFY \) (Ruiz-Garcia et al., 1997). Interestingly, mining of expression profiles show that \( LFY \) expression is affected by the \( ft \) mutation (Schmid et al., 2003). Thus, it is likely that \( FT \) and \( LFY \) do not function entirely in parallel to promote flowering and that some cross regulation between these floral integrator genes occurs.
1.5 **THE GIGANTEA GENE**

1.5.1 **THE PLEIOTROPIC EFFECTS OF THE GI MUTATION**

Rédei (1962) identified the first gi mutant due to its late-flowering phenotype. In general, gi mutants exhibit delayed flowering and undergo a prolonged phase of vegetative growth. This results in a gigantic rosette with more leaves than wild-type, hence the name GIGANTEA (Rédei, 1962; Koornneef *et al.*, 1991) (Figure 1.6). The late-flowering phenotype of gi is most pronounced in LD, consequently GI is proposed to act in the photoperiodic pathway to regulate flowering in *Arabidopsis* (Chapter 1.4.3.5). However, gi mutant plants are still able to respond to day length as mutants also flower later than wild type in SD (Fowler *et al.*, 1999). The late flowering gi mutant phenotype is overcome by growing plants in darkness. This indicates that GI function is not essential in dark grown plants, therefore GI is thought to function in a light dependent pathway to control flowering (Araki and Komeda, 1993a). Interestingly, gi mutants can also be induced to flower early by sucrose; however the significance of this is unclear (Roldan *et al.*, 1999).

The classical gi mutant gi-2 was used in this thesis work. This mutant contains a 7 bp deletion within the GI coding sequence which introduces a premature stop codon and therefore is predicted to encode a truncated protein of 144 amino acids. The expression of the GI transcript is reduced in the gi-2 mutant plant and the flowering time of this mutant is significantly later than wild type Col in both LD and SD conditions (Fowler *et al.*, 1999).
While late-flowering is the most obvious phenotype in the gi mutant, a long-hypocotyl phenotype was observed in three gi alleles tested and this segregates with the late-flowering phenotype (Araki and Komeda, 1993b). The increase in hypocotyl length is caused by an increase in cell size and is proposed to be due to the reduced inhibition of hypocotyl elongation by light. Consistent with a role in light signalling or transduction, the gi mutant was identified in a genetic screen for loci involved in phytochrome signalling (Huq et al., 2000). The mutant isolated displayed elongated hypocotyls under continuous red light, but no change in response to far-red light. As PHYB mediates signalling in response to red light, this is suggestive that the gi mutation disrupts PHYB signalling. It is proposed that gi is defective in selective signalling downstream of PHYB as the mutation does not alter the expression levels of the PHYA or PHYB proteins (Huq et al., 2000). Curiously, the phyB mutant is early flowering and the gi mutant late, signifying that the function of GI in floral regulation is different to that of PHYB.

Further evidence for GI acting in a light signalling pathway was provided by the analysis of 35S:GI plants. Mizoguchi et al. (2005) found that plants over expressing GI had a short hypocotyl phenotype when grown under red light, indicating these plants were hypersensitive to red light. Furthermore, GI was identified in a gene expression screen for shade responsive genes (Devlin et al., 2003), raising the possibility that GI may also act to control flowering in response to light quality. The gi mutant exhibits light-dependent and light–independent circadian clock phenotypes. This and the role of GI in the circadian clock are presented in Chapter 1.5.3.

It has been proposed that GI functions in mediating stress responses in Arabidopsis. Six classical gi mutants display an increased tolerance to oxidative stress induced by paraquat exposure (Kurepa et al., 1998). Tolerance to this herbicide was apparent even in darkness, therefore this is not a light-mediated effect. Recently presented evidence established that the enhanced tolerance of gi mutants to paraquat is associated with the increased expression of anti-oxidative enzymes including ascorbate peroxidise and superoxide dismutases (Cao et al., 2006). Additionally, gi mutant plants show decreased tolerance to freezing and flowering is significantly delayed in response to daily cold treatments (Cao et al., 2005). Analysis of the Arabidopsis transcriptome using microarray technology has shown that GI transcript levels are increased 5-10 fold in response to low temperature (Fowler and Thomashow, 2002). This implies that GI is also required for effective cold tolerance in plants.
Finally, the *gi* mutant has elevated leaf starch and hexose levels in some conditions, as compared to wild-type (Eimert et al., 1995; Hollis, 1999). It has yet to be determined how carbohydrate metabolism affects flowering time in *Arabidopsis*, but for the *gi* mutant the late-flowering phenotype is not a consequence of the accumulation of starch (Eimert et al., 1995). It still remains to be elucidated how the various *gi* mutant phenotypes are connected to the biochemical function of *GI*.

### 1.5.2 ISOLATION AND CHARACTERISATION OF *GI*

The *GI* gene was identified and isolated in a screen for late flowering mutants carried out in this laboratory (Richardson et al., 1998; Fowler et al., 1999). Northern hybridisation experiments have shown that the *GI* transcript is expressed at all stages of development in wild-type *Arabidopsis* (Fowler et al., 1999). This is consistent with previous work by Araki and Komeda (1993b) that suggested that *GI* promotes flowering at the earliest stages of plant development. The *GI* transcript is detected in all tissues analysed, from seedlings to mature plants with developed siliques. As the *GI* transcript is detected throughout the plant, the tissues in which it is needed to activate flowering need to be defined.

*GI* acts as a floral promoter; over expression of *GI* from the 35S promoter results in early flowering in plants grown in both LD and SD conditions (Milich, 2001; Mizoguchi et al., 2005). *GI* is proposed to function downstream of the clock component *LHY* to promote flowering. Over-expression of *GI* can induce early flowering even in the late flowering *lhy-1* gain-of-function mutant (Mizoguchi et al., 2005). The over expression of *CO* is able to compensate for the *gi* mutation by promoting early flowering in both LD and SD. In addition, the transcript levels of the floral inducers *CO* and *FT* are reduced in *gi* mutants, indicating *GI* functions to promote flowering via *CO* (Suarez-Lopez et al., 2001). However, introduction of 35S:*GI* into *co* and *ft* mutant plants results in an intermediate flowering time. Thus the early flowering phenotype of 35S:*GI* is only partially dependent on *CO* and *FT* function (Mizoguchi et al., 2005). Kim et al. (2005a) established that *ELF3* regulates flowering through *GI* in a *CO*-independent process, implying a direct regulation of *FT* by *GI*. Therefore, *GI* may act via multiple floral pathway integrators to promote flowering.
1.5.3 GI EXPRESSION AND THE CIRCADIAN CLOCK

As *gi* mutants predominantly delay flowering in LD it is thought that *GI* is involved in regulating flowering in response to day length (Chapter 1.4.3.5). Given the role of the circadian clock in this response, it is interesting that *GI* transcript levels cycle through the day. Plants entrained in LD exhibit cycling of *GI* expression, which is maintained in plants shifted to continuous light or continuous dark conditions. Peak *GI* expression occurs in the light phase of a day 10 h after dawn and this drops to close to zero during darkness (Figure 1.7a). The *GI* expression pattern also varies between plants grown under LD and SD photoperiods, with expression levels maintained above trough levels for a longer period of time in LD (Figure 1.7b) (Fowler et al., 1999). In addition, *GI* expression is directly affected by light. This light mediated increase in *GI* is thought to be gated by the circadian clock, as the highest increase in *GI* expression occurs when the light treatment occurs close to dawn (Paltiel et al., 2006).

**Figure 1.7** Northern analysis of *GI* expression in *Arabidopsis* plants.

| a | GI expression in plants grown in long day conditions |
| b | Comparison of *GI* expression in short days (SD) and long days (LD). |

The scale at the base of the graph represents light hours and dark hours. From Fowler et al., 1999.

As might be expected, the circadian expression of the *GI* transcript is affected by mutations that affect circadian clock function. Of particular interest is the *GI* expression pattern observed in the *elf3* mutant. *GI* transcript levels do not cycle in *elf3* mutants, but are maintained at high levels at all time-points and this expression pattern correlates with early, day length-insensitive flowering in *elf3* plants (Fowler et al., 1999). Furthermore, the level of the *GI* transcript is low at all times in the *lhy-1* over expression mutant, which flowers late (Fowler et al., 1999). However, the circadian expression of *GI* is damped to
high levels at all times in late-flowering plants over-expressing CCA1, which demonstrates that high levels of GI do not promote flowering in all backgrounds (Fowler et al., 1999). Disruption of another circadian clock regulator TIME FOR COFFEE (TIC) results in early flowering plants that have reduced levels of the GI transcript and a peak of expression of GI earlier in the day (Hall et al., 2003). This suggests that the timing of GI expression is as significant as the levels of the GI transcript in promoting flowering. Further evidence that the timing of GI expression is associated with flowering control comes from investigation of the lhy-11 cca1-1 double mutants. The extreme early flowering phenotype of lhy-11 cca1-1 plants is a result of the constitutive activation of the photoperiodic pathway and is dependent on GI function. The phase of GI expression occurs earlier during the light period in these plants and subsequently an earlier phase of CO and FT expression follows (Mizoguchi et al., 2005). This is consistent with the hypothesis that the coincidence of CO expression with light is important for early flowering (Yanovsky and Kay, 2002) (Chapter 1.4.3.5).

It is hypothesised that GI may function as part of a feedback loop to regulate the clock via the light signalling pathway. Mutations in the GI gene result in a reduced level and amplitude of expression of both LHY and CCA1 (Fowler et al., 1999). These results signify that even though GI is circadian controlled, it does not act in a simple linear pathway downstream of LHY and CCA1, which are closely associated with the clock (Chapter 1.4.3.2). Loss of GI is also associated with a reduced response to light on circadian period length, which is either shorter or longer depending on the gi mutant allele tested (Park et al., 1999; Huq et al., 2000). Furthermore, the over-expression of GI results in an altered circadian period length in plants grown in constant light (Mizoguchi et al., 2005). It is likely that GI functions in a pathway distinct from ZTL, as GI expression is unchanged in late flowering plants over expressing this putative clock photoreceptor (Kim et al., 2005a). GI is more likely to act downstream of ELF3. Functional GI is required for the early flowering of elf3 mutants and ELF3 is proposed to negatively regulate GI (Chou and Yang, 1999; Fowler et al., 1999; Kim et al., 2005a).

The Arabidopsis circadian clock is proposed to act through a traditional feedback loop, however recent mathematical modelling has predicted a more complicated interlocked feedback loop that better accounts for the experimental data (Locke et al., 2005). It is proposed that GI functions as a key component of a second feedback loop that maintains the clock and mediates light input. GI is negatively regulated by clock component TOC1 as
GI transcript levels are low in TOC1 over expression lines (Makino et al., 2002). In turn, GI positively regulates TOC1, which subsequently up-regulates key oscillator genes CCA1 and LHY. This is consistent with the negative feedback loop model, where the synergistic balance between clock genes is important for regulation of the clock.

In contrast to previous work, Mizoguchi et al. (2005) found that GI is required for the correct cycling of the COLD CIRCADIAN REGULATED 2 (CCR2) gene in darkness, signifying that GI has a function independent of light to regulate the circadian clock. In a recent study of QTLs influencing temperature compensation, Edwards et al. (2005) identified GI as a putative component of a temperature compensation mechanism within the Arabidopsis circadian clock (Chapter 1.4.3.4). GI transcript levels are directly affected by changes in temperature (Cao et al., 2005; Paltiel et al., 2006). Furthermore, an increase in temperature was found to be associated with an increase in GI and, in contrast, a decrease in LHY levels. GI is required for robust cycling of TOC1 at 27°C, as TOC1 expression levels are low and constant at high temperature in gi background (Gould et al., 2006). As the balance of oscillator components is significant in clock function, it is likely that GI also has a key role in the regulation of the circadian clock in response to temperature.

Finally, two new gi alleles, the result of single amino acid substitutions, have been identified that exhibit altered circadian rhythms without a late flowering phenotype (Gould et al., 2006) (gi-596, gi-611; Figure 1.8). This indicates that the late flowering normally seen in gi mutants is not just the result of an aberration to the circadian clock. Furthermore, comparisons of Arabidopsis ecotypes Col and CVI to the Ler ecotype have identified amino acid substitutions within GI associated with circadian period QTLs (Swarup et al., 1999; Michael et al., 2003a; Edwards et al., 2005) (CVI, Ler; Figure 1.8). These amino acid differences may be utilised to dissect the GI protein and determine which regions are important for function in the circadian clock and in floral promotion.

1.5.4 The GI Protein

The GI gene is predicted to encode a large 1173 aa protein and database searches indicate that the GI protein has no significant homology with proteins of known function (Fowler et al., 1999). Computer-based predictions suggested that GI is a membrane protein (Fowler et al., 1999; Park et al., 1999), however experimental evidence generated using GUS:GI and
GI:GFP fusion constructs verified that GI is localised to the nucleus (Huq et al., 2000; Mizoguchi et al., 2005). The GI protein interacts in a yeast two-hybrid assay with SPINDLY (SPY), a negative regulator of gibberellin signalling. Analysis of the phenotypes of gi-2 spy double mutants demonstrated that SPY acts downstream of GI and upstream of CO and FT in the photoperiodic flowering. As the spy mutation suppressed the reduction of CO and FT in gi mutants, it is thought that GI acts as a negative regulator of SPY. It is also proposed that SPY and GI function together in cotyledon movement and red light responsive pathways (Southern et al., 2002; Tseng et al., 2004).

GI protein levels fluctuate in response to light/dark cycles, even in plants over expressing the GI transcript. The GI protein accumulates in the light and is actively degraded in the dark via the 26S proteasome pathway (David et al., 2006). Interestingly, the rate of GI degradation in the dark occurs more rapidly when the onset of darkness occurs early in the day, signifying that the stability of GI may be affected by day length and the circadian clock. These variations may contribute to the direct promotional effects of light on flowering in LD.

1.5.5 The characterisation of GI in other plant species

Flowering time genes from genetic pathways that control flowering in response to environmental cues such as photoperiod and vernalisation in Arabidopsis have been recognized in many crop species. The phytochrome and cryptochrome photoreceptors have been identified in all plant taxa examined, suggesting that light-responsive pathways are well conserved in plants (Lariguet and Dunand, 2005). CO is a key flowering regulator in Arabidopsis and CO-like genes have been implicated in flowering in legumes, rice and wheat (Yano et al., 2000; Yan et al., 2004; Hecht et al., 2005). Recent discoveries have verified that GI orthologs exist in other plant species and have common functions to the Arabidopsis GI gene.

The rice GI ortholog OsGI has been isolated and like GI is circadian regulated. In LD conditions over expression of OsGI activates a CO ortholog Hdl1 in rice as it does in Arabidopsis. As opposed to promoting flowering, in rice OsGI expression results in the suppression of flowering in LD (Hayama et al., 2002; 2003). It has been proposed in rice
Figure 1.8 Sequence alignment of GI-like proteins

Multiple amino acid alignment of four GI proteins. Features of the sequence are as indicated: identical residues are shaded black and residues conserved in at least three sequences are grey; areas of lower sequence conservation are underlined by the striped line. Single amino acid substitutions associated with altered circadian phenotypes are under the AtGI (Col) sequence: I113V (Ler), S191F (gi-596), L281F (gi-611), L718F (CVI) (Edwards et al., 2005; Gould et al., 2006). At, Arabidopsis thaliana; Hv, Hordeum vulgare; Le, Lycopersicon esculentum; Os, Oryza sativa; Sb, Sorghum bicolor; St, Solanum tuberosum; Ta, Triticum aestivum; Vv, Vinus vinifera.
that the genetic mechanisms of photoperiodic control are common with those in *Arabidopsis*, but diverge downstream of *GI*, at the regulation of *FT*. In LD conditions, *CO* promotes flowering through *FT* activation in *Arabidopsis* and conversely represses *FT* and flowering in rice, a SD plant (Hayama *et al.*, 2003). The *GI* ortholog in the Common Ice Plant *Mesembryanthemum crystallinum* has recently been identified and it was demonstrated that *McGI* transcripts cycled as in *Arabidopsis*, with a peak in the afternoon and a trough before dawn (Boxall *et al.*, 2005; Hecht *et al.*, 2005). An increase in *GI* expression in response to light and temperature cues was also conserved between *Arabidopsis* and the model legume *Medicago trunculata* (Paltiel *et al.*, 2006).

Curtis *et al.* (2002) have demonstrated that *GI* function is conserved in *Raphanus sativus* (radish), which like *Arabidopsis* is a Brassica plant that flowers in response to LD photoperiods. Expression of an antisense *GI* cDNA from *Arabidopsis* in transgenic radish is able to silence the endogenous *GI* transcript and delay bolting and flowering. These results demonstrate that the regulation of *GI* and its function in the control of flowering in response to photoperiod is conserved in different LD plant species. *GI* orthologs have been recently been identified in wheat (*TaGI1*) and barley (*HvGI*) (Zhao *et al.*, 2005; Dunford *et al.*, 2005). Both *TaGI* and *HvGI* are clock controlled, with a peak of expression at the end of the day. The pattern of expression is very similar to that observed for *GI* in *Arabidopsis*, confirming that the circadian regulation of *GI* is well conserved through plant species. Fusion of the *TaGI1* protein to GFP established that, like its *Arabidopsis* cousin, the *TaGI1* protein is nuclear localised. Over expression of *TaGI1* in *Arabidopsis* can complement the late flowering *gi-2* mutant via up-regulation of endogenous *CO* (Zhao *et al.*, 2005), indicating GI has the same biochemical role in both LD and SD plants.

The sequences in the Genbank database (www.ncbi.nlm.nih.gov) yielded four full length GI-like proteins (sequences provided in Appendix 3.1). Further searches of EST sequences in the TIGR Gene Indices (www.tigr.org) identified many GI-like proteins in a variety of plants species, including angiosperms and gymnosperms. Interestingly, GI-like proteins have not been found in the micro-algae *Chlamydomonas reinhardtii* (Mittag *et al.*, 2005). Comparison of GI to the *Physcomitrella* EST sequences failed to identify any GI-like proteins. The sequences of four full length GI proteins were aligned (Figure 1.8). As might be expected, *AtGI* was the most divergent of the four proteins. Overall, these four GI proteins were well conserved over the entire region of the protein and no obvious domains or regions of high conservation were detected.
1.5.6 IDENTIFYING GI PROTEIN INTERACTORS

Protein::protein interaction and complex formation are essential steps in the control of vital biological processes, such as signal transduction and transcription. Recently, proteomics based methods have proven to be extremely useful for determining the function of genes, particularly novel genes identified by genome sequencing projects. GI encodes a large plant specific protein that has been shown to be localised to the nucleus (Huq et al., 2000). GI is highly conserved throughout, however it contains no protein domains of known function (see Figure 1.8), thus the function of the protein cannot be categorized using sequence comparison. The identification of protein interactors was initiated to help elucidate the biochemical role of GI.

1.5.6.1 THE YEAST 2-HYBRID ASSAY

The yeast 2-hybrid assay is an invaluable molecular genetic tool for identifying protein-protein interactions in vivo. It is a powerful system for large scale screening, as it is possible to screen entire cDNA expression libraries for potential interactors. Importantly, yeast 2-hybrid is sensitive and cost effective (reviewed by Gietz et al., 1997; Causier and Davies, 2002). A reverse 2-hybrid system is useful for further investigation of interactions between known proteins and the detection of domains involved in mediating the protein interactions. For these reasons, in work leading up to this thesis research, the yeast 2-hybrid assay was selected for the identification and analysis of putative GI protein interactors (Snowden and Putterill, this laboratory).

The yeast 2-hybrid system used is based on the reconstruction of a transcription factor, in this case GAL4, in the event of a protein::protein interaction (Figure 1.9a). The subsequent activation by GAL4 of reporter genes allows detection of the interaction. The use of two or more reporter genes is useful for minimising the number of false positives identified. Yeast 2-hybrid assays were performed using the mating technique (Figure 1.9b), which utilise the haploid yeast strain PJ69 (James et al., 1996) of opposite mating types A and α.
Figure 1.9 The yeast 2-hybrid system

a Two plasmid types are constructed; the bait encoding protein of interest X fused in-frame with the DNA-binding domain (BD) of the transcription factor GAL4 and the prey, containing random cDNAs encoding proteins Y, fused in-frame with a transcription activation domain (AD). When both plasmids are co-transformed into a yeast strain interaction between proteins X and Y reconstitutes an active transcription factor which binds elements upstream of the reporter genes to activate their expression.

b Yeast-two hybrid assays using yeast mating. Prey and bait constructs encoding putative interacting proteins X and Y, respectively, are transformed separately into yeast strains of opposite mating type (a and a). After mating, both constructs are present in the same yeast cell and if proteins X and Y interact, reporter gene activity is detected. From Causier and Davies (2002).

1.5.6.2 THE TALE HOMEODOMAIN PROTEINS

The BELL-LIKE HOMEODOMAIN 3 (BLH3) protein was identified in a yeast 2-hybrid as a putative GI interactor (Snowden and Putterill, this laboratory). The predicted BLH3 amino acid sequence was compared to sequence data in Genbank (www.ncbi.nlm.nih.gov/) and it was found that this protein is a member of the TALE superclass of homeodomain proteins. This class is distinct from the typical homeobox genes and contains proteins from diverse species such as yeast, mice and plants. Members include the BELL and KNOX proteins in plants, human TGIF transcription factors and the yeast CUP genes. As might be expected from homeodomain proteins, the TALE proteins function in a wide range of developmental processes, from determining mating type in yeast to patterning in the animal embryo (reviewed by Burglin, 1997).

TALE (three amino acid loop extension) proteins are so called due to the presence of an additional three conserved amino acids between helix I and II in the homeodomain (Pro-Tyr-Pro). These three residues form part of a hydrophobic loop that is thought to be
involved in protein interactions between TALE and other homeobox proteins. Working with *Drosophila*, Passner *et al.* (1999) demonstrated that the three amino acid loop of EXTRADENTICLE (EXD) was important for interaction with the homeotic protein ULTRABITHORAX (UBX). TALE proteins are also unusual in that residue 50 of the third helix of the homeodomain is often a non-polar residue. As this is a position crucial for DNA binding it has been proposed that TALE proteins interact with DNA differently to typical homeodomain proteins (reviewed by Burglin, 1997). For example, EXD binds DNA poorly on its own (Chan *et al.*, 1994), yet binds efficiently as part of the EXD-UBX complex (Passner *et al.*, 1999). Direct interaction of EXD with a third protein, TALE transcription factor HOMOTHORAX, was required for the nuclear localisation of this homeodomain complex (Reickhof *et al.*, 1997; Ryoo *et al.*, 1999). Together these results suggest that in order to function efficiently to target DNA, TALE proteins interact with other homeodomain transcription factors.

In plants, the TALE proteins are represented by the BELL and KNOTTED-LIKE homeobox (KNOX) families. Although these are the only plant specific members of the TALE superclass, the KNOX family shares greater homology with a subset of TALE proteins known as MEINOX proteins due to the presence of extra conserved residues outside the homeodomain (Burglin, 1997). KNOX genes function in boundary maintenance and patterning of leaf and floral tissues (reviewed by Hake *et al.*, 2004).

BELL and KNOX proteins share considerable homology over their homeodomain regions. Twenty-two residues that are conserved over the homeodomain in KNOX proteins are also conserved in their BELL relatives. This similarity is further evident in the preference of BELL and KNOX homeodomain regions for targeting and binding similar DNA sequences containing a TGAC motif (Smith *et al.*, 2002; Chen *et al.*, 2004; Tioni *et al.*, 2005; Viola and Gonzalez, 2006). A single amino acid within helix III of the homeodomain has been identified as the main determinant for the slightly different binding properties of *Arabidopsis thaliana HOMEODOMAIN 1 (ATH1)* and *SHOOT MERISTEMLESS (STM)* (BELL and KNOX proteins, respectively) (Viola and Gonzalez, 2006).

### 1.5.6.3 BLH3 can be further classified as a member of the BELL family

The BLH3 protein can be further characterised as a member of the plant specific BELL family of transcription factors, which has thirteen putative members in *Arabidopsis*. All BELL proteins are defined by three conserved domains: the homeodomain DNA binding
motif, which is highly conserved and shares homology with the homeodomain of the TALE homeobox families including KNOX in plants and MEIS in mice; the BELL domain, and the SKY domain (Bellaoui et al. 2001). A region encompassing the SKY and BELL domains and extending to the homeodomain region is also known as the POX domain (Doerks et al., 2002). Further analyses of the BELL proteins and the relationships within the BELL family are presented in Chapter 3.

The analysis of the BELL family proteins reveals that these putative transcription factors function in diverse areas of plant development. The *Arabidopsis BELL1 (BEL1)* gene is the first and best-characterised member of this family. BEL1 encodes a nuclear localised transcription factor that is proposed to act partially redundantly with AGAMOUS (AG) in ovule development (Ray et al., 1994; Reiser et al., 1995; Western and Haughn, 1999). Mutants lack functional ovules and inflorescence stems exhibit a terminal flower phenotype (Reiser et al., 1995). The PNY gene (also known as BELLRINGER [BLR], VAAMANA [VAN] and REPLUMLESS [RPL]) is the most well studied BELL gene and is thought to function in phyllotactic patterning and stem growth in *Arabidopsis* (Smith and Hake, 2003; Bhatt et al., 2004; Byrne et al., 2003). BLR has been shown to bind directly to AGAMOUS (AG) *in vitro* and is proposed to repress AG *in planta* to allow the normal development of flowers (Bao et al., 2004). Another PNY mutant allele rpl is defective in cell differentiation at the valve borders of the silique, resulting in a decrease in fruit dehiscence (Roeder et al., 2003).

BELL proteins have now been identified in a wide range of plant species, including potato, maize and tomato, in addition to *Arabidopsis*. Recently, the rice BELL gene OsBIHD1 was found to be up-regulated in response to fungal infection, suggesting this gene is associated with a disease resistance response in rice (Luo et al., 2005). BELL proteins from diverse plants such as apple, potato, *Arabidopsis* and barley produce dwarf plants when mis-expressed (Dong et al., 2000; Muller et al., 2001; Chen et al., 2003; Smith and Hake, 2003). Furthermore, mis-expression of JUBEL1 and PENNYWISE (PNY) result in bushy plants, indicating a loss of apical dominance in these plants (Muller et al., 2001; Smith and Hake, 2003). As is evident in the bell mutant, patterning of inflorescence on stems can be altered in these transgenic plants. The apple BELL gene *MALUS DOMESTICA HOMEODOMAIN 1 (MDH1)* may also be involved in regulating patterning, as *Arabidopsis* plants mis-expressing this gene display irregular angles between stem and flower (Dong et al., 2000).
In addition to influencing patterning, some BELL genes affect flowering. *Arabidopsis* transgenic plants expressing an antisense *MDH1* from the CaMV 35S promoter flowered later than wild type plants in SD (Dong *et al.* 2000). *PNY* and homolog *POUNDFOOLISH* (*PNF*) act redundantly and in a dose-dependent manner to control inflorescence development (Kanrar *et al.*, 2006). Neither single mutant exhibits a flowering time phenotype, yet *pny pnf* double mutants do not flower as they are unable to respond to inductive conditions. The transcript levels of key indicators of flowering *LFY* and *APETALA 1* are low in *pny pnf* mutants, suggesting *PNY* and *PNF* regulate these floral meristem identity genes (Smith *et al.*, 2004). BELL family member *ATH1* is up-regulated by light in a manner that is dependent on *COP1*. Therefore *ATH1* has been proposed to function in the *COP1/DET1* light signal transduction pathway (Quaedvlieg *et al.*, 1995). Microarray analyses have detected changes in *ATH1* expression in plants shifted to darkness (Kim and von Arnim, 2006). Preliminary data demonstrates that the mis-expression of *ATH1* confers both flowering and hypocotyl elongation phenotypes in *Arabidopsis* (M. Proveniers, ISPMB conference poster, 2003), suggesting a role for *ATH1* in flowering in response to light.

It is well documented that BELL family proteins interact with the closely related KNOX homeodomain proteins. These interactions have been demonstrated in diverse plant species, from *Arabidopsis* to barley. Further studies have shown that members of the BELL family also interact with the previously uncharacterised family of *Arabidopsis* ovate proteins (Hackbusch *et al.*, 2005). These protein interactions and their functions are discussed fully in Chapter 4.

In summary, BELL proteins function to regulate a wide range of plant developmental process and a number are involved in the control of patterning and floral development. As the BELL proteins are putative transcription factors it may be expected that loss or mis-expression of these genes would result in more severe phenotypes and it is intriguing that this is generally not the case. However, the limited work on double mutants suggest that some BELL proteins may act redundantly and further research is needed to illustrate the extent of the interactions between the genes in this family.
1.6 AIMS OF THIS THESIS

The photoperiodic control of flowering time in *Arabidopsis* is regulated by the interaction of genes and day length signals. The genetic framework for the control of flowering has been constructed through the isolation and analysis of flowering time mutants. The *GI* gene is primarily involved in the control of flowering in response to day length, although *GI* functions in a wide range of developmental processes. The first aim of this thesis research was to characterise *GI* protein interactors, with the expectation of adding to the knowledge of the biological function of *GI* and its placement in the genetic network to control flowering. An approach to identify targets of *GI* is to induce *GI* expression and analyse the global changes in transcript and/or protein levels in response to *GI*. A secondary aim was to generate and characterise a chemically inducible *GI* expression system for this purpose. An introduction to the chemical induction of gene expression and preliminary experiments are described briefly in Appendix VII.

This thesis project endeavoured to characterise the interaction between the *GI* and BELL-like protein BLH3. The protein regions that mediate the *GI*::BLH3 interaction and possible interactions with other BELL proteins were investigated. This work is presented in Chapters 3 and 4. To determine if the *BLH3* and *GI* transcripts overlapped in plants, the expression of *BLH3* was analysed and this work is described in Chapter 5. *Arabidopsis* mutants were identified which contained T-DNA insertions within two BELL-like genes, *BLH3* and a closely related gene *BELL-LIKE HOMEODOMAIN 10 (BLH10)*, which encodes a protein that was shown to interact with *GI*. The final aim of this work was to determine the role of *BLH3* and *BLH10* in plants and establish how this relates to *GI* function. The circadian regulation of *BLH3* and *BLH10* was examined (Chapter 5) and reverse genetics was employed to determine if loss of these genes affected aspects of plant development, such as flowering and seedling de-etiolation, thought to be regulated by *GI* (Chapter 6).
2.1 \textbf{GENERAL}

2.1.1 \textbf{ANTIBIOTICS}

Antibiotics were made up in aqueous solution, filter sterilised and stored as indicated in Table 2.1. Two exceptions were rifampicin and gentamycin, which were dissolved in 100\% methanol and 50\% methanol respectively. Antibiotics were used at the final concentrations listed in Table 2.1.

<table>
<thead>
<tr>
<th>ANTIBIOTIC</th>
<th>SUPPLIER</th>
<th>CONCENTRATION (MG/L)</th>
<th>STORAGE °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{Bacteria}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ampicillin</td>
<td>Scientific Supplies</td>
<td>100</td>
<td>-20</td>
</tr>
<tr>
<td>Gentamycin</td>
<td>Sigma</td>
<td>50</td>
<td>4</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>Invitrogen</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>Rifampicin</td>
<td>Sigma</td>
<td>50</td>
<td>4</td>
</tr>
<tr>
<td>Spectinomycin</td>
<td>Sigma</td>
<td>200</td>
<td>4</td>
</tr>
<tr>
<td>\textit{Plant}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hygromycin</td>
<td>Roche</td>
<td>40</td>
<td>4</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>Invitrogen</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>Timentin</td>
<td>Total Laboratory Systems</td>
<td>100</td>
<td>-20</td>
</tr>
</tbody>
</table>

Table 2.1 Antibiotics added to bacterial and plant media
2.1.2 **BUFFERS AND SOLUTIONS**

<table>
<thead>
<tr>
<th>Buffer/Solution Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acrylamide: BIS 29:1</td>
<td></td>
</tr>
<tr>
<td>BLOTTO</td>
<td>1 × TBS-T, 8% non-fat milk powder</td>
</tr>
<tr>
<td>Chloroform:IAA</td>
<td>Chloroform:isoamyl alcohol (24:1)</td>
</tr>
<tr>
<td>Coomassie stain</td>
<td>20% methanol, 7% acetic acid, 0.1% Coomassie blue</td>
</tr>
<tr>
<td>Denaturation solution</td>
<td>1.5 M NaCl, 0.5 M NaOH</td>
</tr>
<tr>
<td>Electrophoresis loading dye (5x)</td>
<td>30% (v/v) glycerol, 0.04% bromophenol blue, 0.04% xylene cyanol FF, 1× TBE</td>
</tr>
<tr>
<td>GUS staining solution</td>
<td>1 mM 5-bromo-4-chloro-3-indol β-D-glucuronic acid (X-Gluc), 0.5 mM potassium ferricyanide, 0.1% Triton-X-100, 10 mM sodium phosphate buffer pH 7</td>
</tr>
<tr>
<td>Hybridisation buffer</td>
<td>10% dextran sulphate, 1 M NaCl, 1% SDS</td>
</tr>
<tr>
<td>Laemmli Sample Buffer</td>
<td>20% glycerol, 125 mM Tris-HCl (pH 6.8), 2% SDS, 0.02% bromophenol blue, 10% β-mercaptoethanol (added just before use)</td>
</tr>
<tr>
<td>Methylene blue stain</td>
<td>0.04% methylene blue, 500 mM NaAc</td>
</tr>
<tr>
<td>MOPS buffer</td>
<td>200 mM MOPS, 50 mM NaAc, 10 mM Na₂EDTA (pH 7.0)</td>
</tr>
<tr>
<td>Neutralisation solution</td>
<td>1 M NaCl, 0.5 M Tris-HCl (pH 7.5)</td>
</tr>
<tr>
<td>Prewash solution</td>
<td>1 mM EDTA, 1 M NaCl, 0.1% SDS, 50 mM Tris</td>
</tr>
<tr>
<td>RNA loading buffer</td>
<td>50% (v/v) deionised formamide, 1× MOPS buffer, 6% (v/v) formaldehyde, 6.7% (v/v) glycerol, 0.04% bromophenol blue</td>
</tr>
<tr>
<td>Running Buffer (10x)</td>
<td>3% Tris base, 1% SDS, 14.42% glycine</td>
</tr>
<tr>
<td>Solution I</td>
<td>50 mM glucose, 10 mM Na₂EDTA, 25 mM Tris-HCl (pH 8)</td>
</tr>
<tr>
<td>Solution II</td>
<td>200 mM NaOH, 1% SDS (freshly made)</td>
</tr>
<tr>
<td>Solution III</td>
<td>3 M potassium Ac (pH 4.8)</td>
</tr>
<tr>
<td>SSC</td>
<td>150 mM NaCl, 15 mM tri-sodium citrate</td>
</tr>
<tr>
<td>TAE</td>
<td>1M EDTA, 40 mM Tris-Ac (pH 8.0)</td>
</tr>
<tr>
<td>TBE</td>
<td>90 mM H₃BO₃, 2 mM Na₂EDTA, 90 mM Tris-HCl (pH 8.0)</td>
</tr>
<tr>
<td>TBS</td>
<td>200 mM Tris, 1.5 M NaCl (pH 7.6)</td>
</tr>
<tr>
<td>TBS-T</td>
<td>1 × TBS, 0.1% Tween</td>
</tr>
<tr>
<td>TE</td>
<td>1 mM Na₂EDTA, 10 mM Tris-HCl (pH 8.0)</td>
</tr>
<tr>
<td>TER</td>
<td>TE, 20 µg/mL RNase A</td>
</tr>
<tr>
<td>Transfer Buffer</td>
<td>25 mM Tris, 192 mM glycine, 20% methanol, 0.01% SDS</td>
</tr>
</tbody>
</table>

All percentages are weight by volume (w/v) unless otherwise stated.
### 2.1.3 Plasmids

The following plasmids were used in this work.

<table>
<thead>
<tr>
<th><strong>Plasmid</strong></th>
<th><strong>Description</strong></th>
<th><strong>Source</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>BJ69</td>
<td>Binary vector containing T-DNA with hygromycin selection in plants</td>
<td>B. Jansen, Hort Research</td>
</tr>
<tr>
<td>p11OP</td>
<td>Shuttle vector containing 11 optimised lac operator sequences and a minimal CaMV 35S promoter</td>
<td>I. Moore</td>
</tr>
<tr>
<td>p35S:GI-KM</td>
<td>35S:GI construct in pGREEN based binary vector</td>
<td>K. Lee</td>
</tr>
<tr>
<td>p35SBLH3:GFP</td>
<td>BLH3 cDNA cloned into p35S:GFP to form a GFP protein fusion</td>
<td>A. Goldschmidt</td>
</tr>
<tr>
<td>p35S:GFP</td>
<td>CaMV 35S promoter and GFP cloned into pGEMT</td>
<td>A. Goldschmidt</td>
</tr>
<tr>
<td>p35SGFP:BLH3</td>
<td>BLH3 cDNA cloned into p35S:GFP to form a GFP protein fusion</td>
<td>A. Goldschmidt</td>
</tr>
<tr>
<td>pACT</td>
<td>Yeast prey vector for yeast 2-hybrid, contains the GAL4 activation domain</td>
<td>Clontech</td>
</tr>
<tr>
<td>pACT:BLH3</td>
<td>BLH3 partial cDNA, isolated in yeast 2-hybrid</td>
<td>K. Snowden</td>
</tr>
<tr>
<td>pACT:BLH10 dom1</td>
<td>BLH10 BELL domain sequences cloned into pACT</td>
<td>This work</td>
</tr>
<tr>
<td>pACT:ATHdom1</td>
<td>ATH1 BELL domain sequences cloned into pACT</td>
<td>This work</td>
</tr>
<tr>
<td>pACT:BELLdom1</td>
<td>BELL1 BELL domain sequences cloned into pACT</td>
<td>This work</td>
</tr>
<tr>
<td>pACT:dom1</td>
<td>BLH3 BELL domain sequences cloned into pACT</td>
<td>This work</td>
</tr>
<tr>
<td>pACT:dom2</td>
<td>BLH3 HD sequences cloned into pACT</td>
<td>C. Pullen</td>
</tr>
<tr>
<td>pACT:dom3</td>
<td>BLH3 HD and C-terminal sequences cloned into pACT</td>
<td>C. Pullen</td>
</tr>
<tr>
<td>pACT:dom4</td>
<td>BLH3 BELL domain + HD sequences cloned into pACT</td>
<td>This work</td>
</tr>
<tr>
<td>pART27</td>
<td>Binary vector containing T-DNA with kanamycin selection in plants</td>
<td>Gleave, 1992</td>
</tr>
<tr>
<td>pART27:BLH3</td>
<td>BLH3 cDNA cloned into pART27 binary vector</td>
<td>This work</td>
</tr>
<tr>
<td>pART27:BLH10-R</td>
<td>BLH10 RAFL cDNA in pVK binary vector</td>
<td>This work</td>
</tr>
<tr>
<td>pART7</td>
<td>Shuttle vector, CaMV 35S promoter and <em>ocs</em> 3' region flank the multiple cloning site.</td>
<td>Gleave, 1992</td>
</tr>
<tr>
<td>pAS2-1</td>
<td>Yeast bait vector for yeast 2-hybrid, contains GAL4 binding domain</td>
<td>Clontech</td>
</tr>
<tr>
<td>pAS:BLH3</td>
<td>BLH3 HD, BELL + SKY domain sequences cloned into pAS2-1</td>
<td>This work</td>
</tr>
<tr>
<td>pAS:BLH10</td>
<td>BLH10 HD, BELL + SKY domain sequences cloned into pAS2-1</td>
<td>This work</td>
</tr>
<tr>
<td>pAS:del1</td>
<td>GI deletion in pAS2-1</td>
<td>C. Pullen</td>
</tr>
</tbody>
</table>
pAS:del2  
\( GI \) deletion in pAS2-1  
C. Pullen\(^a\)

pAS:del4  
\( GI \) deletion in pAS2-1  
C. Pullen\(^a\)

pAS:del5  
\( GI \) deletion in pAS2-1  
This work

pAS:del6  
\( GI \) deletion in pAS2-1  
This work

pAS:del7  
\( GI \) deletion in pAS2-1  
This work

pAS:GI  
1.5 kb 3’ \( GI \) cDNA fragment cloned into pAS2.1  
K. Snowden\(^a\), this work

pAVA-393  
35S:GFP expression vector  
von Arnim et al., 1998

pBIN:11OP:GI  
11OP:GI expression cassette in binary vector  
This work

pBIN:35SGFP  
35S:GFP cloned into a binary vector  
A. Goldshmidt\(^a\)

pBIN:35SBLH3:GFP  
35SBLH3:GFP cloned into a binary vector  
A. Goldshmidt\(^a\)

pBIN:35SGFP:BLH3  
35SGFP:BLH3 cloned into a binary vector  
A. Goldshmidt\(^a\)

pBluescript (pBS)  
pUC derived cloning vector, \( lacZ \) gene for blue/white selection.  
Stratagene

pBS:BLH10  
pBS containing full length \( BLH10 \) cDNA  
This work

pBS:BLH10/A  
pBS:BLH10 with adaptor containing \( EcoRI \) site introduced into \( Bg/II \) site  
This work

pBS:BLH3  
Partial \( BLH3 \) cDNA subcloned from pACT-BLH3  
This work

pBS:BLH3/A  
pBS:BLH3 with adaptor to produce in frame ATG  
This work

pBS:del6  
\( GI \) deletion in pBS  
This work

pBS:GI  
Full length \( GI \) cDNA in pBS  
S. Fowler\(^a\)

pBS:GI/R  
\( GI \) cDNA fragment from pGEMT:GI/R  
This work

pBS:SstI-  
pBS derivative, \( SstI \) site deleted  
K. David\(^a\)

pGEMT  
Cloning vector  
Promega, Biotek

pGEMT:258  
258 bp \( GI \) PCR fragment in pGEMT  
This work

pGEMT:BLH3  
Partial \( BLH3 \) cDNA amplified from pACT-BLH3, flanking \( EcoRI \) and \( BamHI \) sites introduced  
This work

pGEMT:Del4  
\( GI \) deletion Del4 amplified from pAS-Del4 and inserted into pGEMT  
Trent Bosma\(^a\)

pGEMT:GI5'  
\( GI \) 5’ region amplified from pBS:GI and subcloned into pGEMT, \( KpnI \) enzyme site introduced adjacent to ATG  
This work

pGEMT:GI  
\( GI \) amplified from pAS:GI, contains error  
Trent Bosma\(^a\)

pGEMT:GI/R  
\( GI \) amplified from pAS:GI and inserted into pGEMT  
This work

pMAL  
Vector for expression of protein fused to Maltose Binding Protein  
New England Biolabs

pMAL:BLH3  
\( BLH3 \) from pGEMT-BLH3 subcloned into pMAL  
This work

pMAL:del4  
\( GI \) deletion del4 subcloned into pMAL  
This work

pMAL:GI  
\( GI \) subcloned from pAS:GI into pMAL  
This work
Chapter 2

pR1  GI cDNA cloned into pBluescript. NotI sites removed via mutagenesis.  R. Moyle

pR24  BLH3 cDNA cloned into pART7  This work

pR26.8  5' BLH10 RAFL cDNA cloned into pVK  This work

pR29  GI cDNA cloned into pBS-SstI  This work

pR30  pR29 derivative, GI 5' UTR sequences removed  This work

pR33  GI cDNA from pR30 cloned into p11OP  This work

pR34  11OP:GI subcloned from pR33 into pR7 binary vector  This work

pR7  Derivative of BJ69, CAT reporter gene removed  R. Moyle

pTD1-1  SV40 large T antigen in pACT2  Clontech

pVA3-1  Murine p53 in pGBT9  Clontech

pVK  pART27 containing 35S and ocs 3' sequences from pART7  V. Kelly

pZL:BLH3  Full length BLH3 cDNA clone isolated from cDNA library  C. Pullen

\[ a \]  this laboratory
\[ b \]  University of Oxford, UK
\[ c \]  John Innes Research Centre, Norwich, UK

2.1.4 Oligonucleotides

The oligonucleotide primers used in this work were obtained from Invitrogen. The sequences and uses of the primers are listed below. Restriction enzyme sites are highlighted in bold.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence 5' → 3'</th>
<th>Use</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>5dom1BamHI.Fwd</td>
<td>TGAGTTATCGAGGGATCCG &lt;br&gt; AATGGGAAA</td>
<td>BLH10 amplification primer, introduction of enzyme site for cloning</td>
<td>This work</td>
</tr>
<tr>
<td>5dom1BamHI.Rev</td>
<td>CAACAGGTGGATCCATTC &lt;br&gt; CAAGCTGTT</td>
<td>BLH10 amplification primer, introduction of enzyme site for cloning</td>
<td>This work</td>
</tr>
<tr>
<td>Ap1</td>
<td>GGATCCCTAATACGACTCAC &lt;br&gt; TATAGGGC</td>
<td>Adaptor primer for genome walking</td>
<td>M. Yoon</td>
</tr>
<tr>
<td>Ap2</td>
<td>TATAGGGCTCGAGCCGC</td>
<td>Nested adaptor primer for genome walking</td>
<td>M. Yoon</td>
</tr>
<tr>
<td>ATHBamHI.Fwd</td>
<td>CTTGATGGGGATCCCTAATA &lt;br&gt; ACTCGGAGG</td>
<td>ATH1 amplification primer, introduction of enzyme site for cloning</td>
<td>This work</td>
</tr>
<tr>
<td>ATHXhoI.Rev</td>
<td>AAGAAGGCAATGGCGAGG &lt;br&gt; AACATAGAG</td>
<td>ATH1 amplification primer, introduction of enzyme site for cloning</td>
<td>This work</td>
</tr>
<tr>
<td>BELBamHI.Fwd</td>
<td>GGGTAAACAACAAGGAGTGGATCCG &lt;br&gt; AAGT</td>
<td>BEL1 amplification primer, introduction of enzyme site for cloning</td>
<td>This work</td>
</tr>
<tr>
<td>BELXhoI.Rev</td>
<td>GAGTCATTGGAGCTCAACTAAA</td>
<td>BEL1 amplification primer, introduction of enzyme site for cloning</td>
<td>This work</td>
</tr>
</tbody>
</table>
TTCTGTTGCC

for cloning

GCTGATCCATTA

TA

Del6BamHI.Rev

GCCGCAGATCCATTA

Del6.Fwd

tTTTGTGCCTTTGCCTGT

GA

Del7BamHI.rev

GCCGCAAGATCCATTA

CACCCTAC

Del7EcoRI.Fwd

CTCATCGTTTGGCAGATT

CATATCTTT

Dom1BamHI.fwd

GCTGATCCAACCATCC

AAGTTTGTGATG

Dom1XhoI.Rev

GCCATTAGAGCTCTACG

CTAAGGCTTA

Dom2aXhoI.Rev

GATTAAGAGCTGAGATTA

CAGTCGTAGTC

Dom2BamHI.Fwd

GAGAGATCCCGTCTTAG

GTACTTAGATCA

g702

GACTCGAGTCGACATA

3' UBIQUITIN RT-PCR primer

Frohman et al., 1988

g775

GACTCGAGTCGACATA

RT-PCR (dT)_{17} adapter primer

Frohman et al., 1988

GA9.1F

GATTGGTTCCGGTGATG

BLH3 amplification primer,

introduction of enzyme site

for cloning

GA9.1R

TTTCTACCTTTCTTATTCC

C

GA9.libL

CCAGTGCTGTATATA

CT

BLH3 amplification primer,

introduction of enzyme site

for cloning

GA9L4

GATCACAATCTTTTAGAG

AGTTCCGTCGT

BLH3 amplification primer,

introduction of enzyme site

for cloning

GA9like1.F

TTGTAACAGGAAACGGTAC

C

BLH10 sequencing and

amplification primer

GA9like1.R

TCCTCCTCTCTATCTCTTA

C

BLH10 sequencing and

amplification primer

GA9like2.R

CCTTGGTAGCTACCTCTG

TGTG

BLH10 sequencing and

amplification primer

GA9R2

CTTCAAGAATCATACTCCC

CCAATCAACT

BLH3 amplification primer,

introduction of enzyme site

for cloning

GI3'.Fwd

CAGAAGGAATATCGTGACCT

TTTAAACTGG

GI amplification primer,

introduction of enzyme site

for cloning

GI5'KpnI.fwd

CGTTACCGCTATTAATTGC

TTC

GI amplification primer,

introduction of enzyme site

for cloning

GR1.F

AACCTGCTGGCTTTGAC

CTGAT

GR amplification primer

This work

GR1.R

TTCCCTCCCTTGACGAT

GR amplification primer

This work

C. Pullen a

C. Pullen a

C. Pullen a

C. Pullen a

C. Pullen a

C. Pullen a

C. Pullen a
<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>LB2</td>
<td>GGCT</td>
<td>Nested T-DNA left border primer</td>
<td>TMRI&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>LB3</td>
<td>TAGGCTCTGAATTTTCATAA</td>
<td>T-DNA left border primer in &lt;i&gt;blh3&lt;/i&gt; and &lt;i&gt;blh10&lt;/i&gt; mutants</td>
<td>TMRI&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ls (blunt)</td>
<td>PO&lt;sub&gt;4&lt;/sub&gt;-ATCTGCCAACA</td>
<td>Lower strand adaptor primer for genome walking</td>
<td>M. Yoon&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ls (Taql)</td>
<td>CGACCTGCCGAA</td>
<td>Lower strand adaptor primer for genome walking</td>
<td>M. Yoon&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ls (XbaI)</td>
<td>CTAGACCTGCCGAA</td>
<td>Lower strand adaptor primer for genome walking</td>
<td>M. Yoon&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>M13F</td>
<td>TCCCAATGCTAGACGTCGT</td>
<td>pGEMT sequencing and amplification primer</td>
<td>Promega</td>
</tr>
<tr>
<td>M13R</td>
<td>GGAAACAGCTATGACCATG</td>
<td>pGEMT sequencing and amplification primer</td>
<td>Promega</td>
</tr>
<tr>
<td>oli 7</td>
<td>CCAAATACGACGCCACCAC</td>
<td>GI amplification primer</td>
<td>K Lee&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>oli1</td>
<td>GGTATGCGCATATAAGG</td>
<td>GI amplification primer</td>
<td>K Lee&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>oli12</td>
<td>ACGGCAAGGCAATACAG</td>
<td>GI amplification primer</td>
<td>&quot;</td>
</tr>
<tr>
<td>oli2(R)</td>
<td>TGGTTCAAGAGCTGGAAG</td>
<td>GI amplification primer</td>
<td>&quot;</td>
</tr>
<tr>
<td>oli22</td>
<td>AGTATCGGAGCAATGCGC</td>
<td>GI amplification primer</td>
<td>&quot;</td>
</tr>
<tr>
<td>oli33(R)</td>
<td>GGAAGATTCAAGGCAGAAA</td>
<td>GI amplification primer</td>
<td>&quot;</td>
</tr>
<tr>
<td>oli36(R)</td>
<td>CCTGGGAATTTGCTGATG</td>
<td>GI sequencing and amplification primer</td>
<td>&quot;</td>
</tr>
<tr>
<td>oli37(R)</td>
<td>GGAGAACCACATAGTGGTAGC</td>
<td>GI amplification primer</td>
<td>This work</td>
</tr>
<tr>
<td>p27.35S</td>
<td>CGTCATCCCTTAAGTCTAGTG</td>
<td>CaMV 35S promoter sequencing and amplification primer</td>
<td>K. Richards&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>pACT.Rev</td>
<td>ACTTGCAGGTTTTTTCAGT</td>
<td>pACT amplification primer</td>
<td>C. Pullen&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>pASBamHI.Rev</td>
<td>GGATGATTCCCTAAGAGCTCA</td>
<td>pAS2-amplification primer, introduction of enzyme site for cloning</td>
<td>C. Pullen&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>RAK1</td>
<td>GAGGCTATTCCGCCTGATGAC</td>
<td>5' &lt;i&gt;nptII&lt;/i&gt; sequencing and amplification primer</td>
<td>R. Atkinson&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>RAK2</td>
<td>AATCTCGTGTGGCAGGTTG</td>
<td>3' &lt;i&gt;nptII&lt;/i&gt; sequencing and amplification primer</td>
<td>&quot;</td>
</tr>
<tr>
<td>RB(Q)3</td>
<td>TAACAATTTTCACACAGAAGA</td>
<td>T-DNA RB primer in &lt;i&gt;blh3&lt;/i&gt;</td>
<td>TMRI&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>RB3</td>
<td>CGCCATGCGCAATATGCTGAC</td>
<td>T-DNA RB primer in &lt;i&gt;blh10&lt;/i&gt;</td>
<td>TMRI&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Us</td>
<td>CTAATACGACTCAGCTATAGG</td>
<td>Upper strand adaptor primer for genome walking</td>
<td>M. Yoon&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>UB01</td>
<td>CTACCCTGATCAAGATGCA</td>
<td>5' &lt;i&gt;UBQUITIN&lt;/i&gt; RT-PCR primer</td>
<td>S. Ledger&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
Chapter 2

2.1.5 CHEMICALS AND LABORATORY CONSUMABLES

All chemicals and laboratory consumables used in this thesis work were of analytical grade and were purchased from the following suppliers:

Radioactive 5’-[α-32P] dCTP (370 MBq/mL) and L-[35S] methionine were obtained from Amersham Pharmacia.

2.2 BACTERIA

2.2.1 BACTERIAL STRAINS

The following bacterial strains were used in this work.

*Agrobacterium tumefaciens (Agrobacterium)*

*Agrobacterium* genotype GV3101 (pMP90) was used to mediate gene transfer in *Arabidopsis*. This strain contains the gene encoding rifampicin resistance on the bacterial chromosome. GV3101 carries the pMP90 plasmid which has the following properties: encodes the gene that confers gentamycin resistance, contains the *vir* genes which act in trans to mediate gene transfer in plants and has a deleted T-DNA region (Koncz and Schell, 1986).

*Escherichia coli (E. coli)*

E. coli DH10B [genotype: F' *mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZΔM15 ΔlacX74 deoR recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ rpsL nupG] was used as a host strain in bacterial transformation. This strain allows blue/white screening to distinguish recombinant and non-recombinant plasmids. The *E. coli* strain pRK2013 (Ditta *et al.*, 1980) was used as a “helper” strain during *Agrobacterium* transformation by triparental mating (Chapter 2.2.3.2). Plating bacteria for library screening was the strain XL1-Blue MRF* [genotype: Δ(mcrA)183 Δ(mcrCB-hsd SMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 recA1 lac(F′ proAB lacIΔZΔM15 Tn10 Tet’)]. The BL21-DE3 strain [genotype: B F−...
Chapter 2

dcm amp<sup>T</sup> hsdS (r<sub>B</sub> m<sub>B</sub>) gal<sub>λ</sub> (DE3)] was used for the expression of recombinant BLH3 protein in E. coli.

2.2.2 BACTERIAL MEDIA

Luria (L) Broth  
1% peptone, 1% NaCl, 0.5% yeast extract, pH 7.

NZY media  
0.5% NaCl, 0.2% MgSO<sub>4</sub>, 0.5% yeast extract, 1% NZ amine, pH 7.5

Revco media  
40 mM K<sub>2</sub>HPO<sub>4</sub>, 22 mM KH<sub>2</sub>PO<sub>4</sub>, 1.7 mM sodium citrate, 0.4 mM MgSO<sub>4</sub>.7H<sub>2</sub>O, 30% glycerol.

Rich Broth (RB)  
1% tryptone, 0.5% yeast extract, 0.5% NaCl, 0.2% glucose

SOC media  
2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>, 20 mM glucose.

Terrific broth  
1.2% tryptone, 2.4% yeast extract, 0.4% (v/v) glycerol, pH 7. Following autoclaving, 0.17 M KH<sub>2</sub>PO<sub>4</sub> and 0.72 M K<sub>2</sub>HPO<sub>4</sub>

Yeast nutrient  
0.3% beef extract, 0.5% peptone, 1% yeast extract, 0.14 mM NaCl, pH 7.3.

(YN)

When required, media was solidified with 1.5% Davis agar.

2.2.3 TRANSFORMATION OF BACTERIA

2.2.3.1 TRANSFORMATION BY ELECTROPORATION

To prepare E. coli cells for electroporation, aliquots of DH10B cells (~20 µL) were used to inoculate 2 × 10 mL of L-Broth and incubated over night at 37°C. The next morning 5 mL aliquots were subcultured into 2 × 250 mL of prewarmed L-Broth. These cultures were incubated at 37°C until they reached an OD of 0.8-0.9 and were then incubated on ice for at least 30 min. The cells were pelleted by centrifugation (4700 xg, 15 min) and washed in 250 mL ice cold water. The centrifugation and rinsing steps were repeated, this time with 125 mL cold water. The cells were again pelleted by centrifugation (4700 xg, 15 min), resuspended in 10% glycerol (12 mL) and centrifuged once more. The pellet containing the bacterial cells was resuspended in 10% glycerol (1 mL) and aliquots (40 µL) were dispensed into eppendorf tubes and stored at -80°C.

The transformation of E. coli competent cells was performed using a Bio-Rad Gene Pulser and a Bio-Rad Pulse Controller. Aliquots of competent E. coli cells (40 µL) and DNA to be transformed (10-20 ng) were gently mixed and incubated on ice for 1 min. The cells were transferred to a pre-chilled 1 mm Bio-Rad Gene Pulser<sup>®</sup> cuvette, electroporated (1.8
V, 25 µF, 250 mA) and immediately resuspended in 1 mL of SOC media. The mixture was removed to an eppendorf tube and incubated at 37°C for 1 h. Aliquots of 100-200 µL were spread onto L-plates containing the appropriate antibiotic selection and incubated overnight. In order to determine the transformation efficiency, competent cells were transformed with 1-2 ng of pBS DNA and the number of colonies per µg DNA was calculated. Transformation efficiencies up to $1 \times 10^8$ were obtained.

### 2.2.3.2 Transformation of Agrobacterium

Triparental mating was used to transfer binary plasmids from *E. coli* to *Agrobacterium*; this method is modified from that of Ditta *et al.* (1980). Bacterial cells were removed from separate selective plates containing *E. coli* carrying the binary vector, the *E. coli* helper strain pRK2013 and *Agrobacterium* to be transformed and smeared together on an L-plate. Biparental matings and single parent strains were also smeared onto L-plates as controls. Plates were incubated overnight at 28°C, then cells from all crosses were streaked onto plates selecting for *Agrobacterium*, the binary and the Ti plasmids, allowing only colonies resulting from triparental mating to grow. Plates were incubated for 2-3 d at 28°C and resulting colonies were streaked in duplicate onto selective YN plates. To assess if the culture was *Agrobacterium*, one plate was placed at 28°C and the other at 37°C. Colonies were present at 28°C, but not at 37°C, were accepted as *Agrobacterium*.

### 2.2.4 Long-term Storage of Bacteria

A single bacterial colony was used to inoculate 2 mL Terrific broth and incubated over night at 37°C. The following day cells were collected by centrifugation at 4500 ×g for 1 min and the supernatant was discarded. The bacterial cells were resuspended in 500 µL 50% glycerol and stored at -80°C.

### 2.2.5 Screening a cDNA Library

The *Arabidopsis* cDNA library CD4-15 (Kieber *et al.*, 1993) was screened in this work and a full-length cDNA clone of *BLH10* (1906 bp) was isolated. The CD4-15 library was constructed in the λZAP bacteriophage vector and contained inserts of 2-3 kb from the Col ecotype. CD4-15 was obtained from the *Arabidopsis* Biological Resources Centre. The plating of bacteriophage, library screening and excision of phagemid clones was carried out according to the instructions of the Zap-cDNA Gigapack III Gold Kit (Stratagene). The *BLH10* probe used to screen the library were amplified from genomic DNA by PCR.
(Chapter 2.5.4) and corresponds to the coordinates [-55 to +294], relative to the A of the translation start codon.

2.3 Yeast

2.3.1 Yeast Strain

A *Saccharomyces cerevisiae* (*S. cerevisiae*) strain modified for use in yeast 2-hybrid systems, PJ69-4A (Genotype: *trp1-901 leu2-3, 112 ura3-52 his3-200 gal4 gal80 LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*) was used in this work. Two haploid mating strains (MATa [4a], MAT [4]) were used and mated to generate diploid yeast containing both the pACT and pAS2-1 based plasmids. This host strain contains three reporter genes and is sensitive to weak interactions, allowing efficient selection of interactions and minimising false positives (James *et al.*, 1996).

2.3.2 Yeast Media

<table>
<thead>
<tr>
<th>Media</th>
<th>Constituents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>YPD</strong></td>
<td>1% yeast extract, 2% D-glucose, 2% peptone, pH 5.8</td>
</tr>
<tr>
<td><strong>SC medium</strong></td>
<td>0.67% yeast nitrogen base without amino acids (DIFCO), 2% D-glucose.</td>
</tr>
</tbody>
</table>

Supplementary amino acids were added in the following final concentrations:

- Adenine: 20 mg/L
- Histidine: 20 mg/L
- Leucine: 100 mg/L
- Lysine: 30 mg/L
- Methionine: 20 mg/L
- Tryptophan: 20 mg/L
- Uracil: 20 mg/L

When required, media was solidified with 2% Davis Agar.

2.3.3 Transformation of Yeast

Yeast colonies measuring 2-3 mm in diameter were scraped from a fresh plate and resuspended in 1 mL YPD by vigorous vortexing. This suspension was used to inoculate 50 mL YPD, which was incubated at 30°C until the culture reached stationary phase (16-18 h). The culture was diluted in 300 mL YPD (to OD<sub>600</sub> ~0.2-0.3) and incubated at 30°C until OD<sub>600</sub> had reached 0.4-0.6. The cells were collected by centrifugation at 1000 xg for 5 min and resuspended in 150 mL TE. The centrifugation step was repeated and cells were
resuspended in 1.5 mL TE/LiAc (1 x TE, 100 mM LiAc). These competent cells were used immediately.

Plasmid DNA (100 ng), calf thymus carrier DNA (100 µg) and PEG/LiAc (600 µL; 40% PEG4000, 100 mM LiAc) were added in order to 100 µL yeast competent cells and thoroughly mixed. The competent cells were incubated at 30°C for 30 min with gentle shaking. The cells were subjected to heat shock in a water bath (42°C, 15 min), followed by a 1-2 min incubation on ice. The cells were collected by brief centrifugation and resuspended in 500 µL TE. Aliquots were plated onto the appropriate selective media.

2.3.4 Yeast Mating and Yeast 2-Hybrid

The bait (pAS2.1), prey (pACT) and control vectors (pVA3-1, pTD1-1) were obtained as part of the Matchmaker Two-hybrid System 2 (Clontech). Opposite haploid mating strains of S. cerevisiae PJ69 were transformed separately with ‘bait’ and ‘prey’ vectors; this is outlined below. All pAS based vectors were maintained in the mating strain PJ69 4a. The pACT based vectors were transformed into the opposite yeast mating strain PJ69 4α. Prey and bait strains were streaked onto separate selective plates as outlined below and grown at 30°C for 3 d.

<table>
<thead>
<tr>
<th>PREY</th>
<th>BAiT</th>
</tr>
</thead>
<tbody>
<tr>
<td>PJ69 mating strain</td>
<td>4a</td>
</tr>
<tr>
<td>Vector</td>
<td>pAS2-1</td>
</tr>
<tr>
<td>Selection plate</td>
<td>SC-trp</td>
</tr>
<tr>
<td></td>
<td>SC-leu</td>
</tr>
</tbody>
</table>

Colonies from the prey and bait plates were pressed onto the same velvet so that they were oriented at right angles in a grid-like pattern. The velvet was used to lift colonies from selective plates to a non-selective YPD plate. Plates were incubated overnight at 31°C to allow mating to occur. Colonies were lifted from the YPD plate and plated onto SC-trp-leu plates to select for diploid yeast containing both plasmids. Colonies were also replica plated onto SC media lacking adenine (SC-ade) or histidine (SC-his + 2 mM 3-aminoo-1’,2’,4’-triazol [3-AT]) respectively. Selective plates were incubated for 2-3 d at 30°C and scored for growth. Each interaction was tested using two independent clones and repeated at least 3 times.
2.3.5 **LONG TERM STORAGE OF YEAST**

Yeast colonies measuring ~2 mm in diameter from fresh selective plates were used to inoculate 2-3 mL of the appropriate selective medium. Cells were dispersed by vortexing and incubated 2-3 d at 30°C. The cells were collected by centrifugation at 4500 ×g for 1 min and the supernatant was discarded. The yeast cells were resuspended in 500 µL Revco media and stored at -80°C.

2.4 **ARABIDOPSIS**

2.4.1 **GROWTH OF ARABIDOPSIS**

Wild-type *Arabidopsis thaliana* (L.) Heynh (*Arabidopsis*) ecotype Columbia (Col) and the *gi-2* mutant (*Arabidopsis* Biological Resources Centre, Ohio, USA) were primarily used during this thesis work. The T-DNA mutant plants *blh3* and *blh10* (ecotype Col) were obtained from the Syngenta *Arabidopsis* Insertion Library (Sessions *et al.*, 2002).

2.4.1.1 **PLANT GROWTH MEDIA**

- **Germination medium (GM)**
  10 µM FeEDTA, 1 × Major Salts, 1 × Minor Salts, 1 × MS Vitamins (Murashige and Skoog, 1962) 1% sucrose, pH 5.7.
- **Infiltration medium (IM)**
  10 µM FeEDTA, 1 × Major Salts, 1 × Minor Salts, 1 × MS Vitamins (Murashige and Skoog, 1962) 5% sucrose, pH 5.8. Prior to use 6-benzylaminopurine (in DMSO) was added to a final concentration of 10 µg/L.
- **Murashige and Skoog Medium (MS)**
  10 µM FeEDTA, 1 × Major Salts, 1 × Minor Salts, 1 × MS Vitamins (Murashige and Skoog, 1962), pH 5.7.

When required, media were solidified with 0.8% Davis Agar.

2.4.1.2 **GENERAL GROWTH CONDITIONS**

Prior to sowing, seeds were placed on damp filter paper and stored at 4°C for 2-4 days. Following stratification seeds were sown on damp seed raising mix. Once seedlings had formed true leaves they were transplanted to YATES Kwik Pot trays (42 cell) containing...
Watkins potting mix. Late flowering mutants were treated with Nitrosol liquid fertiliser (Yates NZ) to ensure plants remained healthy before setting seed. Plants were grown in Percival Controlled Environment Chambers (22°C, ~125 μmol.m⁻².s⁻¹ fluorescent light) or under natural light in the glasshouse (supplemented with artificial light during winter months). Plants were grown under LD (16 h light, 8 h dark) and SD (8 h light, 16 h dark) light cycles unless otherwise stated.

2.4.1.3 Harvesting Arabidopsis Seed
All shoots with siliques were fastened inside a Cello bag as soon as the lower siliques had begun to yellow. The plants continued to be watered until flowering had ended, at which time all shoots carrying siliques were harvested and left to dry naturally. Once the plants were sufficiently dry seeds were collected from the bag and stored in microfuge tubes at room temperature.

2.4.1.4 Growing Seedlings on Agar Plates
All seed to be sown on agar plates was sterilised to prevent bacterial and fungal growth. Small amounts of seed (<200) were sealed in individual ‘packets’ of Miracloth (Behring Diagnostics) and placed in a Schott bottle. Bulk seed was sterilised in 50 mL Falcon tubes, with a maximum of 27000 seed (540 mg) per tube.

After soaking in water for 30 min, seeds were pelleted by centrifugation at 1500 ×g for 3 min. The seeds were treated with 95% ethanol for 5 min, followed by centrifugation at 1000 ×g for 2 min. The ethanol was replaced by 7.5% sodium hypochlorite containing 0.5% Triton X-100 for 5-10 min and centrifuged as before. All remaining steps were carried out in a sterile laminar flow cabinet. The seeds were rinsed 5 times in sterile water to remove all sodium hypochlorite, centrifuging between each wash to prevent seed loss during changes of solution. Centrifugation steps were omitted when sterilising seed ‘packets’ in a Schott bottle. The seeds were resuspended in a solution of 0.1% agarose and plated on GM plates containing the appropriate antibiotic. blh3 and blh10 mutants were selected on plates containing the herbicide Basta (18% glufosinate, 10 mg/L). Alternately, seeds were dried slightly and placed directly onto GM/MS plates. Following 3-5 d of cold treatment at 4°C the seedlings were grown in a Percival Controlled Environment Chamber.

2.4.1.5 Analysis of Flowering Time
Plants were grown as described (Chapter 2.4.1.2) in SD (8 h L, 16 h D) or LD (18 h L, 6 h D) conditions. The number of days from germination to the time at which the first floral bud was visible and the total number of leaves (rosette and cauline) at flowering was determined.

### 2.4.1.6 Hypocotyl elongation assays

This protocol was adapted from those described in Moller et al. (2003) and Duek and Fankhauser (2003). Filter paper was sterilised and placed on the surface of large, square MS plates. Seeds were sterilised in 95% ethanol, air dried and plated individually in a grid formation onto prepared MS plates. Seeds were stratified in the dark at 4°C for 3-5 d to ensure uniform germination. Plates were placed in growth cabinets under white light for 1-2 h to induce germination and were then removed to treatment light conditions. Treatment light conditions included continuous white light (LL), constant darkness (DD), low intensity white light (0.1 µmol.m$^{-2}$.s$^{-1}$ fluorescent light), constant blue light (Bc, 450 nm, 15 µmol.m$^{-2}$.s$^{-1}$), constant red light (Rc, 640 nm, 30 µmol.m$^{-2}$.s$^{-1}$) and far red light (FR, 750 nm, ~20 µmol.m$^{-2}$.s$^{-1}$). Bc and Rc light was produced using sheets of LEE photographic filters placed beneath cool white fluorescent bulbs and FR light was generated by high powered light emitting diodes (LED750, Roithner Lasertechnik). The spectra and intensity of the light obtained was measured by a spectroradiometer (Model LI-1800; Li-Cor, Lincoln, NE). Seedlings were grown for 4-10 d depending on the conditions, photographed with a scale and hypocotyl length was measured.

### 2.4.2 Transformation of Arabidopsis

#### 2.4.2.1 Growth of plants for transformation

*Arabidopsis* gi-2 mutant plants were grown to flowering stage as described (Chapter 2.4.1). Wild-type Col plants were grown in Kwik Pot trays (3 × 3 cell) secured with nylon mesh. Seed was sown directly into trays containing dampened soil. The trays were covered with mesh which was secured with tape. As plants formed primary inflorescences these were trimmed to 3-4 cm to encourage growth of auxiliary inflorescences from the rosette, which resulted in more floral buds per plant (Clough and Bent, 1998). Plants were transformed 5-10 days after trimming.

#### 2.4.2.2 Agrobacterium -mediated floral dip transformation
This transformation procedure is an adaptation of *Arabidopsis* vacuum infiltration (Bechtold *et al.*, 1993) by Clough and Bent (1998). For a detailed overview, see the review by Clough (2004). A single colony of *Agrobacterium* containing the desired binary vector was used to inoculate 50 mL of YN containing the appropriate antibiotic and incubated overnight at 28°C. The overnight culture was used to inoculate 2-3 L YN, which was incubated at 28°C until OD$_{600}$ was approximately 0.8. Cells were pelleted by centrifugation (4°C, 4400 xg, 15 min), the pellet resuspended in 1 L IM and placed on ice. Immediately before use 0.02% Silwet-77 (LEHLE SEEDS, Tucson, USA) was added to increase the transformation efficiency (Clough and Bent, 1998). Plasmid DNA was extracted from unused culture (Chapter 2.5.1.3) and analysed by restriction enzyme digestion (Chapter 2.5.3.5) to confirm the identity of the binary vector.

A Kwik Pot tray (3 x 3 cells) with *Arabidopsis* plants was inverted and placed into a shallow beaker holding the *Agrobacterium* suspension so that only aerial tissue was submerged. The beaker was sealed inside a vacuum desicator (Edwards RV12, Edwards High Vacuums International) and a vacuum (-80 kPa) was applied for 10 min and released quickly. Pots were placed in a flat tray covered with clear plastic to maintain humidity. When the plastic was removed (12-16 h later) plants were allowed to recover in LD growth cabinets, where they were grown to maturity and seed harvested.

### 2.4.3 Cross-pollination of *Arabidopsis*

Controlled pollination was performed to backcross blh3 and blh10 mutants to Col and to generate double and triple mutant plants. Floral buds, pre anthesis (stage 10-12; Bowman, 1994) were identified and outer floral organs were removed, leaving the receptive pistil intact. Pollen from freshly opened flowers of the "male" plant was brushed onto the stigma of the exposed pistil. Individual pistils were wrapped in clear plastic film to prevent cross-pollination and drying. To ensure pollination was successful this process was repeated the following day. Siliques were allowed to develop and dry naturally and seed from each was collected separately.

Plants were grown and allowed to self-fertilise through two generations and F$_2$ and F$_3$ progeny were analysed. Plants homozygous for the gi-2 mutation were identified by their late flowering phenotype, which was confirmed by PCR amplification of the region of gDNA spanning the 8 bp gi-2 deletion. Plants homozygous for the blh3 and/or blh10 T-
DNA were identified by PCR amplification of the genomic region of the BLH3 and BLH10 genes.

### 2.4.4 Analysis of Transgenic Plants

#### 2.4.4.1 Screening for Transformants

Seed was sterilised in 50 mL Falcon tubes, with a maximum of 27000 seed (540 mg) in one tube as described (Chapter 2.4.1.4). Seed were plated at a density of 4000 seed/plate on 15 cm GM plates containing the appropriate selection. Ampicillin and timentin were added to restrict bacterial growth. Following 3-5 d of cold treatment at 4°C the seeds were germinated at 20°C in the tissue culture room. Plants were grown under 16 hours of light (47 µmol m⁻²s⁻¹ fluorescent light) per day.

After 10-14 days of growth resistant seedlings were identified and transferred to soil. Plants were grown under plastic for 2-4 d to ensure high humidity and were maintained to maturity in the greenhouse. Seed was collected from each plant individually and stored at room temperature.

#### 2.4.4.2 Determination of T-DNA Copy Number

T₂ seed from individual plants lines to be analysed were sterilised and plated at a density of ~150 seed per plate onto small GM plates containing the appropriate selective antibiotic (Chapter 2.4.4.1). Alternately, blh3 and blh10 plants carrying the gene conferring resistance to the pesticide Basta were grown in soil in normal greenhouse conditions and sprayed with Basta (18% glufosinate, 10 mg/L) after 7 d and 14 d. Seedlings were allowed to grow for 10-14 days, at which time the ratio of antibiotic resistant seedlings (well developed roots, large green leaves) to antibiotic sensitive seedlings (small, yellowing, lacking roots and true leaves) was determined. Seedlings grown in soil that were sensitive to Basta were easily recognised as they died rapidly.

The segregation data were analysed by the Chi-square test to determine whether the observed segregation ratios fit the expected ratio for particular numbers of insertion events. A single dominant gene in the genome, when inherited in a Mendelian manner is expected to give a ratio of antibiotic resistant to antibiotic sensitive plants of 3:1. To determine if the observed segregation ratio represented the expected Mendelian segregation ratio the results
were analysed using the Chi-squared ($\chi^2$) test. $\chi^2$ was calculated using the following formula:

$$\chi^2 = \frac{(o_1 - e_1)}{e_1} + \frac{(o_2 - e_2)}{e_2}$$

\(o_1\) = observed number of sensitive seedlings, \(e_1\) = expected number of sensitive seedlings
\(o_2\) = observed number of resistant seedlings, \(e_2\) = expected number of resistant seedlings

The null hypothesis that "the seedlings were segregating in the expected ratio" was accepted at the 95% level if the $\chi^2$ value was less than 3.841, with one degree of freedom.

To identify lines homozygous for the T-DNA insert, T$_3$ seed from individual plants carrying single T-DNA inserts as determined above, were analysed for the presence of the appropriate resistance-conferring marker, as described above. The transgenic line was deemed homozygous if all T$_3$ progeny were resistant to the appropriate antibiotic or Basta herbicide.

2.4.5 TRANSIENT EXPRESSION OF GFP USING BIOLISTICS

Leaf tissue was harvested from wild type *Arabidopsis* plants and sterilised to minimise fungus contamination. Leaves were soaked in 0.02% Tween (2 min), 70% ethanol (5 sec) and 0.1% hypochlorite (10 min), followed by thorough rinsing in water. Leaves were blotted dry on sterile filter paper and placed on MS plates to recover. Alternately, onion epidermal peels were prepared and placed onto MS plates.

Ultra pure plasmid DNA to be used in bombardment was prepared using the Qiagen MidiPrep Kit following the manufacturer’s instructions. The eluate was extracted once with phenol:chloroform (Chapter 2.5.5.1) and concentrated to at least 1 µg/µL. Plasmid DNA (5 µg) was precipitated onto prepared gold microcarriers (3 mg) with 1 M CaCl$_2$ and 20 mM spermidine. The microcarriers were washed gently in 70% ethanol and twice in 100% ethanol. The gold microcarriers were resuspended in 48 µL 100% ethanol and aliquots of 6 µL (500 µg microcarriers) were used per ‘shot’ into plant tissue. The PDS-1000 He Biolistic Particle Delivery System (Bio-Rad Laboratories Ltd) was used to accelerate DNA coated gold particles into *Arabidopsis* leaf tissue and onion epidermis at 400 psi and 1100 psi respectively. Samples were stored in complete darkness for 48-76 hours after bombardment. GFP fluorescence was visualised using blue light (450-490 nm illumination, 525 nm short pass filter for GFP emission). To visualise nuclei, onion epidermal tissue was mounted in DAPI stain (1/2 MS, 20% glycerol, 2 µg/mL DAPI)
(from Staiger et al., 2003) and viewed using UV light (340-380 nm illumination, >425 nm emission).

### 2.4.6 GUS STAINING AND VISUALISATION

Control plants containing a dexamethasone (dex) inducible GUS construct (OP:GUS) were grown on GM plates and tissue was harvested 24 h after dex induction into 6-cell culture plates. X-Gluc based GUS staining solution (~ 2 mL) (Chapter 2.1.2) was added to each well and the culture plates were wrapped in tin foil to prevent the spread of the stain through the leaf tissue. The leaf tissue was incubated for 16 h at 37°C, followed by rinsing in 95% ethanol. Tissue was stored in 95% ethanol for 1-2 days to remove the chlorophyll and improve the visualisation of the stain. Stained tissue was mounted on standard microscope slides in 1% low melting point agarose for photography.

### 2.5 MOLECULAR TECHNIQUES

#### 2.5.1 PLASMID DNA ISOLATION

##### 2.5.1.1 ISOLATION OF *E. coli* PLASMID DNA

A single colony was used to inoculate 2-3 mL of Terrific broth containing the appropriate antibiotic and incubated at 37°C overnight. The next morning cells were pelleted by centrifugation (12000 \( \times g \), 30 s) and the supernatant removed by vacuum aspiration. The pellet was resuspended in 100 µL ice cold Solution I. Freshly prepared Solution II (200 µL) was added, the solution was mixed by gentle inversion of the tube and ice cold Solution III (150 µL) was added. Samples were incubated on ice for 5 min and cell debris was removed by centrifugation (12000 \( \times g \)) for 5 min. The supernatant was removed to a fresh microfuge tube and extracted with an equal volume of chloroform:IAA (Chapter 2.5.5.1). Total DNA was precipitated by addition of an equal volume of 100% ethanol and centrifugation (12000 \( \times g \)) at 4°C for 10 min. The pellet was rinsed in ice cold 70% ethanol, dried in air and resuspended in 40 µL TER. Yields of 150-250 ng/µL of plasmid DNA were obtained. Plasmid DNA was stored at -20°C.
2.5.1.2 **PURIFICATION OF PLASMID DNA FOR SEQUENCING**

To isolate high quality plasmid DNA for sequencing the Qiagen Miniprep Kit was used, following the manufacturer’s protocol, except that DNA was eluted in deionised water. Yields of ~ 200 ng/µL of plasmid DNA were obtained. DNA was stored at -20°C.

2.5.1.3 **ISOLATION OF PLASMID DNA FROM AGROBACTERIUM**

A single *Agrobacterium* transformant was inoculated into 10 mL of YN containing the appropriate antibiotics and incubated at 28°C for 24 h. Alternately, *Agrobacterium* culture remaining from plant transformation (Chapter 2.4.2) was used. The cells were pelleted by centrifugation (4000 xg, 4°C, 5 min), the pellet resuspended in ice cold Solution I (650 µL) containing 4.8 mg/mL lysozyme and incubated at room temperature for 30 min. Freshly prepared Solution II (1350 µL) was added and the solution incubated for a further 30 min. Following addition of ice cold Solution III the samples were mixed gently and incubated on ice for 5 min. Cell debris was separated from solution by centrifugation at (3500 xg, 4°C, 10 min), the supernatant extracted two times with an equal volume of phenol:chloroform and once with an equal volume of chloroform:IAA. DNA was precipitated by addition of two volumes of absolute ethanol and incubation at -20°C for 1-2 h. DNA was pelleted by centrifugation (12000 xg, 4°C, 20 min). The pellet was rinsed in 70% ethanol, dried in air and resuspended in 50 µL TER. An aliquot of the plasmid DNA was used to transform *E. coli* (Chapter 2.2.3.1) and was subsequently isolated for further analysis.

2.5.1.4 **ISOLATION OF PLASMID DNA FROM YEAST**

This method for isolation of yeast plasmid DNA was adapted from that of Hoffman and Winston (1987). Fresh yeast cells were used to inoculate 2-10 mL of appropriate culture media and were incubated 2-3 d at 30°C. The cells were collected by brief centrifugation and resuspended in 200 µL isolation buffer (2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl [pH 8], 1 mM Na₂EDTA). Glass beads (0.3 g) and phenol:chloroform (200 µL) were added to each tube, which were shaken at high speed on a vortex for 2 min. Following centrifugation at 12000 xg for 5 min the supernatant was removed to a fresh tube. DNA was precipitated by the addition of two volumes of absolute ethanol and 16 µL 5 M NH₄Ac, followed by centrifugation (12 000 xg, 10 min). The pellet was dried in air and resuspended in 20 µL of water. An aliquot of the plasmid DNA (1 µL) was used to transform *E. coli* (Chapter 2.2.3.1) and was isolated for further analysis.
2.5.2 ISOLATION OF NUCLEIC ACIDS FROM *Arabidopsis*

2.5.2.1 SMALL SCALE EXTRACTION OF GENOMIC DNA

The following method was used to isolate DNA from small tissue samples (1-2 leaves) and was used when screening for transformants. The concentration and quality of DNA obtained was suitable for PCR amplification.

*Arabidopsis* leaf tissue was harvested into microfuge tubes and snap frozen in liquid N₂. The tissue was then stored at -70°C or used immediately. Frozen plant tissue was placed in a pre-chilled microfuge tube under liquid N₂ and reduced to a fine white powder by grinding with a microfuge pestle using an electric hand drill. The powder was resuspended in 20% SDS (50 µL) and 750 µL of extraction buffer (50 mM EDTA, 10 mM β-mercaptoethanol, 500 mM NaCl, 100 mM Tris [pH 8]) by inversion. The samples were incubated at 65°C for 10 min. Potassium acetate (5 M, 250 µL) was added and samples were incubated on ice for 20 min. Cellular debris was pelleted by centrifugation (13000 ×g, 20 min). DNA was precipitated from the aqueous phase by the addition of isopropanol (500 µL) and incubation at -20°C (30-60 min). DNA was pelleted by centrifugation (13000 ×g, 10 min) and resuspended in 300 µL TE. Genomic DNA was stored at 4°C. The yield of *Arabidopsis* genomic DNA obtained was approximately 1 µg.

2.5.2.2 MINIPREP SCALE EXTRACTION OF GENOMIC DNA

This method was used to isolate larger yields and higher quality genomic DNA from approximately 0.5 g of *Arabidopsis* leaf tissue. Leaf tissue was harvested and ground to a fine powder as described (Chapter 2.5.2.1). The powder was resuspended in 560 µL of extraction buffer (50 mM EDTA, 500 mM NaCl, 100 mM Tris [pH 8]) and 20% SDS (40 µL). Samples were incubated at 65°C for 10 min. RNase A was added (30 µg) and the samples were incubated at 65°C for 1 h. Aliquots of 5 M potassium acetate (200 µL) were added and samples were incubated on ice (20 min). Debris was pelleted by centrifugation (13000 ×g, 10 min) and DNA was precipitated from the aqueous phase by addition of isopropanol (500 µL), followed by centrifugation (13000 ×g, 5 min). The loose pellet containing the DNA was washed with 70% ethanol (400 µL) and resuspended in 250 µL TE. Following the addition of 250 µL CTAB buffer (2% CTAB, 50 mM EDTA, 2 M NaCl, 200 mM Tris [pH 7.5]) the samples were incubated at 65°C for 1 h. The solution was extracted twice with chloroform:IAA and the DNA was precipitated by the addition of 100% ethanol (1 mL). To pellet the DNA, samples were centrifuged at 13000 ×g for 5 min.
The pellet was washed with 70% ethanol, dried in air and was resuspended in 100 µL TE. An approximate yield of 8-10 µg of genomic DNA was obtained.

2.5.2.3 Doyle and Doyle Extraction of Genomic DNA
This large scale extraction of DNA was adapted from that of Doyle and Doyle (1990) and was used to isolate DNA from plant tissue for use in Southern analysis.

Leaf tissue (5-6 g) was ground to a fine powder in a chilled mortar and pestle and scraped into 15 mL preheated CTAB extraction buffer (2% CTAB, 1.4 M NaCl, 0.2% β-mercaptoethanol, 20 mM EDTA, 100 mM Tris-HCl [pH 8]). Samples were incubated at 60°C for 30 min and the aqueous phase was extracted twice with chloroform:IAA. Ice cold isopropanol (10 mL) was added to the aqueous phase and DNA was precipitated at -20°C for 2-18 h. The DNA was recovered by centrifugation at 1000 ×g for 5 min. The pellet was resuspended in 25 mL wash buffer (76% ethanol, 10 mM NH₄Ac) and incubated at room temperature for 1-2 h. The DNA was again recovered by centrifugation at 1600 ×g for 20 min. The pellet was allowed to air dry and was resuspended in 3 mL TE. DNA was treated with RNaseA (30 ng) and incubated at 37°C for 30 min. Samples were diluted with 2 volumes of TE. DNA was precipitated by the addition of 2.5 volumes of absolute ethanol and 2.5 M NH₄Ac, followed by centrifugation at 10000 ×g (20 min, 4°C). The pellet was air dried and resuspended in 300 µL TE. An approximate yield of ~200 µg of genomic DNA was obtained.

2.5.2.4 Isolation of RNA from Arabidopsis
To avoid contamination and degradation of RNA, precautions were taken during the treatment of samples. Gloves were worn when handling pipette tips and microfuge tubes, plastic-wear was treated with H₂O₂ and glass-ware was baked at 200°C for 4 h. Solutions were prepared using RNase free chemical stocks, treated with 0.1% DMDC (v/v) and autoclaved. Solutions that contained Tris-HCl were prepared from DMDC treated and autoclaved RNase free stocks.

Leaf tissue was harvested into microfuge tubes and snap frozen in liquid N₂. Using a microfuge pestle and an electric power drill, the frozen tissue was ground to a fine powder. The powder was resuspended in extraction buffer (500 µL, 5 mM EDTA, 150 mM LiCl, 5% SDS, 80 mM Tris-HCl [pH 8]) and phenol:chloroform (500 µL) by vigorous vortexing
and placed on a shaker for 5 min. The samples were centrifuged for 5 min (13000 \times g), the aqueous phase was transferred to a fresh microfuge tube and extracted with phenol:chloroform. RNA was precipitated from the aqueous phase with isopropanol (500 µL) and 3 M sodium acetate (50 µL). The samples were centrifuged (13000 \times g, 4°C, 10 min) and the pellet containing RNA was washed with 70% ethanol, air dried and resuspended in 40 µL water.

An aliquot of RNA (20 µL) was diluted to 1 mL and the OD_{260} measured to estimate the concentration of RNA in solution. Samples were also checked by agarose gel electrophoresis (Chapter 2.5.3.2), to ascertain that the RNA had not degraded and that the concentration was consistent with that obtained by the OD_{260} results. RNA samples were stored at -70°C until required.

### 2.5.2.5 Purification of RNA

RNA was purified prior to use in RT-PCR. RNA was treated with 1 µL DNaseI (1 U/µL) in the presence of an RNase inhibitor (RNaseOUT™). RNA was incubated at 37°C for 30 min to degrade contaminating genomic DNA. RNA was then purified using the RNeasy™ Plant Miniprep Kit (Qiagen) according to the manufacturer’s instructions.

### 2.5.3 Manipulation and Examination of Nucleic Acids

#### 2.5.3.1 Enzymes and Markers

Invitrogen, Biolab Scientific, Life technologies, New England Biolabs, or Roche Molecular Biochemicals supplied DNA modification enzymes, restriction enzymes and associated buffers. Platinum Taq Polymerase was obtained from Invitrogen and the Expand High Fidelity PCR System from Roche. All standard markers, including low and high–mass DNA markers, the 1kb+ DNA size ladder and the RNA size ladder, were purchased from Invitrogen.

#### 2.5.3.2 Agarose Gel Electrophoresis

Nucleic acid fragments of different sizes were separated on horizontal agarose gels (0.7%-2% in 1 × TBE) submerged in TBE (Sambrook et al., 1989). DNA samples were mixed with 0.2 volumes of 5× electrophoresis loading dye, loaded into the wells of an agarose gel and electrophoresed for 40 min to overnight in a field of 1-6 Vcm⁻¹. Gels were
stained in ethidium bromide (~ 0.5 mgL⁻¹) for ~20 min and nucleic acids were visualised and recorded using a UVP Gel Documentation System Imagestore 7500, version 7.12.

For RNA samples, special precautions were taken to avoid the degradation of RNA (see Chapter 2.5.2.4). In contrast to DNA gels, agarose gels for RNA were prepared and submerged in TAE. Otherwise, the protocol was followed as described above.

2.5.3.3 Polyacrylamide gel electrophoresis (PAGE)
To separate and distinguish between small nucleic acid fragments of very similar sizes, DNA samples were separated on vertical acrylamide gels (8% acrylamide, 1 x TBE, 0.075% ammonium persulphate, 0.001% TEMED) submerged in 1 x TBE. DNA samples were mixed with 0.2 volumes of 5× electrophoresis loading dye, loaded into the wells of an acrylamide gel and electrophoresed for ~50 min at 200 V. Gels were stained in ethidium bromide (~ 0.5 mgL⁻¹) for ~20 min and nucleic acids were visualised and recorded using a UVP Gel Documentation System Imagestore 7500, version 7.12.

2.5.3.4 Dephosphorylation of vector DNA
To minimise religation of vector fragments, shrimp alkaline phosphatase was used to remove 5’ phosphate residues. Reactions were carried out as recommended by the suppliers. Dephosphorylated DNA was purified prior to use in ligations (Chapter 2.5.5.1).

2.5.3.5 Ligation of DNA
Ligation reactions were performed in volumes of 10 µL using 80-150 ng vector and insert DNA with molar ratios of vector to insert of 1:2, 1:3 and 1:5, except when ligating adaptor primers, where ratios of 1:5 and 1:10 were used. T4 DNA ligase (1 U/µL) and 5× ligase buffer, supplied by the manufacturer, were used. For ligations where larger DNA fragments were cloned, a concentrated T4 DNA ligase (5 U/µL) was used. Reactions were incubated overnight at 14°C. Aliquots of the reaction were used to transform E. coli as described in Chapter 2.2.3.1.

2.5.3.6 Restriction enzyme digestion
Restriction enzyme digests of plasmid DNA were performed in 10-50 µL volumes. Reactions contained 1-5 U of enzyme per µg DNA and the appropriate reaction buffer
suggested and supplied by the manufacturer. Reactions were incubated at 37°C unless otherwise specified for 2 h to overnight. For digestion reactions containing more than one enzyme a buffer was selected that optimised the activity of both enzymes as specified by the manufacturer.

2.5.3.7 QUANTIFICATION OF NUCLEIC ACIDS
Nucleic acid concentration was determined using the Nanodrop™ ND-1000 Spectrophotometer (Biolab Scientific Ltd). The concentration of DNA in a sample was also determined using electrophoresis through an agarose gel (Chapter 2.5.3.2), followed by staining with ethidium bromide and visualisation in ultraviolet (UV) light. The amount of fluorescence of a sample is proportional to the total mass of DNA, so the quantity of DNA was estimated by comparison of band intensity to bands of known concentration from DNA mass ladders (Chapter 2.5.3.1).

2.5.3.8 DNA SEQUENCING
Sequencing of prepared plasmids (Chapter 2.5.1.2) and PCR products (Chapter 2.5.5.3) was performed using the dideoxynucleotide chain terminator method developed by Sanger et al., (1977) on Applied Biosystems 377 sequencer (Perkin-Elmer, Applied Biosystems). Sequencing was carried out by K. Boxen (Centre for Gene Technology, School of Biological Sciences, University of Auckland).

2.5.4 POLYMERASE CHAIN REACTION (PCR)

2.5.4.1 PCR FROM DNA
Amplification reactions were carried out in 20-50 µL volumes containing genomic DNA (1-10 ng), plasmid DNA (5-50 pg) or first strand cDNA (1 µL) (Chapter 2.5.4.2) as a template. The reaction mixture included 1 × PCR buffer, 1.5-2.5 mM Mg, 1-3 U thermostable Platinum Taq polymerase (Invitrogen), 0.2-0.5 µM dNTPs (Roche) and 0.5 µM of the appropriate oligonucleotide primers (Chapter 2.1.5). For long or difficult PCR amplification reactions, HiFi Taq polymerase (Roche) was used according to the manufacturer’s instructions. When screening bacterial colonies for transformants a single colony was resuspended in 30 µL of water and 1 µL of the slurry was used as a template. PCR reactions were carried out in a GeneAmp PCR System 2400 thermocycler from
Chapter 2

Perkin Elmer and reaction products were separated by electrophoresis through an agarose or acrylamide gel (Chapter 2.5.3.2 and 2.5.3.3).

2.5.4.2 **Reverse Transcriptase (RT)-PCR**

The first step in RT-PCR was the synthesis of first strand cDNA from total RNA. cDNA synthesis was performed in 20µL volumes containing 5 µg total RNA (Chapter 2.5.2.5), 1 µg adaptor primer (Chapter 2.1.5), 1× First Strand buffer, 10 mM DTT, 1 mM dNTPs and DMDC-treated water. Reactions were incubated at 70 °C for 10 min and placed on ice briefly while 200 U SuperScript™ III RT was added. Samples were then incubated at 37 °C for 5 min, followed by an incubation for 1 h at 42°C. The enzyme was inactivated by incubation at 95°C for 5 min. Samples were diluted to 200 µL and used directly as templates for PCR (Chapter 2.5.4.1) or stored at -20°C.

2.5.4.3 **Genome Walking to Isolate Flanking DNA**

Genome walking was performed to isolate genomic DNA sequence flanking T-DNA insertions in the *blh3* and *blh10* transgenic plant lines. The primers and protocol used were modified from Spertini *et al.* (1999) by M. Yoon (this laboratory). The upper strand (us) and lower strand (ls) adaptors were mixed together and diluted to 50 µM. Adapter combinations were incubated at 75°C (10 min) and allowed to anneal by slow cooling to room temperature.

Genomic DNA (~150 ng) was digested by the appropriate restriction enzyme (Chapter 2.5.3.6), followed by incubation at 80 °C (20 min) to inactivate the enzyme. The annealed adaptor primers (2.5 µM) were ligated to ~5 ng digested DNA (Chapter 2.5.3.5). The ligation was diluted tenfold in TE and 2 µL used as a template for the first round of PCR amplification. A T-DNA specific primer and adaptor primer (ap1) were used. First round PCR products were diluted 1000-fold in TE and 2 µL used as a template in the second round PCR amplification. The T-DNA specific primer and a nested adaptor primer (ap2) were used. The PCR products were purified (Chapter 2.5.5.3) and analysed by sequencing (Chapter 2.5.3.8).

2.5.5 **Purification of DNA**

2.5.5.1 **Purification by Phenol:Chloroform Extraction**
The solution containing the DNA was mixed thoroughly with an equal volume of phenol:chloroform. The solution was centrifuged (12000 \( \times \) g, 5 min) and the aqueous phase was transferred to a fresh microfuge tube, where it was mixed with an equal volume of chloroform:IAA and centrifuged as before. Nucleic acids were precipitated from the aqueous phase with 0.1 volumes 3 M sodium acetate and 2 volumes 100% ethanol. The solution was incubated at -20°C for 20 min and the nucleic acids were pelleted by centrifugation (9500 \( \times \) g, 4°C, 10 min). The pellet was rinsed with 70% ethanol, vacuum dried and resuspended in 10-20 \( \mu \)L of water.

2.5.5.2 Extraction of DNA from agarose gels

Nucleic acid fragments produced by PCR or resulting from enzyme digestion of plasmid DNA were separated by agarose gel electrophoresis (Chapter 2.5.3.2) and stained with ethidium bromide. Fragments were visualised and excised from the gel under UV light. Nucleic acids were purified from the agarose slice using the Qiagen Gel Purification Kit, following instructions provided by the manufacturer. The purified DNA was used in ligation reactions (Chapter 2.5.3.5) or analysed by sequencing (Chapter 2.5.3.8).

2.5.5.3 Purification of PCR products

PCR products were purified to remove enzyme, salts and unused nucleotides using the Qiagen PCR product purification kit, following the instructions provided by the supplier. The purified products were analysed by sequencing (Chapter 2.5.3.8) or were used to prepare probes for hybridisation analysis (Chapter 2.5.6.2).

2.5.6 Hybridisation analysis

The method used for hybridisation analysis has been adapted from that described by Fourney et al. (1988).

2.5.6.1 Transfer of nucleic acids

RNA was size-fractionated by electrophoresis through a denaturing agarose gel (0.7 % agarose, 1 \( \times \) MOPS, 1.8 % formaldehyde). RNA samples (10-15 \( \mu \)g) were mixed with 15 \( \mu \)L RNA loading dye, denatured by heat treatment (65°C, 15 min) and electrophoresed in 1 \( \times \) MOPS buffer at 80-90 V. An RNA size ladder was also electrophoresed in every gel. RNA was stained with ethidium bromide (~0.5 mgL\(^{-1}\)) and visualisation in ultraviolet (UV) light. After photography, the RNA was transferred to a nylon membrane (Roche
Molecular Biochemicals) by capillary action. The RNA was coupled to the membrane by UV-crosslinking using a Stratalinker™ (Stratagene). The RNA bound to the membrane was visualised with an ethylene blue stain (0.04 % methylene blue, 0.5 M NaAc).

DNA was electrophoresed through an agarose gel (Chapter 2.5.3.2). The gel was treated in denaturation solution (30 min) and neutralisation solution (30 min). The DNA was then transferred to a nylon membrane (Roche Molecular Biochemicals) by capillary action. UV-crosslinking permanently fixed DNA to the membrane as described above.

2.5.6.2 **PREPARATION OF RADIO-LABELLED DNA PROBES**

DNA fragments to be used as hybridisation probes were isolated from agarose gels (Chapter 2.5.5.2) or generated by PCR (Chapter 2.5.4) and denatured before use by incubation at 100°C for 5 min. DNA was labelled with $[\alpha - ^{32}P]$ dCTP using the RadPrime DNA Labelling System (GibcoBRL) according to the supplier’s instructions. ProbeQuant™ G-50 Micro Columns (Pharmacia Biotech) were used to remove unincorporated nucleotides. The probe was denatured (100°C, 5 min) before being added to the hybridisation buffer (Chapter 2.5.6.3).

2.5.6.3 **HYBRIDISATION AND QUANTIFICATION OF GENE EXPRESSION**

Membranes were prehybridised (1 h, 65°C) in hybridisation buffer (15 mL) containing denatured salmon sperm DNA (100 mg/L). The probe DNA (Chapter 2.5.6.2) was added to the hybridisation buffer and allowed to hybridise to the membrane overnight (65°C). The membrane was rinsed twice in $2 \times$ SSC at room temperature. The membranes were then washed twice in 0.1% SDS and 0.1-0.5 $\times$ SSC for 20 min at 65°C. The membranes were rinsed in 0.1 $\times$ SSC before sealing in a plastic bag.

Expression levels of RNA were quantified using a phosphorimager. The radioactive signal was detected by exposure of membranes to a Fujifilm BAS imaging plate (10 min-10 d) and scanning the plate in a Fujifilm FLA-2000 phosphorimager using the Imagereader version 1.3E software. The visual images were quantified and printed using MacBAS version 2.5 software. Alternately, the radioactive signal was visualised by exposure to x-ray film (X-OMAT-K, Kodak).
To remove the previous probe, membranes were soaked in 0.1% SDS at 90-100°C and allowed to cool to room temperature. Membranes were re-hybridised with a 25s/26s rDNA probe (pTIP6, King and Davies, 1992) as described above. rRNA expression levels were quantified as above and used to normalise initial gene expression levels.

2.5.7 PROTEIN WORK

2.5.7.1 PROTEIN EXTRACTION FROM TOTAL PLANT TISSUE
Leaf tissue (~100 mg) was harvested into microfuge tubes and snap frozen in liquid N₂. Using a microfuge pestle and an electric power drill, the frozen tissue was ground to a fine powder. The powder was resuspended in ~200 µL Laemmli Sample Buffer and homogenised using a microfuge pestle. Samples were incubated for 5 min at 95°C and cellular debris was precipitated by centrifugation (14 000 ×g, 10 min). The supernatant was removed to a fresh tube and the protein concentration of each sample was determined by Bradford Assay (Bradford, 1976). Protein samples were used immediately or stored at -20°C for up to 1 week.

2.5.7.2 SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE)
Protein samples were diluted in Laemmli Sample Buffer and incubated at 95°C for 5 min before loading. Standard SDS-PAGE gels were used, consisting of a stacking gel (4% acrylamide, 125 mM Tris-HCl [pH 6.8], 0.1% SDS, 0.05% ammonium persulphate, 0.001% TEMED) and resolving gel (10-15% acrylamide, 375 mM Tris-HCl [pH 8.8], 0.1% SDS, 0.05% ammonium persulphate, 0.0005% TEMED). Products were separated by electrophoresis in 1 x Running Buffer at ~100 V and visualised by staining with Coomassie Blue.

2.5.7.3 WESTERN TRANSFER AND IMMUNODETECTION
Proteins to be analysed were resolved by SDS-PAGE and transferred to a Protan® nitrocellulose transfer membrane (Sartorius) by electroblotting, as described by Sambrook et al. (1989). Proteins were transferred to the membrane by electrophoresis either at 100 V for 1 h or 30 V overnight. The membrane was incubated with Ponceau Stain (0.2% Red Ponceau, 1% acetic acid) to visualise proteins and ensure transfer was successful. To block non-specific binding, membranes were incubated in BLOTTO for 1 h at room temperature or at 4°C overnight. Primary anti-BLH3 antiserum was diluted 1:1000 in BLOTTO and incubated with the membrane for 1-2 h at room temperature. The membrane was subject to
five wash steps, for 10 min each, in fresh TBS-T. The secondary anti-rabbit antiserum (conjugated to horseradish peroxidase) was diluted 1:5000 in BLOTTO and incubated with the membrane for 1 h at room temperature, followed by a repeat of the wash steps as above. Detection of the antibody was carried out using the ECL™ horseradish detection system (Amersham), following the manufacturer’s instructions. The presence of the BLH3 protein recognised by the anti-serum was detected by exposure of the membrane to x-ray film (X-OMAT-K, Kodak).

2.5.7.4 SYNTHESIS OF RADIO-LABELLED PROTEIN

Radio-labelled GI protein was synthesised in vitro for use in pull down assays. Highly concentrated plasmid DNA (>1 µg/µL) was prepared for use as a template for transcription using the Qiagen Midi-Prep Kit according to the manufacturer’s protocol. Protein was labelled with L-[^35S] methionine (Redivue, Amersham Pharmacia) using the TnT Quick Coupled Transcription/Translation System (Promega). In order to analyse the translation product, an aliquot was diluted with Laemmli sample buffer and separated by SDS-PAGE (Chapter 2.5.7.2). The gel was fixed for 30 min in 50% methanol/10% acetic acid and a further 30 min in Amplify Reagent (Amersham Pharmacia). The gel was soaked for 5 min in 7% methanol/7% acetic acid/1% glycerol to prevent the gel from cracking before drying under vacuum using a gel dryer (DrygelSr SE1160, Hoeffer Scientific Instruments). Radio-labelled products were visualised using autoradiography, with exposure times of 1 – 7 days at -70°C. The size of the protein was estimated by comparison to a ^14C -Labelled Protein Molecular Weight Standard (Gibco BRL).

2.5.7.5 PURIFICATION OF MBP FUSION PROTEINS

A single BL21-DE3 clone containing the appropriate pMAL-based plasmid was used to inoculate 2 mL of Rich Broth (RB) and was incubated overnight at 37°C. An 800 µL aliquot of the overnight culture was used to inoculate 80 mL RB and was incubated, with shaking, at 37°C until the OD_{600} reached 0.5. To induce MBP expression 300 µM IPTG was added to the culture, which was then incubated for a further 2 h. The cells were harvested by centrifugation at 4000 xg for 10 min and the pellet was resuspended in 5 mL chilled Column Buffer (20 mM Tris-HCl (pH 7.4), 0.2 M NaCl, 1 mM EDTA, 10 mM β-mercaptoethanol, added just before use). The cells were frozen in Column Buffer at -20°C overnight. The cells were gently thawed in an ice cold water bath and sonicated in 4 x 20 sec pulses (40 Amps). The crude extract (supernatant) and insoluble fraction (pellet) were
separated by centrifugation (20 min, 4°C, 9000 xg). The insoluble fraction was resuspended in 5 mL cold Column Buffer and both the crude extract and the insoluble fraction were stored in aliquots at -20°C.

MBP fusion proteins were purified from the crude extract (and the insoluble fraction as a control) using amylose resin (New England Biolabs). Amylose resin (50 µL per sample) was washed 2 x in 1.5 mL Column Buffer. The resin was resuspended in 50 µL Column Buffer, 50 µL of crude extract was added and samples were incubated on ice for 15 min to facilitate the binding of the MBP to the amylose resin. Samples were centrifuged briefly and the supernatant discarded. The pellet containing the protein bound to the resin was washed 2 x with 1 mL Column Buffer. The resin was resuspended in 50 µL Laemmlli Sample Buffer and incubated at 95°C for 5 min to release the protein bound to the resin. Samples were centrifuged briefly to pellet the resin and 20 µL of the supernatant was analysed by SDS-PAGE. Translation products were visualised with Coomassie stain and quantified.

Bulk protein purification was necessary for use in the inoculation of rabbits for antibody production. The protocol was as described above, but cells were harvested from larger 80 mL overnight cultures and cells were resuspended in 20 mL Column Buffer. Protein was isolated from 10 mL crude extract onto Poly-Prep Columns (Biorad) and following washing with 20 mL Column Buffer, eluted in a total of Column Buffer + 10 mM maltose. The final concentration of protein acquired was ~1 mg/mL.

**2.5.7.6 IN VITRO PROTEIN BINDING ASSAY**

Recombinant MBP (43 kDa) and MBP:BLH3 (~78 kDa) protein fusions were expressed in *E. coli* strain BL21-DE3, purified using amylose beads and quantified (Chapter 2.5.7.4). GI protein labelled with 35S-Methionine was synthesised *in vitro* from 1 µg plasmid DNA (Chapter 2.5.7.3). MBP (5 µg) and MBP:BLH3 (5 µg) purified on amylose beads were incubated in binding buffer (150 mM NaCl, 2 mM EDTA, 20 mM Tris-HCl pH 7.5, 0.1% Nonidet P-40) (Bhalerao *et al.*, 1999) with radio-labelled GI protein (25 µL) at 4°C for 1 hour with gentle agitation. Amylose beads alone were also incubated with the labelled GI protein as a negative control. The beads were washed 6 times with ice cold binding buffer and resuspended in 20 µL Laemmlli buffer. The proteins were released from the beads by boiling and separated by SDS-PAGE. Radio-labelled GI was visualised by
autoradiography and the size of the protein estimated by comparison to a $^{14}$C-Labelled Protein Molecular Weight Standard (Gibco BRL).

2.6 COMPUTER ANALYSIS

2.6.1 GENERAL COMPUTING

General computing was carried out on a Dell Pentium IV using the Microsoft XP operating system. Word processing, basic drawing and graphing were performed using Microsoft Word 2001, Microsoft Powerpoint 2001 and Microsoft Excel 2001 (Microsoft Corporation). The programs Amplify version 1.2 (© W. R. Engels) and Oligo version 4.0-s (© W. Rychlik, 1991) were used for oligonucleotide primer design and were carried out on an Apple Power Macintosh G3 using the Mac OS 9.0 operating system.

2.6.2 SEQUENCE ANALYSIS

DNA sequences were compared and analysed using the Wisconsin Package Version 9.1, (Genetics Computer Group [GCG], Wisconsin), ClustalX 1.64 (Thompson et al., 1997) and BLAST (GenBank [www.ncbi.nlm.nih.gov]) (Altschul et al., 1997) programs. Arabidopsis genomic sequences were examined using TAIR SeqViewer (www.arabidopsis.org). DNA sequence data was analysed and assembled using the Vector NTI Advance program, ContigExpress® module (Informax™, Invitrogen life science software).

2.6.3 PHYLOGENETIC ANALYSES

A total of forty-three protein sequences were used in phylogenetic analyses. Thirty-seven of these, including thirteen Arabidopsis genes, were derived from cDNA sequences. The remaining six were predicted protein sequences from the annotated Arabidopsis and rice genome sequences. The Drosophila homeodomain protein DmANTP (M20704) was included as an outgroup. Initial analyses also included the plant KNOX protein OsOSH1 (D16507) as an outgroup. The sequences and Genbank accession numbers used for the sequence analyses are located in Appendix 1.2. The amino acid sequences encompassing the three conserved domains and including an extra nine highly conserved amino acids outside the homeodomain region (BLH3 aa 167-417) were used. Amino acid sequences
were aligned using ClustalX 1.64 (Thompson et al., 1997) and the protein matrix MacClade 3.08 (Sinauer Associates Inc., Sunderland, MA, USA). Construction of phylogenetic trees was carried out using PAUP 4.0b3 (Sinauer Associates Inc.). Bootstrap replicates of 1000 were employed to support the tree and only branches with bootstrap values greater than 70% were retained.

2.6.4 MICROSCOPY, PHOSPHORIMAGES AND PHOTOGRAPHS

Tissue was examined for GFP expression using a Leica DMRE microscope and images were captured using a Leica DC500 digital camera. For the visualisation of fluorescent images a mercury vapour lamp was used as the light source and light was filtered through a GFP filter (excitation 450-490 nm; emission ~525 nm). Images of leaf tissue exhibiting GFP fluorescence were also captured using a Leica TCS 4D confocal scanning microscope. A 488 nm laser was used and GFP was detected at ~500 nm.

Phosphorimages were analysed and quantified using the MacBAS version 2.5 program (Fuji Photo Film Company). Photographs were scanned using an Agfa Snapscan 1212 (USB version) and image size and contrast were manipulated using Photoshop™ version 4.0 (Adobe) software. The content of these images was not altered during this process.

2.7 COMPLIANCE

The creation, importation and genetic modification of plants, bacteria and yeast performed as part of this thesis research complied with Environmental Risk Management Authority regulations and were carried out under approval number GM099/UA020 (updated in 2000).
CHAPTER THREE: **The BLH3 Homeodomain Protein Interacts with GI**

### 3.1 Introduction

Previously, a putative GI protein interactor was identified in a yeast 2-hybrid screen. It was found that this protein was a member of the BELL family of plant specific transcription factors and was identical to the BELL-LIKE HOMEO DOMAIN 3 protein (BLH3).

In the first part of this chapter further yeast 2-hybrid assays and *in vitro* protein assays are undertaken to confirm the interaction of GI with BLH3. The second part of this chapter reports on the analysis of the BELL family. Sequence and phylogenetic analyses were employed to identify BELL family members, to analyse the structure of these proteins and establish their relationship to each other. Finally, BELL-LIKE HOMEO DOMAIN 10 (BLH10), the BELL protein most closely related to BLH3, was examined further.

#### 3.1.1 Previous Work

A yeast 2-hybrid screen was performed by Kim Snowden (this laboratory) to isolate proteins that are capable of interacting with GI. Classical late flowering *gi* mutants encode truncated proteins that lack the C-terminal region, indicating that this region is important for GI function. Consequently, a DNA fragment encoding the C-terminal portion of the GI protein (aa 740-1173) was used to form the prey vector pAS:GI. The *Arabidopsis* expression library CD4-22 (Kim *et al.*, 1997) was screened for protein interactions. Four positive clones were identified in the original screen. Three clones encoded unknown proteins which contained repeated myosin-like helical domains that have been known to mediate non-specific protein interactions. For this reason these clones were not investigated further as part of this thesis work.
Sequence analysis established that the fourth clone contained a partial cDNA fragment which encoded the homeodomain protein BLH3 and that this cDNA was in frame with the activation domain in pACT (Carly Pullen, this laboratory). This clone, pACT:BLH3 (encoding aa 180-524) was predicted to encode the majority of the BLH3 protein. Interestingly, a second longer BLH3 clone was also isolated from the 2-hybrid screen (encoding aa 62-524). However, reconstruction assays showed that this longer clone interacted weakly with the empty vector pAS2.1 (Kim Snowden). This type of false positive is thought to occur when the cDNA clone within pACT interacts with the HA tag in the pAS2.1 vector (Gietz et al., 1997). This result indicated that BLH3 contains a region between aa 62-180 that can mediate a non-specific protein interaction. On this basis, the smaller clone encoded by pACT-BLH3 was used in all further assays in this thesis project.

For use in further interaction assays, two GI fragments were amplified from pAS:GI and inserted into pGEMT (pGEMT:GI; pGEMT:Del4) by Trent Bosma (this laboratory).

### 3.2 Testing the GI and BLH3 Protein Interaction

#### 3.2.1 Identification and Repair of a PAS:GI Error

The sequence of the pGEMT:GI clones (Chapter 3.1.1) were analysed before use in further cloning steps and it was evident that all clones contained the same sequence error. This deletion of one base pair at nucleotide 3661 near the 3′ end of the GI cDNA brought into frame a premature stop codon 10 bp downstream. This resulted in the loss of 59 amino acids from the C-terminal of the GI protein.

Further sequencing of pAS:GI revealed that the one base pair deletion was present in the original pAS:GI vector (Figure 3.1a). The error was repaired by replacing this region of GI with a GI cDNA fragment from pBS:GI. The GI fragment was released from pBS:GI by digestion with NcoI and XbaI and inserted into NcoI/XbaI digested pAS:GI, replacing the GI sequence containing the error (Figure 3.1b). Sequencing of pAS:GI and the GI cDNA used for the repair and comparison of these to the predicted GI sequence confirmed that the one base pair deletion was no longer present in the GI sequence. The repaired pAS:GI vector was used for the rest of this thesis work.
Figure 3.1 The detection and repair of a deletion in pAS:GI

a The constructs pAS:GI and pBS:GI were sequenced using the oli12 primer. The sequences obtained were compared to the genomic sequence of GI. The region of interest is highlighted by the black box and the site of the deletion is indicated by the arrow above the pAS:GI sequence.

b Strategy for the repair of the deletion in pAS:GI. The pAS:GI construct was subject to restriction enzyme digestion by NcoI and XbaI, releasing the 1.3 kb fragment of GI containing the deletion error. The equivalent 1.3 kb fragment of GI was purified from pBS:GI after digestion with NcoI and XbaI and inserted into NcoI/XbaI digested pAS:GI.

Further work with the reconstructed pAS:GI vector in yeast was challenging as it was not possible to maintain the yeast containing this vector for periods longer than 1 week. The pAS:GI plasmid had to be re-transformed regularly into PJ69 4a and used immediately for interaction assays. The pAS:GI plasmid was isolated from yeast and the identity was confirmed by restriction enzyme digestion (data not shown). However, subsequent sequencing of the plasmids isolated from yeast by Revel Drummond (this laboratory) revealed that they contained a mixed population, including intact pAS:GI and pAS:GI vectors containing a new single base pair deletion in the 5’ end of GI. The cause of this error in the GI sequence is unclear, although it is suggestive that pAS:GI is unstable in
3.2.2 Confirmation of the GI::BLH3 Protein Interaction

In order to confirm the GI::BLH3 protein interaction and test the interaction with the repaired pAS:GI vector (Chapter 3.2.1), the yeast 2-hybrid interactions performed previously were reconstructed. The GI (pAS:GI) and BLH3 (pACT:BLH3) vectors were re-transformed into the PJ69 yeast mating strains 4a and 4α respectively.

The yeast 2-hybrid interactions were replicated and the results obtained verified that the GI and BLH3 proteins interact with each other (Figure 3.2). The control plasmids pVA3-1 and pTD1-1 (Clontech) were also included in the interaction assay. These strongly interacting clones are provided as a positive control with the Matchmaker 2-Hybrid System (Clontech). Additionally, they act as negative controls, as neither pAS:GI nor pACT:BLH3 should interact with these proteins. The 2-hybrid results obtained verify that pACT:BLH3 does not interact with a control bait protein (pVA3-1) or the empty vector, but interacts specifically with pAS:GI (Figure 3.2). These results also show that GI does not interact non-specifically with pACT alone nor pACT containing another protein (pTD1-1). These results were observed with three independent repeats and demonstrate that the GI::BLH3 protein interaction is specific in a yeast 2-hybrid assay.

**Figure 3.2 Yeast 2-hybrid assay showing interactions between GI and BLH3**

The prey vector (pAS:GI), encoding a GI peptide (aa 740-1173) fused to the GAL4 BD, was transformed into the yeast strain PJ69 4A. The pACT:BLH3 prey plasmid (aa 180-524) was transformed into the opposite mating strain PJ69 4α. Following mating, diploid yeast containing both bait and prey plasmids were selected on SC media lacking tryptophan and leucine (SC-trp-leu). Clones exhibiting adenine prototrophy were identified by replica plating onto SC media lacking adenine (SC-ade). The interaction between pVA3-1 and pTD1-1 was used as a positive control. Empty bait and prey vectors pAS and pACT were included as negative controls.
3.2.3 INVESTIGATION OF THE GI::BLH3 INTERACTION BY IN VITRO PROTEIN INTERACTION ASSAY

The yeast 2-hybrid assay is a useful system for detecting putative protein interactors; however it is important to test these protein interactions outside of the yeast cellular environment. Despite considerable improvements in the 2-hybrid system over recent years that minimise false positives, it is possible for these assays to detect indirect interactions. To verify the yeast 2-hybrid results and confirm the direct physical interaction of GI and BLH3, a standard in vitro binding assay was employed. The system used requires one protein partner to be fused to Maltose Binding Protein (MBP) and expressed in E. coli. The recombinant proteins are then affinity purified on amylose beads. The second protein partner is radio-labelled by expression in vitro from the T7 promoter in the presence of $^{35}$S-labelled methionine. The steps required to generate these proteins and the subsequent interaction assays are described in the following sections.

3.2.3.1 CREATING MBP FUSION CONSTRUCTS

Three constructs were generated using the pMAL expression vector to produce recombinant BLH3 and GI proteins fused to MBP for expression and purification in E. coli. The purified proteins were later used as input for in vitro interactions (Chapter 3.2.3.4). In brief, partial BLH3 and GI cDNA fragments were amplified by PCR using adaptor primers to introduce restriction enzyme sites onto the end of the PCR products. The restriction sites were engineered so that the BLH3 and GI genes could be cloned in frame with MBP in the pMAL expression vector in subsequent cloning steps. These cloning steps are described in more detail below.

First, the partial BLH3 cDNA was subcloned from pACT and inserted into pMAL to create a MBP:BLH3 fusion. The BLH3 clone (aa 180-524) was amplified from pACT:BLH3-1 using the pACT.fwd and pACT.rev primers, which contain flanking restriction enzyme sites onto the end of the PCR products. The restriction sites were engineered so that the BLH3 and GI genes could be cloned in frame with MBP in the pMAL expression vector in subsequent cloning steps. These cloning steps are described in more detail below.
the insert with respect to MBP were determined by restriction enzyme digestion (Figure 3.3a and b). A single clone containing BLH3 in the correct orientation was selected and the junction between the BLH3 and MBP genes was subject to sequence analysis to ensure that BLH3 was fused in the correct frame (Figure 3.3c). The pMAL:BLH3 construct was used to transform E. coli for the expression of recombinant protein (Chapter 3.2.3.2).

![Restriction maps of pMAL:BLH3, with the BLH3 insert cloned in the forward and reverse orientations with respect to MBP. SspI restriction enzyme sites and diagnostic restriction fragments, used to confirm the orientation of the insert, are indicated by the black arrows (▼).](image)

![Digestion with SspI was used to verify the orientation of the BLH3 insert. Fragments indicated by black arrows (1.0 kb, 3.6 kb) are common to both orientations. Fragments labelled with red arrows are unique to clones with the insert present in the incorrect orientation (1.9 kb, 1.5 kb, clone 1) and blue arrows are unique to clones where the insert is present in the correct orientation (2.6 kb, 0.8 kb, clones 2 and 3). 1kb: 1 kb+ DNA size marker.](image)

![The pMAL:BLH3 construct was sequenced at the junction between the BLH3 and MBP genes to ensure that the BLH3 gene was in frame. The sequence was translated and the protein sequence compared to the expected MBP and BLH3 sequences. The direction of the protein coding sequence is indicated by the arrows. The EcoRI restriction enzyme site used for cloning is highlighted by the green box.](image)

In order to test the GI::BLH3 protein interaction with the opposite protein partner fused to MBP, two GI deletions were also cloned into pMAL. Analysis of truncated GI proteins revealed that a small GI deletion (Del4, aa 740-810) interacted with BLH3 (see Chapter 4 for yeast 2-hybrid deletion analysis). For this reason, Del4 and the partial GI sequence
from pAS:GI were selected to be fused to MBP. The 1.5 kb and 0.3 kb inserts were released from pGEMT:GI and pGEMT:Del4 respectively (Chapter 3.1.1) by digestion with EcoRI and BamHI and inserted into the pMAL vector. This created a Maltose Binding Protein (MBP) fusion to the N-terminus of GI (pMAL:GI) and Del4 (pMAL:Del4). The identity of the pMAL:GI and pMAL:Del4 plasmids were confirmed by restriction enzyme digestion (data not shown).

### 3.2.3.2 Expression of MBP Fusion Proteins

The expression vectors pMAL, pMAL:GI, pMAL:Del4 and pMAL:BLH3 were transformed independently into *E. coli* BL21-DE3 cells (Chapter 2.2.3), which facilitate the expression of recombinant proteins. The expression of the recombinant proteins was induced by IPTG, followed by incubation for 2 h at 37°C. Cell extracts were prepared and MBP:BLH3 was identified in the soluble fraction of the cellular extract. However, a significant amount of the recombinant protein was also present in the insoluble fraction. Recombinant MBP:GI and MBP:Del4 proteins were not detected (data not shown).

In an attempt to maximise the amount of soluble recombinant protein, cells were allowed to grow at a lower temperature of 30°C for longer incubation times of 4 h, 6 h and 20 h following induction with IPTG. The 4 h incubation time proved adequate for the induction of suitable levels of MBP:BLH3 and MBP control (Figure 3.4a, ii and iii). The lower incubation temperature was successful, as the majority of the MBP:BLH3 fusion was present in the soluble cellular extract (ce column, Figure 3.4).

![Figure 3.4](image.png)

**Figure 3.4 Expression and quantification of MBP:BLH3 fusion proteins**

**a** MBP and the MBP fusion proteins MBP:Del4, MBP:BLH3 and MBP:GI were induced in *E. coli* and the crude soluble cellular extract (ce) was prepared. The recombinant proteins containing MBP were also separated from the crude extract on amylose beads (pu). The two samples were separated by SDS-PAGE and visualised by Coomassie staining. The sizes of the MBP and MBP:BLH3 proteins were estimated by comparison with a standard protein molecular weight ladder (L). i; MBP:Del4, ii; MBP:BLH3, iii; MBP

**b** Quantification of MBP:BLH3 and MBP proteins. Samples of MBP:BLH3 and MBP as
Again, the MBP:GI and MBP:Del4 proteins were not detected at any reasonable level following induction with IPTG (Figure 3.4a). A very small amount of MBP:Del4 may be detected after purification (Figure 3.4a, i), although this was certainly not enough to be used in a protein interaction assay. Furthermore, it is impossible to resolve the size of the MBP:Del4 fusion using SDS-PAGE. The Del4 protein is only 7 kDa, as a result MBP:Del4 (50 kDa) and MBP control (43 kDa) appear to be the same size. For this reason it was difficult to determine if the MBP:Del4 fusion was successful.

The sequences of the clones were previously assessed in pGEMT and the GI and Del4 inserts were in the correct frame for cloning with EcoRI and BamHI. The pMAL:GI and pMAL:Del4 plasmid DNA was extracted from E. coli and analysed by restriction enzyme digest. The results indicated that the expected pMAL:Del4 plasmid was present (data not shown). In contrast, restriction digestion of pMAL:GI produced fragments of unexpected size, signifying that an aberrant pMAL:GI plasmid was present in E. coli. This very likely explained the lack of expression obtained from this construct. As the MBP:BLH3 protein fusion was successful at this time, it seemed sensible to move forward and use only the MBP:BLH3 proteins for further interaction assays. The pMAL:GI and pMAL:Del4 clones were not used any further.

The fusion protein MBP:BLH3 and the MPB control were affinity purified from the soluble fraction on amylose beads (Chapter 2.5.7.5). Purified MBP:BLH3 was separated by SDS/PAGE and quantified by comparison against a BSA loading control (Figure 3.4b). The size of the recombinant protein was slightly smaller than expected (predicted to be ~75 kDa), however it was larger than the MBP control (43 kDa). It is not unusual for recombinant proteins to migrate differently to standard protein preparations through SDS-PAGE gels, depending on the conditions. A 5 µg standard of purified MBP:BLH3 and MBP control were used for each subsequent in vitro interaction assay.

3.2.3.3 Expression of radio-labelled proteins in vitro

The interaction system used in this work employs an in vitro transcription/translation kit to radio-label a protein interacting partner (Chapter 2.5.7.4). In this case, the BLH3 partner was successfully expressed and purified in E. coli, therefore it was necessary to express and radio-label the GI partner. Three GI clones were selected for translation in vitro: the
first a full length GI cDNA encoding the entire GI protein, second the truncated C-terminal GI clone that was used as prey in the initial yeast 2-hybrid screen and lastly, the smallest GI deletion Del4 was used.

In order to translate the GI clones in vitro, they were required to be inserted downstream of the T7 promoter. The C-terminal GI clone and the Del4 deletion were amplified from pAS:GI using the pASEcoRI.fwd and pASBamHI.rev primers, resulting in 1.5 kb GI and 0.3 kb Del4 products. These PCR products were subcloned into the pGEMT vector. To ensure that the GI and Del4 inserts were in the correct orientation with respect to the T7 promoter, positive clones were subject to PCR (Figure 3.5). The resulting pGEMT:GI and pGEMT:Del4 constructs were analysed by sequencing to ensure that they contained no errors introduced during PCR (data not shown).

The GI and Del4 products in pGEMT encoded their own start sites for in vitro protein translation. The pGEMT vectors were unable to be used as templates for in vitro translation as they contained start sites upstream of the start codon in the GI and Del4 inserts. Fortunately, the orientation of the inserts with respect to the T7 promoter were the same as that required for cloning into the pBS vector. The GI and Del4 fragments were released from pGEMT by digestion with EcoRI and BamHI and inserted directionally into pBS to create pBS:GI/R and pBS:Del4 respectively. Positive clones were selected and their structure was verified by restriction enzyme digestion (data not shown).

**Figure 3.5 Verifying the orientation of GI inserts in pGEMT by PCR**

a  Diagram outlining the PCR strategy for identifying clones containing GI inserts in the correct orientation with respect to the T7 promoter. The primers used, oli5 and M13R, are indicated by the black arrows.

b  DNA was extracted from nine positive clones (1-9). A 1:100 dilution was used as a template for PCR amplification (Chapter 2.5.4) using the oli5 and M13R primers. The presence of the expected 500 bp product indicated that two clones contained the insert in the correct orientation. 1kb; 1 kb+ DNA size marker.
The constructs pBS:GI/R, pBS:Del4 and the full length GI cDNA (in pBS:GI) were selected as templates for in vitro translation reactions. The first attempts at producing protein were not successful, likely due to RNase contamination of the plasmid templates and a lower than optimal concentration of DNA templates. An alternative purification protocol was utilised to produce high concentration, ultra pure plasmid DNA for use as templates. Plasmid DNA was extracted using the Qiagen Midiprep kit and purified by phenol:chloroform extraction (Chapter 2.5.5.1). The three GI proteins were translated in vitro from 1 µg plasmid DNA, incorporating $^{35}$S-Methionine to generate radio-labelled proteins. The amount and size of the GI proteins obtained were analysed by SDS-PAGE (Figure 3.6).

The translation of both GI and GI/R was successful, although not as efficient as the luciferase control (Figure 3.6). The full length GI protein is predicted to be >100 kDa and is not resolved well by these gels. As a result it is difficult to estimate the size of the product obtained, although it is larger than the 97 kDa size marker. The GI/R protein is predicted be 47 kDa and at ~30 kDa is slightly smaller than expected (Figure 3.6), although this is not unusual. Of concern was the fact that there was also a smaller product of ~20 kDa present, though at much lower concentrations. It is possible that this is a GI protein product that resulted from translation at an alternate start site within pBS:GI/R. It is likely that the translation of Del4 was successful, but the expected product is less than 10 kDa and is too small to be resolved by these gels. Regrettably, Del4 cannot be separated from the unincorporated nucleotides at the bottom of the gel (Figure 3.6). For this reason Del4 was not used in any in vitro interaction assays.

![Figure 3.6](image)

**Figure 3.6  Expression of radio-labelled GI proteins in vitro**

The GI proteins and a luciferase control were translated in vitro and separated by SDS-PAGE (Chapter 2.5.7.2). Proteins were visualised by exposure of the gel to film for 36 h. **luc**, luciferase control; **GI**, full length GI protein >100 kDa; **GI/R**, GI/R protein ~30 kDa; **Del4**, Del4 protein <10 kDa; **L**, radio-labelled protein ladder.
3.2.3.4  PROTEIN INTERACTION ASSAYS

The proteins generated in Chapter 3.2.3.2 and 3.2.3.3 were expressed and purified on the day of use for protein interaction assays. Radio-labelled full length GI protein was incubated with extracts containing MBP:BLH3 and MBP. Amylose beads alone were also incubated with the labelled GI protein as a negative control. Bound protein was released from the beads, separated by SDS/PAGE and visualised by autoradiography. Figure 3.7a shows the retention of labelled GI protein only with purified MBP:BLH3. No interaction was observed with purified MBP or amylose beads alone indicating that the interaction is not due to non-specific binding by GI to the components of the binding matrix. This experiment was repeated and similar results were obtained (data not shown).

The in vitro interaction was performed using the smaller clone GI/R, encoding the C terminal region of the GI protein. This truncated version of GI initially showed interactions with both MBP:BLH3 and MBP alone, although significantly more interaction with MBP:BLH3 than with MBP alone (Figure 3.7b). In an attempt to increase the specificity of the GI/R interaction the experimental conditions were altered. The interaction was repeated with an increased salt concentration in the wash buffer to provide a more stringent wash. This resulted in the very efficient removal of all radio labelled GI/R from both the MBP control and MBP:BLH3 (data not shown). This may be suggestive of a weak interaction between GI and BLH3.

![Figure 3.7](image-url)

**Figure 3.7  BLH3 and GI interact in vitro**

Purified MBP:BLH3 recombinant protein and MBP control were quantified against a BSA loading control (5 µg) (input gels). MBP and MBP:BLH3 were incubated with ^35^S-methionine labelled GI translation product (~120 kDa) and GI protein bound to MBP:BLH3 was detected by autoradiography (pulldown gels). Amylose beads (Am) were incubated with labelled GI protein as a negative control. L: standard protein molecular weight ladder.

- **a** The GI protein binds the MBP:BLH3 fusion protein in vitro.
- **b** The GI/R protein binds the MBP:BLH3 fusion protein with greater affinity than to MBP.
As the signal obtained from GI after the interaction was weak, an attempt to increase the signal was made by adding double the amount of radio labelled GI to the purified MBP:BLH3. This led to an increase in signal, however GI was also present in the MBP control sample, though at much lower levels compared to the samples containing MBP:BLH3 (Figure 3.7c). This was most likely due to a surplus of GI protein, which was unable to be removed by the series of six wash steps. It should be noted that even with a considerable excess of radio-labelled protein, the GI protein bound to MBP:BLH3 significantly more than to MBP alone and did not bind to the amylose bead matrix (Figure 3.7c).

In summary, the GI and BLH3 proteins interact in yeast 2-hybrid and these interactions appear to be specific to these two proteins. Furthermore, the BLH3 and GI proteins interact outside the yeast environment in \textit{in vitro} interaction assays, providing substantial evidence that this is a legitimate protein interaction.
3.3 A STUDY OF THE BELL PROTEINS

3.3.1 INTRODUCTION

Previously, the GI protein interactor BLH3 was revealed to have homology to the Arabidopsis BELL1 homeodomain protein (Chapter 3.1.1). When this thesis project began the BELL1 protein was the only BELL-like protein sequence available in the Genbank database. Additional putative BELL-like proteins were identified and analysed as sequence data became available. Alignment of these proteins revealed information on protein structure, such as conserved motifs within the sequence. Phylogenetic analyses of the BELL proteins were used to provide information about the evolutionary relationships within the gene family. These results revealed relationships between closely related pairs and groups of genes, which were particularly informative when deciding which gene(s) may be functionally similar to BLH3 and thus useful to work with in future experiments.

3.3.2 IDENTIFICATION OF BELL-LIKE PROTEINS

Initially, mining the Arabidopsis genome (AGI, 2000), thirteen BELL genes were identified. The thirteen Arabidopsis BELL sequences are the putative amino acid sequences that correspond to known cDNAs, therefore all are expressed in Arabidopsis. Pairwise comparisons of the thirteen full length Arabidopsis BELL proteins were calculated using ClustalX and presented in a square matrix (Appendix 1.1). The results show that the BELL proteins have a relatively broad range of divergence and that the similarity between proteins is generally quite low. In fact, the majority of the proteins have less than 40% identity with the other BELL proteins and the most similar proteins, BLH2 and BLH4, share only 66% identity. BLH3 shares between 21-61% amino acid identity with the 12 Arabidopsis BELL genes, with the predicted BLH10 protein the most similar (Appendix 1.1).

In order to identify all possible members of the BELL family, the predicted amino acid sequence of BLH3 was used to search the NCBI and TIGR databases for similar proteins via a series of BLAST searches (Altschul et al., 1997). Newly identified members of the BELL family, particularly those not highly similar to BLH3, were also used as query sequences in order to increase the probability of identifying more diverse BELL proteins.
A total of forty-five amino acid sequences representing the BELL family of homeodomain genes were obtained from ten plant species. These proteins have been identified in diverse plants species, including tomato, rice and the gymnosperm *Gnetum*. Numerous putative BELL proteins were identified in the TIGR Gene Indices, from tree plants such as poplar and pine to crop species such as onion and cotton. These were not used for analyses here as they were generally represented by small partial EST sequences.

The majority of BELL sequences from species other than *Arabidopsis* correspond to known cDNA sequences in Genbank, therefore are known to be expressed genes. However, six of the rice sequences used in this work are predicted protein sequences from the annotated rice genome sequences. Five of these rice BELL genes were supported by cDNA clones, but these did not encode the full length protein or encoded an altered protein compared to that predicted by the genomic annotation. For the OsBELL14 gene no cDNA sequences were found in the Genbank database. Alternatively spliced cDNA clones representing another four rice BELLs were identified. In the case of OsBEL, the cDNA encoded a full length protein that included all three conserved domains. This was in contrast to the genomic annotation, which predicted a smaller protein that lacked the SKY domain and part of the BELL domain. In this case, it is likely that the annotation of OsBEL in the database is incorrect and demonstrates the necessity of obtaining cDNA sequence to confirm intron/exon boundaries and the coding regions of genes.

### 3.3.3 Analysis of BELL Protein Structure

To analyse the structure of the BELL proteins, a series of multiple sequence alignments were created using the GAP program (GCG). The two most similar proteins from *Arabidopsis* (BLH10, 68% identity and BLH7, 36% identity) and the most similar from rice (OsBELL5, 50% identity) were aligned with BLH3 (Figure 3.8). An alignment of all forty-five BELL proteins identified was also performed (Appendix 1.3).

These alignments revealed the structural characteristics that define the BELL family. Three conserved regions are used to classify BELL proteins; the SKY domain, BELL domain and homeodomain (Figure 3.8).
Figure 3.8 Comparison of the predicted amino acid sequence of four BELL proteins

The BLH3 and BLH10 homeodomain proteins from *Arabidopsis* were aligned with BLH7 and BELL5, two similar proteins from *Arabidopsis* and *Oryza sativa* respectively. Features of the sequence are as indicated: identical residues are shaded black and residues conserved in at least three sequences are grey; the SKY domain is underscored by the red box; the green box underlines the BELL domain; the homeodomain motif is underlined by the grey box; putative amphipathic α-helical regions are indicated above the sequence; low complexity regions within BLH3 are indicated by blue lines; repeats of three or more glutamine or serine residues within BLH3 are in bold lettering; conserved TALE residues (PYP) are indicated by the arrow; the QGLSLSL and VSLTLGL boxes are highlighted in yellow. At, *Arabidopsis thaliana*; Os, *Oryza sativa*.

It is interesting to note that approximately 30% of the BELL proteins in the alignment contain SKY and/or BELL domains that are not particularly well conserved (Appendix 1.3). In fact, there are no amino acid residues within the SKY or BELL domains that are absolutely conserved in all BELL proteins. The SKY and BELL domains are predicted to contain α-helices and these are marked above the sequence. It has been proposed that these α-helices are amphipathic, meaning hydrophobic residues are aligned down one side of the helix. The output of multiple protein prediction programs, such as Jpred, PHD and
SSPRED (ExPASy proteomics server, http://ca.expasy.org/) support this theory. These programs suggest that hydrophobic residues lie down the face of helices within the SKY and BELL domains; therefore it is probable that these helices are amphipathic. The third conserved domain is the DNA binding homeodomain (HD) region, consisting of three α-helices. There are 21 residues within the homeodomain that are absolutely conserved in all 45 BELL proteins, including the three amino acid loop extension (PYP) that define the TALE class of transcription factors (Appendix 1.3).

Outside of the three conserved domains there is little amino acid conservation between the 45 BELL proteins, possibly due to lower constraints on function. There are small regions of amino acid identity within subgroups of proteins, particularly between the BELL and homeodomain regions (Appendix 1.3). Examining the BELL proteins of potato, Chen et al. (2003) recognised a motif C-terminal of the homeodomain which consists of seven amino acids and named it the VSLTLGL box. Analysis of the complete alignment of the BELL proteins show that this conserved VSLTLGL box is present in 33 of the BELL proteins identified (Appendix 1.3). Close examination of the sequence of the BELL family revealed that there is a similar motif located in the region of the protein N-terminal to the SKY domain. It also consists of seven amino acids (QGLSLSL), although it is not as well conserved as the VSLTLGL box. This motif is present in 70% of proteins containing the VSLTLGL box. Interestingly, the QGLSLSL box is present in the BLH3 and BLH10 proteins, although they lack the VSLTLGL box (Figure 3.8). The function of these regions is unknown.

Examination of BLH3 with programs which recognize protein motifs (ExPASy proteomics server, http://ca.expasy.org/) revealed that BLH3 also contains two regions comprising repeats of glutamine and serine residues (Figure 3.8). Glutamine rich regions and homopolymer repeats are proposed to act as transcriptional activation domains in some organisms (Gerber et al., 1994; Lincoln et al., 1994), which is particularly interesting as BLH3 is a putative transcription factor.

### 3.3.4 A Phylogenetic analysis of the BELL family

The evolutionary relationship between the BELL proteins was established by phylogenetic analysis, with the aim of providing more information on BLH3 and closely related genes on which to base further experiments. A total of forty-three unique BELL proteins
identified in Genbank were used, representing eight plant species (see Appendix 1.2 for BELL sequences and gene identifier numbers). This is the most extensive BELL phylogeny to date. Previous phylogenies have not included all of the *Arabidopsis* proteins and only a limited number of proteins from other species (Becker *et al.* 2002; Chen *et al.* 2003). The two BELL sequences from the TIGR gene indices were not included as they had not yet been identified.

The amino acid sequences encompassing the three conserved domains and including nine extra conserved amino acids outside the homeodomain (corresponding to BLH3 aa 167-418) were aligned using ClustalX and the protein matrix MacClade (Chapter 2.6.3). The *Drosophila* homeodomain protein DmANTP was included as an outgroup. Heuristic and Neighbour Joining trees based on distance analysis were constructed and both yielded the same clade distribution. The support for the trees in the data was estimated using bootstrap analysis. Only branches with greater than 70% bootstrap support after 1000 replicates were retained. The rice KNOX gene OsOSH1 was also included in the initial phylogenetic analyses to ensure that the identified proteins were true members of the BELL family, not the related KNOX family. All BELL proteins grouped together with high bootstrap support, with the KNOX protein as an outlier (data not shown). These phylogenies provide additional evidence that the BELL proteins are a family that are likely to have a common evolutionary history and therefore related functions.

The BELL family separates into seven groups (Figure 3.9), which are arbitrarily numbered I to VII. These clades are well supported by bootstrap values of $\geq 73\%$. The *Arabidopsis* BELL genes are represented in all seven clades. The largest groupings are clades IV and V, which contain eight BELL proteins each. The majority of the BELL proteins contain a 7 amino acid motif, the VSLTLGL box, located toward the C-terminal of the protein (Chapter 3.3.3). This VSLTLGL motif is found within proteins in clades I-V, with two exceptions. Interestingly, BLH3 and BLH10 do not contain the VSLTLGL box, suggesting that it has been lost in BLH3 and BLH10. Of these proteins in clades I-V, 68% also contain the less conserved 7 amino acid motif, the QGLSLSL box at the N-terminal of the protein. The proteins grouped within clades VI and VII do not contain the conserved VSLTLGL box within the N-terminal region of the protein and also lack the QGLSLSL box, with one exception LeBL1. Proteins in clades VI and VII may have diverged from other BELL proteins before the evolution of these conserved boxes. The significance of these conserved regions is unknown.
Figure 3.9 Phylogram of predicted BELL family proteins

The sequences of 43 BELL proteins, including 13 Arabidopsis BELL proteins, were used to construct a heuristic tree. The Drosophila homeobox protein DmANTP was used as an outgroup. Amino acids spanning the three conserved protein domains (BLH3 amino acids S167-G418) were used and bootstrap values on each branch are based on 1000 trials. The seven clades are labelled with roman numerals. At, Arabidopsis thaliana; Dm, Drosophila melanogaster; Gg, Gnetum gnemon; Hv, Hordeum vulgare; Le, Lycopersicon esculentum; Md, Malus x domestica; Os, Oryza sativa; St, Solanum tuberosum; Zm, Zea mays.

Proteins from Arabidopsis and rice are present in all clades, therefore it is likely that the BELL gene family existed before the divergence of eudicots and monocots. The single protein from a non-flowering plant, GgMELBEL1, does not group within any of the seven clades, however it is placed well within the BELL tree (Figure 3.9). A previous phylogeny revealed that GgMELBEL1 was located within the BELL1 clade (Becker et al., 2002). The additional sequences included here and the stricter requirement of >70% bootstrap support for branches would have affected this clade distribution in our phylogeny.
3.3.4.1 BLH3 AND BLH10 GROUP TOGETHER

The BLH3 and BLH10 proteins group together in clade V with very strong bootstrap support (Figure 3.9). High sequence identity and relatively short branch lengths suggest that BLH3 and BLH10 are the result of a duplication or expansion event. Phylogenetic trees were also constructed using data from parsimony analysis. Clades I, III, IV and VII were supported by parsimony trees, although the three other clades were split (data not shown). However, BLH3 and BLH10 remained grouped together with strong bootstrap support.

These results have implications for potential loss-of-function experiments involving BLH3 and BLH10, as it is possible that these genes may compensate for each other. In addition, it is possible that BLH6 and BLH7, which are also grouped in clade V, may overlap functionally with BLH3/BLH10. In this case the production of triple (or quadruple) mutants may be necessary to determine the function of the proteins in this clade. In contrast, BELL proteins that are not closely related may also share a common function with BLH3. Preliminary data revealed that the over-expression of BLH2 is able to partially compensate for the loss of BELL1 in bel1 mutants (G. Haughn, pers. comm.). This indicates that the BLH2 and BELL1 genes have overlapping functions although these proteins are not particularly closely related. BLH2 and BELL1 are grouped in clades I and III respectively (Figure 3.9). Only two genes grouped in clade V with BLH3 have been characterised; the potato BELL StBEL30 and OsBIHD1 from rice. StBEL30 is expressed in all tissues in potato, with the highest levels detected in flowers (Chen et al. 2003). OsBIHD1 is thought to function in disease response in rice as the transcript is induced in response to fungal infection and treatment with benzothiadiazole, an agent known to induce disease resistance (Luo et al., 2005).

3.3.5 GENOMIC ORGANISATION OF THE BELL FAMILY

Analysis of the complete Arabidopsis genome sequence (The Arabidopsis Genome Initiative, 2000) provided clues to the evolutionary history of the BELL family. In this section, genomic approaches have been used to further illustrate the relationships of the BELL genes in Arabidopsis, including the expansion of the family and the divergence of individual genes such as BLH3.
Figure 3.10 Genomic organisation of the BELL family genes

A schematic of the five *Arabidopsis* chromosomes (labelled I-V) is shown with the thirteen *Arabidopsis* BELL genes are marked by the black arrows. Genes and chromosomes are not drawn to scale. Genes are coloured white if they show no homology to each other at the amino acid level, with the direction of the arrows indicating the orientation on the chromosome. The coloured arrows indicate genes that encode proteins with homology to the following proteins: RING finger proteins; WD repeat containing proteins; phosphoglyceride transfer proteins; wound inductive proteins; expressed protein 1; glycosyl transferases; LON domain proteases; bZIP proteins; purine transporter family; hypothetical protein; purple acid phosphatases; expressed protein 2; THY proteins.
The BELL gene family and flanking genes were mapped and their relationships determined. The thirteen *Arabidopsis* BELL genes are distributed on chromosomes I, II, IV and V, but are absent from chromosome III (Figure 3.10). The BELL genes are not clustered, as has been observed for some gene families within the *Arabidopsis* genome (AGI, 2000). Assessment of up to eight genes on either side of the BELL gene loci was carried out to determine if BELL genes had been subject to small duplication events. These genes were compared to genes flanking other BELL genes using the BLAST algorithm and many of these show 25-82% identity at the amino acid level (Figure 3.10) (Appendix 1.4).

Of particular interest are the *BLH3* and *BLH10* loci. Shared gene pairs are located around *BLH3* and *BLH10*, with a number of shared pairs also located around the *BLH6* and *BLH7* loci. Both *BLH3* and *BLH10* are located on chromosome I at approximate positions of 28.3 Mb and 6.8 Mb respectively. Using BLAST searches it was found that six genes flanking the *BLH3* loci share 38-82% identity at the amino acid level to genes flanking the *BLH10* loci (Figure 3.11a) (Appendix 1.4). Furthermore, analysis of the genomic sequence of these genes demonstrated that the *BLH3* and *BLH10* genes have a highly conserved intron/exon structure (Figure 3.11b). Both genes have four exons and four introns, including one leader intron in the 5’ UTR. The results suggest that these two genes are the result of a duplication of a region of chromosome 1. Another BELL gene, *BLH11*, is also located close to *BLH3*; however the presence of this gene is difficult to explain as it shares greater sequence homology with other BELL proteins.

Four genes flanking the *BLH3* and *BLH10* loci also have high sequence identity at the amino acid level to genes flanking *BLH6* and *BLH7* (Figure 3.10). The BLH6 and BLH7 proteins themselves have high sequence identity and group together within clade V, as do BLH3 and BLH10. It is likely that the duplication events involved in expanding clade V were also followed (or preceded) by a small duplication event between chromosomes II and IV, as genes are shared between the *BLH6* and *BLH7* loci that are not shared with the *BLH3* and *BLH10* loci on chromosome I (Figure 3.10). This is supported by results obtained by Blanc *et al.* (2003), which show that the *BLH6* and *BLH7* loci correspond to a duplication block.
Figure 3.11 Schematic diagram of the positions of BLH3, BLH10 and flanking genes on chromosome I.

**a** Genes sharing $\geq 35\%$ identity at the amino acid level are connected by lines. Arrows indicate genes sharing identity with genes flanking other BELL genes BLH6 and BLH7. The ID numbers on flanking genes, preceded by At1g, are AGI identification numbers. Not drawn to scale. **b** The BLH3 and BLH10 genes have a similar genomic structure. Exons are shaded black, introns are grey and non-coding 5’ and 3’ UTR are striped.

### 3.4 Analysis of the BLH3 and BLH10 Genes

Results from the previous sections suggested that BLH3 and BLH10 are highly similar, particularly over the three conserved domains and consequently may share some similarity in function. For this reason both genes were selected for further investigation. Clones containing the entire open reading frame of BLH3 and BLH10 were obtained by screening of Arabidopsis expression libraries. At the time of screening, no full length BLH3 or BLH10 cDNA clones were present in the Genbank database. These cDNA clones were necessary for in use future experiments (Chapters 4, 5 and 6). Comparison of the BLH3 and BLH10 genes revealed more similarities, supporting the decision to investigate both of these genes in future experiments.

#### 3.4.1 Identification of the BLH3 cDNA

Previously, the Arabidopsis expression library CD4-7 (D’Alessio et al., 1992) was screened for BLH3 by Carly Pullen (this laboratory). A BLH3 cDNA of 1946 bp in length was isolated (AY598452) (Appendix 1.5). This cDNA includes 147 bp 5’ UTR and 226 bp 3’ UTR and is predicted to encode a 524 aa protein (60 kDa).
At the time, no other BLH3 cDNA sequences were available in the Genbank database. During this thesis work, the BLH3 clone was compared to sequences in the Genbank database and other BLH3 cDNA clones were identified. Comparison of the BLH3 sequence to a 1950 bp clone (AF353093) revealed that the BLH3 clone contained a 114 bp longer 5' UTR and 116 bp shorter 3' UTR. A third cDNA clone (AY062545) contains a 57 bp longer 3' UTR and an 82 bp shorter 5' UTR than the cDNA clone identified here. Recently, a fourth cDNA clone was released. This 1824 bp clone (AY085278) contains a 28 bp 5' UTR and 222 bp 3' UTR, both of which are shorter than those in the BLH3 cDNA. These comparisons suggest that BLH3 has at least four polyadenylation sites.

### 3.4.2 IDENTIFICATION OF THE BLH10 cDNA

#### 3.4.2.1 WHICH FORM OF THE BLH10 cDNA IS EXPRESSED IN PLANTS?

Due to the close phylogenetic relationship between BLH3 and BLH10 and the evidence that these genes are the result of a duplication event, the BLH10 gene was also analysed in this work. At the time, the only cDNA clone available in Genbank database corresponding to BLH10 was a RIKEN Arabidopsis Full-Length (RAFL) cDNA clone (2059 bp) (AY052366). The translation of the RAFL cDNA encoded a truncated protein. Nineteen conserved residues within the BELL domain were altered, followed by the introduction of a premature stop codon resulting in the loss of the rest of the BELL domain and the entire homeodomain region (Figure 3.12a). Analysis of the sequence of the RAFL cDNA clone revealed it was lacking part of what was predicted to be exon 2, most likely due to the alternative splicing of intron 2 (Figure 3.12b).

To establish which BLH10 transcript was expressed in plants and hence which cDNA clone should be used for further experiments, the expression of BLH10 in plants was determined. An RT-PCR protocol was developed to determine if the RAFL clone was expressed in plants (Figure 3.12b). The results show that only the BLH10 cDNA predicted by the annotation of the BLH10 gene was present in wild type plants (Figure 3.12c). The 800 bp product representative of the predicted BLH10 cDNA was detected at all times of the day tested. The 650 bp product indicative of the presence of the RAFL clone was not detected, confirming that this transcript is not expressed in plants grown in our standard conditions.
Which BLH10 transcript is expressed in plants?

**a** Translation of the cDNA sequence at the predicted intron/exon boundary. BLH10 cDNA; part of the conserved BELL domain is shown, RAFL cDNA; encodes 19 altered residues and a premature stop codon as highlighted.

**b** Schematic diagram of three of the possible BLH10 transcripts that may be expressed in plants. 1, gDNA, intron 2 not spliced; 2, predicted annotation of cDNA, intron 2 spliced; 3, RAFL cDNA, intron 2 and part of exon 2 spliced. Black arrows represent internal BLH10 specific primers.

**c** Analysis of BLH10 expression by RT-PCR. Aerial tissue was harvested from plants of the 6 leaf stage at the times after dawn indicated. Total RNA (4µg) was used as a template for reverse transcription, followed by amplification with BLH10 specific primers. Amplification products of BLH10 of the expected size were visualised on an agarose gel. The Ubiquitin (UBQ) RT-PCR was used as an internal control.

Surprisingly, an 890 bp product was detected at two time points (Figure 3.12c). This product suggests the presence of gDNA or an unspliced cDNA clone containing intron 2. The likelihood of contamination with gDNA is small as RNA samples were pretreated with DNase before use as templates in RT-PCR and these samples were subject to control PCR amplification reactions without the prior reverse transcription step (data not shown). As a precaution, the 800 bp and 890 bp PCR products were subject to direct sequencing. Analysis of the sequence established that both products were amplified from the expected region of BLH10, and that the 890 bp product contained intron 2. These results suggest that cDNA clones different to the RAFL clone in the database are present in plants; as a result cDNA libraries were screened in an attempt to identify these alternate BLH10 clones.
3.4.2.2 SCREENING A CDNA LIBRARY

Four cDNA libraries were selected to undergo preliminary screening for the presence of BLH10 clones. Internal BLH10 specific primers were used to amplify samples of the CD4-6 (Wiegel et al., 1992), CD4-7 (D’Alessio et al., 1992), CD4-15 (2-3 kb) and CD4-16 (3-6 kb) (Keiber et al., 1993) libraries. The expected product was obtained for three out of four samples, demonstrating that BLH10 was present in all libraries tested except CD4-7. To increase the possibility of isolating a full length clone, primers designed to the 5' UTR of BLH10 were used to amplify BLH10 from the remaining three libraries (Figure 3.13a). It was found that only CD4-15 included clones that contained the BLH10 5' UTR. Therefore, the CD4-15 library was selected and 5 x 10⁵ pfu were screened for the presence of BLH10. Initially, the radio-labelled BLH10 probe used to screen this library was designed to correspond to the predicted 5' UTR of BLH10 (Figure 3.13b). The resulting first round screen of the library resulted in no convincing positive plaques. As the PCR results verify that BLH10 is present in this library, it was likely that the negative result was obtained because a BLH10 cDNA with a long 5' UTR was not present.

A second BLH10 probe was amplified from genomic DNA, corresponding to the following coordinates relative to the A of the translation start codon; -55 to +294 (Figure 3.13b). The screening was successful and twenty positive plaques were selected from the primary round. Dilutions of these were subject to PCR using internal BLH10 specific primers to determine that the positive clones detected were not identical to the alternately spliced RAFL clone. All twenty putative positives contained the correctly spliced BLH10 cDNA (data not shown).
Three positive plaques were selected (7.1, 9.1, 12.1), plated at 100 pfu per plate and subject to hybridisation with the BLH10 probe. Hybridisation to ~20 plaques over the three plates was detected. Five positive plaques were selected and the phage excised. The plasmid DNA was isolated from two independent clones representing each of the five positive colonies (Chapter 2.2.5) and the cDNA inserts were released by restriction enzyme digestion with XbaI/XhoI. The resulting restriction pattern indicated that the size of the cDNA inserts ranged from ~1.3 kb to 4 kb, with the size of BLH10 expected to be ~2 kb (Figure 3.14a).

**Figure 3.14 Identification of clones containing BLH10**

a Restriction enzyme digests of five positive clones with XbaI/XhoI. The cDNA insert is released from the 3 kb pBS vector. Each positive is represented by two independent clones.

b A schematic drawing of the pBS:BLH10 vector is outlined below. Key restriction sites are indicated. Not to scale.

c Restriction enzyme digestion of BLH10 cDNA clone 9.1.2. Plasmid DNA was digested with the following restriction enzymes and products of the expected size were obtained. HindIII; 3100 bp, 1100 bp, 500 bp, 160 bp, KpnI; 3200 bp, 1700 bp, PstI; 3400 bp, 1500 bp, KpnI/PstI; 2900 bp, 1250 bp, 400 bp, 250 bp. Fragment sizes are approximate.

1 kb; 1 kb+ DNA size marker with 1, 2 and 3 kb markers indicated.

Sequencing of the five clones confirmed that three clones contained the BLH10 cDNA. Two clones containing the 1.9 kb and 1.3 kb inserts encoded non-related proteins and were discarded. The two clones with larger 4 kb inserts contained BLH10 and the insertion of an extra gene and were also discarded. Only clone 9.1.2 contained the BLH10 cDNA of the expected size and this was named pBS:BLH10. The presence and orientation of BLH10 with respect to pBS was confirmed by analysis of further restriction enzyme digests using restriction sites expected to be present in BLH10 (Figure 3.14b and c).
3.4.2.3 CHARACTERISING THE BLH10 cDNA

The BLH10 cDNA isolated in Chapter 3.4.2.2 was fully sequenced and the contiguous sequence assembled (Genbank number AY570508) (Appendix 1.6). The BLH10 clone was 1906 nucleotides long and encodes a 538 aa protein of approximately 61 kDa. This clone includes 76 bp 5' UTR and 215 bp 3' UTR. The BLH10 protein sequence was compared to the annotated BLH10 gene in Genbank and it was established that this clone encoded the full length protein as predicted by the genomic annotation. Later an additional BLH10 cDNA clone (BX813490) was identified in the Genbank database. The 3' UTR of this cDNA clone is 42 bp shorter and the 5' UTR 89 bp longer than the BLH10 clone isolated during this work. This result indicates that BLH10 has alternate polyadenylation sites.

3.4.3 WHICH ATG IS THE TRANSLATION START SITE?

The predicted translation start sites of BLH3 and BLH10 are strong plant initiation consensus sequences (Lutcke et al., 1987), so it is predicted that upstream sequences in the cDNA clones isolated are part of the 5' UTR and not the coding region. The 5' UTR of BLH10 has in-frame stop codons upstream of the putative translation start site and no open reading frames of significant length in the alternate reading frames (Figure 3.15). This provides evidence that the correct start codon within BLH10 has been identified.

<table>
<thead>
<tr>
<th>BLH3</th>
<th>ttaaacaacaaaaacctttccgatttattattccaatagctgttgtattaccctaatagtgcggc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame 1</td>
<td>L N N K T L S I Y Y P M A V Y Y P N S V G</td>
</tr>
<tr>
<td>Frame 2</td>
<td>* T T K H F R F I I Q W L C I T L I V S</td>
</tr>
<tr>
<td>Frame 3</td>
<td>K Q Q N T F D L L S N G C V L P * * C R</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BLH10</th>
<th>ctgtaataaaaacttttaaaaaggtgtaaaccaaatagcagtttattaccacaagtagtgcggt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame 1</td>
<td>L * * N F * K V * T P M A V Y Y T S N V G</td>
</tr>
<tr>
<td>Frame 2</td>
<td>C N K T F K R C K H Q W Q F I T Q V M S</td>
</tr>
<tr>
<td>Frame 3</td>
<td>V I K L L K G V N T N G S L L H K * C R</td>
</tr>
</tbody>
</table>

Figure 3.15 Comparison of the translation start site in BLH3 and BLH10

The BLH3 and BLH10 cDNA clones were translated in all three reading frames. The predicted ATG codon and protein sequence is highlighted in bold type. Only frame 1 contained suitable open reading frames. The translation of BLH10 in frame 2 contains a stop codon 54 residues downstream from the possible start site (not shown).
The 5' UTR of *BLH3* does not contain any start codons upstream of the predicted start site, nor does it contain upstream stop codons. This is also true for an alternate clone with a longer 5' UTR (AY062545). Yet sequence similarity suggests that both the *BLH3* and *BLH10* genes start at equivalent sites (Figure 3.15), which provides support that the identified translation start site is correct.

### 3.4.4 Analysis of the Sequences Upstream of *BLH3* and *BLH10*

*GI* is a circadian clock controlled gene which is up-regulated in response to poor light in shade conditions. As BLH3 has been shown to interact with the GI protein, it was important to investigate if *BLH3* expression may be controlled in a manner similar to *GI*. Likewise *BLH10*, due to the high degree of identity this gene shares with *BLH3*. Devlin *et al.* (2003) detected particular hexamer sequences that are over-represented in the promoter regions of genes that are light responsive and the *GI* promoter contained 14 of these. Based on this discovery, genomic regions directly 5' to the *BLH3* and *BLH10* translation start sites and up to 1000 bp upstream were examined. It was found that *BLH3* and *BLH10* contain 14 and 13 hexamer sequences respectively, some of which were repeated up to 4 times. Of the 14 hexamers detected in sequences upstream of *BLH3*, 7 were identical to hexamer sequences in the *GI* promoter region (data not shown).

The *BLH3* and *BLH10* genes are closely related and group together with *BLH6* and *BLH7* in phylogenetic analyses (Chapter 3.3.4). Hence, sequences upstream of *BLH6* and *BLH7* were included in this investigation to determine if the promoter sequences of these genes shared similarities with *BLH3* and *BLH10*. The Transcription Regulatory Element Search (TRES) (www.bioportal.bic.nus.edu.sg) seek out common palindrome motifs upstream of gene start sites which have unique features of symmetry and the ability to form hairpins or loops. It is thought that conserved palindromes may function in transcriptional regulation. The ~1000 bp upstream of the transcription start sites of *BLH3*, *BLH10*, *BLH6*, *BLH7* and *GI* were analysed and compared using this tool. Nineteen palindromic sequences were identified that were common to all four BELL proteins, some repeated up to nine times. Ten of these are summarised below (Table 3.1). These results suggest that not only are sequence elements common in the promoter regions of all four BELL genes tested, but that these genes also share possible promoter motifs with *GI*. The possibility that *BLH3* and *BLH10* are transcriptionally regulated in a manner that is similar to that of *GI* was examined further in Chapter 5.
### Table 3.1 Identification of repeated motifs in the upstream sequences of GI and four BELL genes

The sequence up to 1000 bp upstream of GI and the four BELL genes were analysed for common palindromic sequences using the TRES sequence analysis tool.

<table>
<thead>
<tr>
<th>Palindrome Sequence</th>
<th>GI</th>
<th>BLH3</th>
<th>BLH10</th>
<th>BLH6</th>
<th>BLH7</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAATTT</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>TTTAAA</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>AAANTTT</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>TTTNAAA</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>AAANNTTT</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>TTTNAAA</td>
<td>4</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>AAANNNTTT</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>TTTNNAAA</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>AAAANNNTTT</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>
| TTTNNNNAAA          | 2  | 5    | 2     | 0    | 3    

### 3.5 DISCUSSION

This chapter presents work performed to establish the BLH3 protein as a GI protein interactor, followed by the examination of the BELL family of proteins. The interaction between the BLH3 and GI proteins was confirmed using a yeast 2-hybrid assay. The BLH3 and GI proteins interacted exclusively, as neither protein interacted with control proteins. Yeast 2-hybrid is a useful system for identifying potential protein interactors, however this system occasionally generates false positives. It is also possible that yeast proteins mediate or assist in the protein interactions. For these reason it is crucial to confirm that the protein interaction(s) detected in the yeast system are able to be replicated using an alternative method. For this thesis project an in vitro co-immunoprecipitation assay was carried out to confirm the ability of proteins to interact outside of the yeast environment. The selected in vitro assay allows for the precise control of physical and chemical conditions in the protein environment. The results of in vitro interaction assays supported the yeast 2-hybrid findings as BLH3 and GI continued to interact outside of the yeast cellular environment. Ideally, one would demonstrate the interaction of two plant proteins in vivo, as this would strongly suggest that the protein interaction was biologically functional in plants. For this purpose expression analyses were employed to determine if BLH3 and BLH10 overlapped with GI in plants. These experiments are described in Chapter 5.
3.5.1 THE BELL FAMILY

Mining the *Arabidopsis* genome revealed that BLH3 is one of 13 BELL proteins in *Arabidopsis*. This family is characterised by the presence of three conserved domains; the SKY domain, BELL domain and homeodomain. Further examination of the BELL family of proteins by phylogenetic and sequence analyses provided support that the BELL genes identified have a common evolutionary history and possible similarity in function. The conserved regions and sequences unique to the BELL family were able to be identified using multiple sequence alignments. These also proved useful for analysis of proteins closely related to BLH3. Comparison of the BLH3 and BLH10 proteins using the BLAST algorithm had shown that these proteins were closely related. Further sequence alignments using the amino acids that span the three conserved protein domains (aa 167-408) of BLH3 and BLH10 show that these domains are particularly highly conserved, sharing 81% identity. The BELL proteins PNY and PNF share 30% identity at the amino acid and act redundantly in *Arabidopsis* to control floral induction (Smith *et al*., 2004). Given the similarity between BLH3 and BLH10, it is possible that these proteins share overlapping roles in *Arabidopsis*. These results provide justification for further work with *BLH10*.

3.5.1.1 THE EXPANSION OF THE BELL FAMILY

The expansion of the BELL family in plants was likely to have occurred within the lineage of Viridiplantae, as no BELL-like proteins have been identified in species outside of the plant kingdom. It is interesting to note that the comparison of the BLH3 sequence to ESTs and genomic sequence from the green algae *Chlamydomonas* failed to identify any BELL-like proteins, yet did detect a closely related KNOX-like transcription factor. The BELL gene family is present in the genomes of a diverse range of seed plants, including the divisions Magnoliophyta (flowering plants), Pinophyta (conifers) and Gnetophyta. Comparison of the BLH3 sequence to the annotated *Physcomitrella patens* genome failed to reveal any BELL-like proteins in this moss species. Only the complete genomic sequences of other embryophytes, including the bryophytes (non-vascular plants) and tracheophytes (vascular plants), will ultimately determine how early within the Viridiplantae the expansion of the BELL family occurred.

Recent studies suggest that three large-scale duplications or polyploidisation events have a major role in the evolution of the *Arabidopsis* genome since the split of this lineage from *Oryza* (Simillion *et al*., 2002; Blanc *et al*., 2003; Bowers *et al*., 2003). The genomic
regions containing BLH3 and BLH10 correspond approximately to an interchromosomal duplication of a segment of chromosome I (AGI, 2000). The approximate age of this duplication block is 75 mya (Simillion et al., 2002), therefore this is deemed a relatively recent duplication event. This is compatible with the high degree of similarity seen not only between BLH3 and BLH10 but also the surrounding genes (Figure 3.11). These results imply that BLH3 and BLH10 result from a recent chromosomal duplication event in Arabidopsis.

By comparing all the BELL loci to the duplication events identified by Simillion et al. (2002) and Blanc et al. (2003), the only other recent duplication event within the BELL family is between BLH2 and BLH4. More ancient duplications events are associated with the PNY, PNF, BELL1 and ATH1 loci, hence the synteny between these loci has been lost over time. In accord with this theory, the genes flanking ATH1, PNY and PNF share little to no homology to each other or with genes flanking other BELL genes (Figure 3.10). It has been suggested that smaller scale duplications and/or gene loss events may have occurred after segmental duplication creating hidden duplication blocks (AGI 2000; Simillion et al., 2002). This could account for the fact that some loci share only a single gene pair, for example the RING finger proteins flanking BELL1, BLH1 and BLH4 (Figure 3.10). Together these results suggest that more than one duplication event has given rise to the expansion of the Arabidopsis BELL genes. Although the results imply an ancient origin for the BELL gene family, the presence of well conserved motifs within the proteins signifies that some common function within this family has been maintained.

3.5.1.2 Function of the BELL Proteins

What is the role of the BELL proteins and why are there multiple members of this family? Do these genes have a common function or have they evolved to perform distinct tasks? The majority of BELL proteins, with the exception of PNF and BLH11, interact with other BELL proteins in yeast 2-hybrid (Hackbusch et al., 2005). Interestingly, these interactions occur between proteins in both closely and more distantly related clades. The facility for interaction with KNOX proteins is also well conserved throughout the BELL family. Proteins known to be involved in BELL::KNOX and BELL::BELL protein interactions are represented in all seven clades, which includes twelve of the thirteen Arabidopsis proteins. Importantly, this verifies that within the BELL family the ability to mediate protein interactions is well conserved. The function of BELL proteins in protein interactions is addressed in more detail in Chapter 4.
Excluding protein interactions, less is known about the function of the BELL proteins in plants and if these functions are conserved within clades or groups. The accumulation of higher levels of mRNA in flowers (ATH1, PNY) (Quaedvlieg et al. 1995; Byrne et al. 2003) and ovules (BEL1, MDH1) (Reiser et al. 1995; Dong et al. 2000) is evident for genes in four separate clades. The gene products of ATH1, BLH3, PNY and PNF (clades II, VI, VII) can act in the induction and development of flowering (Smith et al., 2004; Cole et al., 2006; Kanar et al., 2006). Dwarf plants and reduction in leaf size is a common phenotype of plants mis-expressing PNY, JUBEL2, MDH1, and StBEL5 (Dong et al. 2000; Muller et al. 2001; Byrne et al. 2003; Chen et al. 2003; Bhatt et al. 2004). These genes are represented in four of the seven clades. It is evident that the BELL proteins have a complex role in plant patterning and reproductive development.

The BLH11 protein is unique among the BELL proteins analysed as it contains a 7 aa insertion within the homeodomain, between the conserved TALE (PYP) residues (Appendix 1.3). It is extremely likely that this insertion affects the DNA binding function of the homeodomain in BLH11. Not only are the PYP residues absolutely conserved within the entire TALE superclass of transcription factors, previous work with the KNOX protein OSH1 showed that an amino acid substitution in the TALE region is associated with loss of normal function of this protein (Sato et al., 2002). This raises the possibility that BLH11 does not function as a normal transcription factor. BLH11 encodes the conserved BELL and SKY domains that are characteristic of the BELL family and is placed firmly within the BELL phylogenetic tree (clade VI, Chapter 3.3.4). Therefore, BLH11 is classified as the thirteenth family member in this thesis work. As conserved regions outside of the homeodomain are responsible for mediating protein interactions, one may speculate that BLH11 may be involved in protein interactions.

It is also likely that the rice proteins OsBELL10 and OsBELL11 are not functional BELL proteins, as they each encode a protein that lack a complete BELL domain and homeodomain respectively (pileup, Appendix 1.3). Interestingly, cDNA clones corresponding to these genes have been found, indicating that they are expressed in plants. It is tempting to speculate that these proteins could act in a dominant negative manner in plants, adding a further level of complexity to the control and function of BELL proteins.

### 3.5.2 Comparison of the BLH3 and BLH10 cDNAs
Further similarities between the \textit{BLH3} and \textit{BLH10} genes were seen after the examination and comparison of the cDNA sequences. Putative full length cDNA clones of \textit{BLH3} and \textit{BLH10} were isolated and their sequences analysed (Chapter 3.4). Recent examination of over 20 000 \textit{Arabidopsis} cDNA sequences found that the average cDNA length is 1055 bp and the average size of the 5' UTR is 82 bp (Alexandrov \textit{et al}., 2006). Comparison to the \textit{BLH3} and \textit{BLH10} cDNA demonstrate that both overall length of the cDNAs and the 5' UTR sequences are longer than the plant average. The \textit{BLH3} and \textit{BLH10} clones show 61\% identity to each other at the nucleotide level. The sequences of these cDNAs were compared to the genomic sequences to confirm intron/exon borders. This revealed that \textit{BLH3} and \textit{BLH10} have a similar intron/exon structure, with both genes containing an intron upstream of the predicted translation start site and three smaller introns within the coding sequence.

3.5.2.1 LEADER INTRONS AND THE 5' UTR

One interesting feature of the \textit{BLH3} and \textit{BLH10} cDNAs is that they contain leader introns within the 5' UTR. The 274 bp intron in the \textit{BLH10} 5' UTR and the 452 bp intron in the \textit{BLH3} 5' UTR are larger than average and contain numerous stop codons in all three reading frames. Introns within the 5' UTR and genes are important in the control of transcription and translation of genes in \textit{Arabidopsis} (Rose, 2004; Curi \textit{et al}., 2005). The comparison of cDNA sequences and gene expression in \textit{Arabidopsis} indicates that five or more exons (four introns) are necessary for maximum average expression, which reflects the role of introns and splicing in the control of gene expression (Alexandrov \textit{et al}., 2006). In rice, the Os\textit{TUB16} gene contains a long intron of 863 bp within the 5' UTR that is absolutely required to maintain high levels of expression (Morello \textit{et al}., 2002).

The \textit{BLH7} gene, closely related to \textit{BLH3} and \textit{BLH10}, also contains a leader intron. Furthermore, the presence of a leader intron is spread through the BELL family, with the more distantly related genes \textit{BLH1} and \textit{ATH1} both containing two introns each within the 5' UTR regions. The \textit{GI} gene also contains a leader intron and the light regulated \textit{PHYA} contains a long leader intron of 923 bp within the 5' UTR. Experimental evidence indicates that despite conservation of intron position throughout the profilin gene family in \textit{Arabidopsis}, each intron has a different function in controlling gene expression (Jeong \textit{et al}., 2006). This implies that differing intron sequences within a conserved family can provide a further level of control of gene expression. Taking these results into account, it is
feasible that the leader introns located in the 5' UTRs of the BLH3 and BLH10 genes have a role in the regulation of these genes. The fusion of BLH3 and BLH10 promoter deletions to a reporter gene such as GUS and subsequent comparison of the expression pattern generated to that of intact promoter sequences could determine the role of the leader introns in the regulation of BLH3 and BLH10.

### 3.5.2.2 ALTERNATIVE SPLICING AND DIFFERENT cDNA FORMS

Functionally different proteins can be produced from a single primary transcript that is alternatively spliced and processed into different forms. The detection of alternative splicing in plants is increasing, although in many cases the biological significance is unknown (reviewed in Lorković et al., 2000). A study of over 15000 Arabidopsis transcriptional units found that 11.6% had splicing variants (Iida et al., 2004). Two distinct BLH10 cDNAs have been isolated; one an alternative acceptor site splice variant encoding a truncated protein, the second encoding a full length putative BELL protein. Although the splice-variant transcript was not detected in wild type plants grown in standard conditions (Chapter 3.4), it is possible that it is expressed in plants when conditions are altered. Some alternative splicing events are under developmental or tissue-specific control and different splicing profiles have been detected in plants subject to environmental stress conditions (reviewed in Lorković et al., 2000; Iida et al., 2004). For example, the transcripts of three rice KNOX transcription factors HOS58, HOS59 and HOS66 are alternatively spliced in an organ specific manner and this affects the ability of their products to control the transcription of target genes (Ito et al., 2002). The processing of the flowering time gene FCA gene is well studied in Arabidopsis. Alternative splicing and polyadenylation produces four mRNA species, only one of which is functional and promotes early flowering. Furthermore, the presence of introns within the FCA transgene affects not only expression of the transcript, but also the levels of the FCA protein (Macknight et al., 2002; Quesada et al., 2003).

The alternative splicing of transcripts was also found in nine of the fifteen rice BELL genes identified (Chapter 3.3). The alternatively spliced rice cDNA clones encoded truncated BELL proteins, up to 394aa smaller than those predicted by the genomic sequence. In the cases of OsBELL1, OsBELL3, OsBELL4 and OsBELL8, the genomic annotations were supported by other full length cDNA clones, suggesting that these genes do encode functional proteins in plants. This implies that there may be an additional level
of post transcriptional control in BELL proteins in rice. However, the role of the alternative splicing of BELL genes, particularly BLH10, has yet to be elucidated.

3.5.3 SUMMARY

The BLH3 protein was confirmed as a putative GI protein interactor using yeast 2-hybrid assays and in vitro protein assays. The phylogenetic and genomic structure analyses carried out in this chapter have contributed to the understanding of the evolutionary relationships within the BELL family. It was found that BLH3 was closely related to another Arabidopsis BELL gene, named BLH10. This has provided useful information for the design of further experiments, as it is possible that BLH3 and BLH10 have similar and/or conserved functions. To determine if the ability to interact with GI is conserved in the BELL family, particularly by BLH10, further yeast 2-hybrid assays were performed (Chapter 4).

The identification of three conserved motifs in BLH3 and BLH10 provides the foundation for deletion experiments. Deletions of BLH3 and BLH10 were generated and a reverse yeast 2-hybrid system was employed to determine the protein domains important for mediating interaction with GI. These experiments are also described in Chapter 4.
CHAPTER FOUR: **ANALYSIS OF THE GI AND BLH3 PROTEINS USING A REVERSE 2-HYBRID SYSTEM**

4.1 **INTRODUCTION**

In the previous chapter the interaction between the BLH3 and GI proteins in a yeast 2-hybrid system was described. The reverse 2-hybrid assay is a powerful tool for analysing proteins known to interact and for identifying regions or domains within the proteins that mediate interactions. This chapter illustrates the use of deletion clones of BLH3 and GI to determine the regions within these proteins that are involved in the GI:BLH3 interactions. The first section briefly outlines the generation of GI and BLH3 deletions constructs and the second section describes the analysis of interaction between the various deletions. Lastly, the question of whether GI interacts with other BELL proteins is addressed.

4.1.1 **BELL PROTEIN INTERACTION DOMAINS**

Analyses of the putative BELL protein sequences illustrate the presence of three distinct domains: the SKY domain, BELL domain and homeodomain regions (Chapter 3.3.3). Both the SKY and BELL domains contain putative amphipathic α-helices, which are implicated in protein recognition and interaction. In recent years, it has been established that the SKY and BELL domains are involved in protein-protein interactions, as deletion of part of one or both of these domains can abolish protein interactions (Muller *et al.*, 2001; Smith *et al.*, 2002; Chen *et al.*, 2003). The region spanning the BLH3 BELL domain is sufficient for interaction with the KNOX protein STM (Cole *et al.*, 2006). Conversely, experimental work by Bellaoui *et al.* (2001) indicates that at least two additional regions N-terminal and C-terminal of the BELL domain within BELL1 can interact with KNOX proteins. The examination of nested deletions found that the homeodomain region of BELL proteins mediated interactions with AtOFP proteins (Hackbusch *et al.*, 2005).
The initial yeast 2-hybrid screen identified two partial BLH3 clones that encoded proteins which interacted with GI (Chapter 3.1.1). The longer clone of the two, which contained an intact SKY domain, interacted non-specifically with the empty pAS vector. In contrast, the shorter clone which lacked part of the SKY domain interacted specifically with GI. This result revealed that sequences N-terminal to the BELL domain are not required for the interaction of the BLH3 and GI proteins. For this reason, the SKY domain and N-terminal sequences were not included in the deletion assays. The investigation of interactions of BLH3 deletions with GI are described in Chapter 4.4.3.

4.1.2 Which region of GI is important for mediating protein interactions?

The logical follow up experiments to the yeast 2-hybrid work described in Chapter 3 would be to identify regions or domains within the GI protein that are involved in mediating the protein interaction with BLH3. The predicted GI protein holds no clues, as it contains no recognisable domains of known function. Previous computational analyses suggested that GI contains at least five transmembrane domains (Fowler, 2000). In contrast, experimental work shows that GFP-tagged GI protein is located in the nucleus (Huq et al., 2000; Mizoguchi et al., 2005). Despite the enormous amount of sequence data now available in the Genbank and TIGR databases, homology searches carried out using GI still do not reveal any domains of homology with known proteins.

Initially, a series of deletions of the C-terminus of GI were created to pinpoint any specific regions necessary for interaction with BLH3. Deletions encompassing the N-terminus of GI were also constructed. These results of these assays are reported in Chapter 4.4.3.

4.1.3 Does GI interact with other BELL proteins?

It has been well documented that the KNOX and BELL proteins interact, not only in *Arabidopsis*, but in other plant species such as potato, rice and barley. A common feature of these interactions is that the KNOX protein partner often interacts with a subset of BELL proteins (Muller et al., 2001; Chen et al., 2003; Hackbusch et al., 2005; Cole et al., 2006). Potentially, interaction with a protein such as GI may be conserved between members of the BELL family. The BLH10 protein was identified and is highly similar to
BLH3 over the three conserved domains (Chapter 3.3.3). As BLH10 interacts with a subset of proteins which interact with BLH3 (Hackbusch et al., 2005), the possibility exists that this protein may also interact with GI. Interactions with ATH1 and BELL1 were also tested to establish if BELL proteins less closely related to BLH3 interact with GI. These experiments are described in the penultimate section of this chapter (Chapter 4.5).

### 4.2 Creation of the BLH3 Deletion Constructs

Four BLH3 deletion constructs were used in the deletion assays (pACT:Dom1, pACT:Dom2, pACT:Dom3 and pACT:Dom4) (Figure 4.1). These clones encompassed the BELL, homeodomain and C-terminal sequences. The plasmids pACT:Dom2 (aa 298-429) and pACT:Dom3 (aa 298-524) were constructed previously (C. Pullen, this laboratory). The plasmid maps for these two vectors are provided in Appendix 2.2.

The plasmids pACT:Dom1 (aa 180-343) and pACT:Dom4 (aa 180-429) were constructed as part of this thesis work. The clone encoding the partial BLH3 protein, isolated from the original 2-hybrid screen, was inserted into the bait vector (pAS2-1) (Figure 4.1) in order to test for dimerisation with the BLH3 clones in pACT. The construction of the three vectors (pACT:Dom1, pACT:Dom4, pAS:BLH3) is described in the following sections.
4.2.1 The Construction of pACT:Dom1 and pACT:Dom4

The proteins encoded by the Dom1 and Dom4 deletions encompass the BELL domain and BELL and homeodomain regions of BLH3 respectively (Figure 4.1). These deletions were amplified by PCR from the BLH3 cDNA with primers designed to introduce BamHI and XhoI restriction sites to the 5′ and 3′ ends of the sequence. The 5′ BamHI site was also required to create a fusion in frame with the GAL4 activation domain. The PCR products were ligated into pGEMT and subject to sequence analysis to ensure no PCR induced errors were present (data not shown). The deletion clones were released by digestion with BamHI/XhoI and inserted into BamHI/XhoI digested pACT to create pACT:Dom1 and pACT:Dom4 (Figure 4.2). The identity of the clones was confirmed by restriction enzyme digestion and sequence analysis (data not shown).

4.2.2 The Construction of PAS:BLH3

At the time of this work, it was not known if the BLH3 protein was able to homodimerise. The partial BLH3 clone (aa 180-524) that was isolated in the original yeast 2-hybrid screen was used to investigate this and was cloned into the opposite yeast 2-hybrid vector, pAS2-1. A BLH3 restriction fragment was released from pMAL:BLH3 (Chapter 3.2.3.1) by digestion with restriction enzymes EcoRI and BamHI. The 1032 bp fragment was inserted into EcoRI/BamHI digested pAS2-1 to generate pAS:BLH3 (Figure 4.3a). Two positive
clones were selected and the identity of the insert was confirmed by restriction enzyme digestion (Figure 4.3b).

![Restriction map of pAS:BLH3](image)

**Figure 4.3 Generation of the deletion construct pAS:BLH3**

* Figure 4.3 Generation of the deletion construct pAS:BLH3. Each digest was carried out on two independent clones. The approximate sizes of the fragments produced by digestion and visible on the gel are: *BamHI/EcoRI* (B/E); 8.4, 1.4 kb, *HindIII* (H); 4.6, 2.2, 0.9, 0.7, 0.7, 0.45 kb. **1** kb; 1 kb+ DNA size marker, only fragments greater than 300 bp are visible on this gel.

### 4.2.3 Summary

Three *BLH3* clones encoding truncated proteins were generated as part of this thesis work for use in deletion assays. The *pACT:Dom1* and *pACT:Dom4* constructs represented deletions of *BLH3* encompassing the BELL domain to test if this region was important for interaction with GI. The third construct was a longer *BLH3* clone inserted into the opposite mating vector *pAS2-1*, which was used to determine if BLH3 homodimerised in yeast 2-hybrid.

These three constructs (*pACT:Dom1*, *pACT:Dom4* and *pAS:BLH3*), as well as two constructs generated previously (*pACT:Dom2* and *pACT:Dom3*) were transformed into the yeast strain PJ69. The *pACT* based plasmids were transformed into the mating strain PJ69 4α and the *pAS2-1* based plasmid was transformed into the opposite mating strain PJ69 4a (Chapter 2.3.3). Putative transformants were screened by PCR using insert specific primers and two individual transformants representing each deletion clone were selected. Plasmid DNA was extracted from yeast colonies (Chapter 2.5.1.4) and retransformed into *E. coli* (Chapter 2.2.3.1) to confirm the identity of the plasmids by restriction enzyme digestion (data not shown).
4.3 THE CREATION OF GI PROTEIN DELETIONS

4.3.1 ANALYSIS OF THE STRUCTURE OF GI PROTEINS

Comparison of the sequence of GI to recognised proteins has yielded no clues to the role of GI; the GI protein has no domains of known function. In an attempt to identify any particularly well conserved regions or domains within GI, a search for related GI sequences was performed. Full length GI proteins were identified in four species, though a pileup of these sequences failed to locate specific regions or domains likely to mediate protein interactions (Chapter 1.5.5; Figure 1.8).

Fortunately, GI-like proteins have now been identified in a wide range of plant species, from angiosperms such as tomato and wheat, to gymnosperms such as pine. To test if the addition of more GI sequences may help resolve the protein pileup and pinpoint conserved domains, five additional partial GI-like sequences from other plant species were selected from the TIGR Gene Indices (sequences provided in Appendix 3). The EST clones encoded partial protein sequences, so the C-terminal ends of the five GI-like proteins were aligned against AtGI, OsGI, TaGI and HvGI using the GAP program (GCG) (Figure 4.4). Again, these proteins show a high level of conservation over the majority of the protein sequence, with much of the conservation at the level of identity. A small region spanning thirty four amino acids adjacent to the C-terminal end of GI was the only significant region of low sequence conservation.

An important addition to this pileup is a partial GI sequence from Pinus taeda (PtGI, loblolly pine). This sequence represents the only GI protein in this analysis from the gymnosperms and thus has a higher degree of divergence from GI proteins in the other plant species. There remain 66 residues conserved in the C-terminal region of PtGI compared to the GI proteins from the angiosperm species (Figure 4.4). This level of identity suggests a constraint due to function on this region of the GI gene. This is supported by a previous proposal that the C-terminal end of the GI protein is important in its role in flowering, as a number of late flowering plants have been identified that encode a truncated GI protein that lack the C-terminal (Fowler et al., 1999).
Figure 4.4 Partial Sequence alignment of GI-like proteins

Multiple amino acid alignment of the C-terminal end of the GI proteins. Features of the sequence are as indicated: residues identical in at least 5 sequences are shaded dark grey and residues conserved in at least five sequences are light grey. A small PtGI EST is at the bottom of the alignment, residues absolutely conserved in all sequences including PtGI are highlighted black. A region of lower conservation underlined with striped line. At, Arabidopsis thaliana; Hv, Hordeum vulgare; Le, Lycopersicon esculentum; Os, Oryza sativa; Pt, Pinus taeda; Sb, Sorghum bicolor; St, Solanum tuberosum; Ta, Triticum aestivum; Vv, Vitis vinifera.
4.3.2 THE GI DELETION CONSTRUCTS

The first step in analysing interactions between GI and BLH3 deletions was to generate deletion constructs in yeast plasmids. Previously, a series of increasingly smaller deletions to the C-terminal of the bait construct were made to pinpoint the region of the protein that is important for interaction with BLH3. The 3' end of the GI construct pAS:GI was progressively deleted by DNase digestion, resulting in the removal of 63 aa, 149 aa and 363 aa to generate deletions pAS:Del1 (aa 740-1110), pAS:Del2 (aa 740-1024) and pAS:Del4 (aa 740-810) respectively (C. Pullen, this laboratory) (Figure 4.5). The restriction enzyme maps used to verify the identity of these plasmids is provided in Appendix 2.3.

![Deletion constructs used to identify GI interacting domains](image)

**Figure 4.5** Deletion constructs used to identify GI interacting domains

Truncated GI (pAS:GI, pAS:Del1:7) sequences used in yeast 2-hybrid interactions. The amino acids encoded by each clone are indicated in superscript. NLR, putative GI nuclear localisation region, aa 543-783.

As part of this thesis work, three further GI deletion constructs (pAS:Del5, pAS:Del6 and pAS:Del7) were generated. Fusions of GI deletion clones to the C-terminus of the GAL4 binding domain were made using the pAS2-1 plasmid. The pAS:Del5 deletion (aa 811-1173) encompasses all sequences within pAS:GI that are exclusive of those in the smallest deletion pAS:Del4 (Figure 4.5). In contrast, pAS:Del6 contains N-terminal sequences exclusive of the pAS:GI bait clone. Finally, the pAS:Del7 clone is a small N-terminal deletion located between predicted transmembrane domains. The construction of these deletion vectors is described in the next section.
4.3.2.1 CONSTRUCTION OF THE pAS:DEL5 DELETION

The Del5 deletion (aa 811-1173) includes sequences within pAS:GI that are exclusive of those in the smallest deletion Del4 (Figure 4.4). Del5 was amplified by PCR from the GI cDNA with primers designed to introduce the required restriction sites EcoRI and BamHI. Initially, amplification of these deletions was carried out using pAS:GI as a template. However, sequence analysis of the PCR products revealed an identical error appearing in all products (Chapter 3.2). The pAS:GI vector was reconstructed from the GI cDNA and subsequent yeast 2-hybrid assays and amplification reactions were carried out using the reconstructed pAS:GI vector (Chapter 3.2). Del5 was re-amplified; the 1086 bp PCR product was inserted into the shuttle vector pGEMT and subject to sequence analysis to ensure no PCR induced errors were present. The Del5 fragment was released from pGEMT by digestion with the restriction enzymes EcoRI and BamHI and subcloned into EcoRI/BamHI digested pAS2-1 to create pAS:Del5 (Figure 4.6a). The identity of the resulting plasmid was established by restriction enzyme digestion and the junction with pAS2-1 was confirmed by sequencing (Figure 4.6b).

Figure 4.6 Analysis of the pAS:Del5 deletion construct

a  Restriction map of pAS:Del5, drawn approximately to scale. Key restriction enzyme sites are shown. The direction of the open reading frames are indicated by the arrows. Pr, ADH1 promoter; BD, GAL4 binding domain; 3’, ADH1 3’ terminator.

b  Sequence analysis of pAS:Del5 to check the junction of the Gal4 Binding Domain and GI to ensure that the GI gene was in frame. The EcoRI restriction site used to insert GI downstream of the Gal4 Binding Domain is highlighted by the black box. The translation of the GI sequence is shown above the nucleotide sequence.

4.3.2.2 CONSTRUCTION OF THE pAS:DEL6 DELETION
The second GI deletion, Del6 (aa 1-740) contains sequence that is exclusive of the bait vector pAS:GI and is comprised of the remaining 2540 bp of GI coding sequence upstream of the GI sequence cloned into pAS:GI. Firstly, the 3' end of the GI cDNA was removed by digestion of pBS:GI with XcmI and BamHI. The resulting plasmid contained 2352 bp of GI 5' sequence, including the ATG (Figure 4.7a).

Figure 4.7 Generation of the deletion construct pAS:Del6

a Diagram outlining the strategy used to construct the pAS:GI vector. In summary, a 258 bp PCR product was amplified from the GI cDNA and inserted into pGEMT. The pGEMT:258 and pBS:GI plasmids were digested with XcmI and BamHI and ligated, generating pBS:Del6. Del6 was released from pBS:Del6 and ligated into pAS2-1. The direction of open reading frames are indicated by the arrows. Pr, ADH1 promoter; BD, GAL4 binding domain; 3', ADH1 3' terminator.

b Restriction map of pAS:Del6, approximately to scale. The restriction enzyme sites for HindIII (▼) and BglII/PstI (▲) are indicated.

c Restriction enzyme digests used to verify the structure of pAS:Del6. Each digest was carried out on two independent clones. The approximate sizes of the fragments produced by digestion are: HindIII; 0.9, 2.2, 2.9, 4.6 kb, BglII/PstI; 1.0, 3.2, 6.4 kb.

The remaining 258 bp of GI required was amplified by PCR from the cDNA sequence and inserted into pGEMT. Following sequencing analysis, the 258 bp product was subcloned.
into \textit{XcmI/BamHI} digested pBS:GI to create pBS:Del6. The 2.5 kb Del6 restriction fragment was released by restriction enzyme digestion with \textit{VspI} and \textit{BamHI} and subcloned into \textit{NdeI/BamHI} digested pAS2-1 (note that \textit{VspI} and \textit{NdeI} digestion create compatible overhanging ends) to produce pAS:Del6 (Figure 4.7a). The recombinant plasmid was used to transform \textit{E. coli} and two single transformed colonies were selected. Plasmid DNA was extracted and analysed by restriction enzyme digestion (Figure 4.7b and c).

### 4.3.2.3 Construction of the pAS:Del7 GI Deletion

The third \textit{GI} deletion, \textit{Del7} (aa 109-215) encodes a small N-terminal portion of the \textit{GI} protein. The 318 bp \textit{Del7} deletion was amplified by PCR from the \textit{GI} cDNA and inserted into pGEMT. The \textit{Del7} insert was subject to sequencing analysis to ensure that no PCR induced errors were present. The \textit{Del7} restriction fragment was released by digestion with \textit{EcoRI} and \textit{BamHI} and was subcloned into \textit{EcoRI/BamHI} digested pAS2-1, producing pAS:Del7 (Figure 4.8). The identity of the pAS:Del7 plasmid was confirmed by restriction enzyme digestion (data not shown).

#### Figure 4.8  
Restriction enzyme map of pAS:Del7

\textit{GI} deletion \textit{Del7} was cloned into the yeast 2-hybrid vector pAS2-1. Restriction enzyme map (not to scale) shows the restriction sites used to confirm the identity of the \textit{Del7} insert.

### 4.3.3 Summary

In summary, three \textit{GI} clones encoding protein deletions were inserted into the pAS2-1 vector, these were pAS:Del5, pAS:Del6 and pAS:Del7. These three vectors and four constructed earlier, (pAS:GI, pAS:Del1, pAS:Del2 and pAS:Del4) were transformed into the yeast strain PJ69 4a. Putative transformants were screened by PCR using insert specific primers and two individual transformants representing each deletion clone were selected. Plasmid DNA was extracted from yeast colonies (Chapter 2.5.1.4) and retransformed into
E. coli (Chapter 2.2.3.1). Subsequently, positive clones were selected and the identity of the plasmids was confirmed by restriction enzyme digestion (data not shown).

4.4 Analysis of the BLH3 and GI deletions

The BLH3 and GI vectors encoding protein deletions generated in the previous sections (Chapter 4.2 and Chapter 4.3) were maintained in yeast strain PJ69 and used in reverse yeast 2-hybrid assays. Briefly, haploid transformants were maintained on plates lacking tryptophan and leucine. After mating to produce diploid clones containing both bait and prey plasmids, interaction assays were carried out on plates lacking either adenine or histidine. A positive interaction was recognised by clones displaying histidine and/or adenine prototrophy.

4.4.1 The Bell and Homeodomain Regions are Involved in the BLH3::GI Interaction

It had already been established that the SKY domain within BLH3 is not necessary for interaction with GI, as the shorter of the two BLH3 clones originally isolated from the 2-hybrid screen (pACT:BLH3) lacks most of this region (Figure 4.1). Therefore, deletions that contained either the BELL or homeodomain regions were constructed in the prey vector pACT (Chapter 4.2). The BELL domain is the most likely protein interacting domain as it is predicted to contain amphipathic $\alpha$-helical regions.

The two clones containing the BELL domain (Dom1 and Dom4) were the only clones that interacted with pAS:GI (Figure 4.9). In fact, the BELL domain (Dom1, aa180-343) was the smallest independent region of BLH3 that mediated interaction with pAS:GI (Table 4.1). Conversely, the BELL domain alone was not sufficient to interact with the deletions of GI (Del4, Del5) that were able to interact with pACT:BLH3. Only Dom4, which spanned the BELL and homeodomain regions, interacted with the identical subset of GI clones as pACT:BLH3 (Table 4.1). The homeodomain region alone (Dom2, aa298-429) does not interact with any GI clone tested. The entire BLH3 protein C-terminal to the BELL domain, including the homeodomain (Dom3, aa298-524) did interact with one GI deletion (Del5) (Table 4.1). Yet, this clone did not interact with pAS:GI.
Figure 4.9  Yeast 2-hybrid interactions between GI and BLH3 deletions

The bait vectors (pAS:GI, pAS:Del4, pAS:Del6) were transformed into the yeast strain PJ69 4A. The prey vectors (pACT:BLH3, pACT:Dom1, pACT:Dom4) were transformed into the opposite mating strain PJ69 4α. Following mating, diploid yeast containing both bait and prey plasmids were selected on SC media lacking tryptophan and leucine (SC-trp-leu). Clones exhibiting adenine prototrophy were identified by replica plating onto SC media lacking adenine (SC-ade). Empty bait and prey vectors pAS and pACT were included as negative controls.

Table 4.1  Summary of interactions between GI and BLH3 deletions

The results of the reverse 2-hybrid assays using the GI and BLH3 deletions. Amino acids encoded by deletion clones are in brackets. Transformants were scored for growth on SC-ade, except where * indicates selection on SC-his/2 mM 3-AT. The + indicates interaction and - indicates no interaction with the corresponding protein fusion; +/- indicates a weak/suspect interaction; nt, not tested; Each interaction was tested using two independent clones and repeated at least 3 times.

<table>
<thead>
<tr>
<th></th>
<th>pACT</th>
<th>BLH3</th>
<th>Dom1</th>
<th>Dom2</th>
<th>Dom3</th>
<th>Dom4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(180-524)</td>
<td>(180-343)</td>
<td>(298-429)</td>
<td>(298-524)</td>
<td>(180-429)</td>
<td></td>
</tr>
<tr>
<td>pAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GI</td>
<td>(740-1173)</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Del1</td>
<td>(740-1110)</td>
<td></td>
<td>+</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>Del2</td>
<td>(740-1024)</td>
<td></td>
<td>+</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>Del4</td>
<td>(740-810)</td>
<td></td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Del5</td>
<td>(811-1173)</td>
<td></td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Del6</td>
<td>(1-740)</td>
<td></td>
<td>+/-</td>
<td>+/-</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>Del7</td>
<td>(109-215)</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BLH3</td>
<td>(180-524)</td>
<td></td>
<td>*</td>
<td>nt</td>
<td>nt</td>
<td>*</td>
</tr>
</tbody>
</table>

4.4.2  BLH3 HOMODIMERISES IN YEAST 2-HYBRID
The partial $BLH3$ clone isolated from the original yeast 2-hybrid screen was cloned into the bait vector pAS2-1 to test for homodimerisation (Chapter 4.2.2). While screening for interactions, the pAS:BLH3 construct was found to be auto-activating when selected for adenine prototrophy. This implies that the BLH3 protein contains an activation domain that is able to initiate transcription of the adenine reporter gene. This was not unexpected, as it is common for transcription factors to activate reporter genes when used as bait in yeast 2-hybrid. To overcome auto-activation of the adenine reporter gene, interactions with pAS:BLH3 were selected for histidine prototrophy in the presence of 2 mM 3-AT. In these conditions the pAS:BLH3 clone did not self activate the histidine reporter gene and interacted with pACT:BLH3 (see Figure 4.13) (Table 4.1). The pAS:BLH3 bait also interacted with the BLH3 deletions Dom1 and Dom4, which contain the BELL domain.

### 4.4.3 Deletions of GI Interact with BLH3

Six deletion constructs representing regions of the GI proteins were used in this deletion assay. The deletions of C-terminal sequences of GI in the pAS:GI bait were tested first. All four deletions interacted with pACT:BLH3 (Del1-5, Table 4.1). Surprisingly, both Del5 (aa811-1173) and the small Del4 deletion (aa740-810) interacted with pACT:BLH3. These clones are mutually exclusive, demonstrating this region of GI contains at least two independent sequence motifs capable of mediating interaction with BLH3 in a 2-hybrid assay. Both the Del5 and Del4 deletions interacted with the BLH3 clone containing both the BELL and homeodomain regions (Dom4), although neither interact with the BELL domain (Dom1) nor homeodomain (Dom2) alone. Importantly, none of the deletions interacted with the pACT vector alone or the pTD1-1 control, confirming that the interaction of even the smallest GI deletion Del4 remained specific to BLH3 (Figure 4.9). Furthermore, the Del4 deletion did not simply interact with all BLH3 deletions, demonstrating that some specificity for interaction has been retained in the small 70 aa Del4 protein.

The two N-terminal deletions of GI that contain sequences exclusive to those in pAS:GI were also tested for interaction with BLH3. The small Del7 clone did not interact with BLH3 or with any deletion of BLH3 (Table 4.1). The Del6 deletion interacted with pACT:BLH3 and the two deletions Dom1 and Dom4. Yet growth of colonies containing the Del6 deletion on selective media was inconsistent and often these colonies grew poorly
(Figure 4.9). For this reason, the results gained with the Del6 construct may not be reliable (see Chapter 4.4.4).

### 4.4.4 Proteins expressed by the deletion constructs in yeast

Western analyses of the proteins expressed from the BLH3 and GI vectors were carried out by Revel Drummond (this laboratory) using antibodies to the GAL4 activation and binding domains. The results demonstrated that the deletion constructs generally produced the expected proteins in yeast. However, at 70 kDa the pAS:BLH3 protein was larger than the expected size of 50 kDa. Sequencing of the pAS:BLH3 vector indicated that 400 bp of non-\textit{BLH3} sequence was inserted at the 3` end of \textit{BLH3}. Surprisingly, the Del6 protein was not detected. The presence of a 69 bp insertion of unknown origin within the C-terminal region of the coding sequence introduced six nonsense amino acids followed by a premature stop codon (R. Drummond). No Del6 protein was detected; therefore it is likely that the six nonsense amino acids and/or the premature stop codon caused the Del6 protein to become unstable. Consistent with this result, the growth of yeast colonies after interaction with Del6 was unpredictable. In Figure 4.9 Del6 appears to interact strongly with pACT:BLH3 and Dom4, as demonstrated by the substantial growth of colonies. In contrast, the interaction with the Dom1 deletion could be called weak at best due to the very small growth of colonies. It is difficult to explain why this Del6 clone appeared to interact with some clones and not others, however the interactions were often weak. In light of this finding, any results gained using the pAS:Del6 deletion construct were discounted.

### 4.5 Does GI interact with other BELL proteins?

Due to the high degree of similarity between BLH3 and BLH10, particularly over the conserved domains, we wished to determine if the BLH10 protein was also capable of interaction with GI. The BELL domain was shown to be the smallest region of BLH3 able to interact with GI (Chapter 4.4.1). The region encompassing the BELL domain of BLH10 encodes a peptide that shares 83% identity at the amino acid level with the equivalent BLH3 domain (Figure 4.10b). It is possible that BELL domains are inherently ‘sticky’ and residues located outside of the BELL domain may be required to increase the specificity of protein binding. To test the hypothesis that BELL domains are simply regions that are able
to interact with GI without specificity, the BELL domains of two more distantly related BELL proteins, BELL1 and ATH1, were also inserted into pACT (Figure 4.10).

The KNOX homeodomain proteins are closely related to the BELL family and are able to form hetero- and homodimers via sequences N-terminal to the homeodomain (Bellaoui et al. 2001; Muller et al. 2001). These results suggested the possibility that BLH3 and BLH10 may also dimerise. This was tested using the constructs generated in this section.

![Deletion constructs used to identify interacting BELL proteins](image)

**Figure 4.10** Deletion constructs used to identify interacting BELL proteins

a) Truncated BLH10 (pACT:BLH10, pAS:BLH10), ATH1 (pACT:ATH1) and BELL (pACT:BELL1) sequences used in yeast 2-hybrid interactions. The amino acids encoded by each clone are indicated in superscript. SKY; SKY domain, BELL; BELL domain, HD; homeodomain.

b) Pileup of BELL domains used in yeast 2-hybrid assays. Comparison of the BELL domains of BLH3, BLH10, BELL1 and ATH1. The features of the sequence are as indicated; identical residues shaded black and residues conserved in at least three sequences in grey.

### 4.5.1 The Generation of Two BLH10 Constructs

A truncated BLH10 cDNA corresponding to the conserved BELL domain was inserted into pACT. The BLH10 deletion was amplified from cDNA using primers designed to introduce flanking BamHI restriction enzyme sites. Two attempts at inserting the BLH10 PCR product into pGEMT were not successful. The PCR amplification was repeated with a new BLH10 cDNA template and the 410 bp product was digested directly with BamHI.
Diagram outlining the PCR strategy for identifying pACT:BLH10 clones containing BLH10 inserts in the correct orientation with respect to the Gal4 activation domain. DNA was extracted from selected positive clones and used as a template for PCR amplification. The presence of a product of 410 bp with the BLH10.for and pACT.rev primers (R) indicated that clone 2 contained the insert in the correct orientation. The presence of a 410 bp product with the BLH10.for and pACT.for primers (F) indicated that the insert is in the incorrect orientation in clone 1. BamHI restriction sites are indicated by the blue arrow heads.

Restriction enzyme map of pACT:BLH10 (not to scale) showing the restriction enzyme sites used to confirm the identity of the insert.

Restriction enzyme digest of pACT:BLH10. Plasmid DNA was extracted and digested with BamHI, BglII and EcoRI to confirm the structure of the plasmid. Fragments of the following expected sizes were obtained: BamHI, 7.6 kb, 0.4 kb; BglII, 7.6 kb, 0.4 kb; EcoRI, 3.2 kb, 3.2 kb, 1.7 kb. 1kb; 1 kb+ DNA size marker.

Restriction enzyme map of pAS:BLH10 (not to scale) showing the restriction enzyme sites used to confirm the identity of the insert.

Figure 4.11  Generation of pACT:BLH10 and pAS:BLH10

Diagram outlining the PCR strategy for identifying pACT:BLH10 clones containing BLH10 inserts in the correct orientation with respect to the Gal4 activation domain. DNA was extracted from selected positive clones and used as a template for PCR amplification. The presence of a product of 410 bp with the BLH10.for and pACT.rev primers (R) indicated that clone 2 contained the insert in the correct orientation. The presence of a 410 bp product with the BLH10.for and pACT.for primers (F) indicated that the insert is in the incorrect orientation in clone 1. BamHI restriction sites are indicated by the blue arrow heads. Pr, promoter; AD, activation domain; 3’, 3’ terminator sequences.

Restriction enzyme map of pACT:BLH10 (not to scale) showing the restriction enzyme sites used to confirm the identity of the insert.

Restriction enzyme digest of pACT:BLH10. Plasmid DNA was extracted and digested with BamHI, BglII and EcoRI to confirm the structure of the plasmid. Fragments of the following expected sizes were obtained: BamHI, 7.6 kb, 0.4 kb; BglII, 7.6 kb, 0.4 kb; EcoRI, 3.2 kb, 3.2 kb, 1.7 kb. 1kb; 1 kb+ DNA size marker.

Restriction enzyme map of pAS:BLH10 (not to scale) showing the restriction enzyme sites used to confirm the identity of the insert.
The BLH10 PCR product was inserted into BamHI digested pBS and subject to sequencing analysis to ensure that no PCR induced errors were present (data not shown).

The BLH10 restriction fragment was released by digestion with BamHI and inserted into BamHI digested pACT to generate pACT:BLH10. As this cloning step was non-directional, clones containing the BLH10 insertion in the correct orientation were determined by PCR amplification of the BLH10/pACT junction (Figure 4.11a) and restriction enzyme digestion (Figure 4.11b and c).

Previous experiments demonstrated that BLH3 was able to homodimerise via the BELL domain (Chapter 4.34.2). To establish if BLH10 was also able to homodimerise, a BLH10 cDNA clone encoding all three conserved domains was inserted into the opposite plasmid, pAS2-1. First, an oligonucleotide linker containing an EcoRI site flanked by BglII overhanging ends was introduced into the BglII site in the BLH10 coding sequence. Subsequently, a 1353 bp restriction fragment was released by digestion with EcoRI/PstI and inserted into EcoRI/PstI digested pAS2-1 to create pAS:BLH10 (aa 88-466) (Figure 4.11 d).

### 4.5.2 The Generation of Alternative BELL Constructs

To assess if BELL domains from BELL proteins less closely related to BLH3 interacted with GI and/or BLH3, the BELL domain regions of ATH1 and BELL1 were used in yeast 2-hybrid assays. To create pACT:ATH1 (aa 255-355) and pACT:BELL1(aa 242-380), sequence encoding the putative BELL domains of BELL1 and ATH1 were amplified from the cDNA library CD4-15 (Kieber et al., 1993) using gene specific primers to introduce the restriction enzyme sites required for cloning. The PCR products were inserted into pGEMT and subject to sequence analysis to ensure no PCR induced errors were present (data not shown). The ATH1 and BELL1 restriction fragments were released by digestion with BamHI and XhoI and subcloned into BamHI/XhoI digested pACT to create pACT:ATH1 and pACT:BELL1 (Figure 4.12). The identity of the clones was verified by restriction enzyme digestion (data not shown).
4.5.3 Testing BELL Interactions with Yeast 2-Hybrid

The BLH10, ATH1 and BELL1 deletion constructs generated in the previous sections (Chapter 4.5.1 and Chapter 4.5.2) were transformed into yeast strain PJ69 for use in yeast 2-hybrid interaction assays. Haploid transformants were maintained on plates lacking tryptophan and leucine. After mating to produce diploid clones containing both bait and prey plasmids, interaction assays were carried out on plates lacking either adenine or histidine. A positive interaction was established by clones displaying histidine and/or adenine prototrophy.

The initial 2-hybrid interaction assay proved that the BLH10 protein was able to interact with GI. The BELL domain alone was sufficient for interaction with pAS:GI (Figure 4.13b). Unsurprisingly, the BLH10 BELL domain did not interact with the smaller deletions of GI (Table 4.2). This result is reminiscent of the earlier 2-hybrid assays where the BELL domain of BLH3 was not sufficient for interaction with GI deletions pAS:Del4 and pAS:Del5 (Table 4.1). Notably, pACT:ATH1 and pACT:BELL1 did not interact with pAS:GI (Figure 4.13a and b). These results demonstrate that GI does not interact with all BELL proteins in yeast 2-hybrid assays, but specifically with a subset that includes BLH3 and BLH10.
Figure 4.13  Yeast 2-hybrid interactions between GI and BELL deletions

The bait vectors (pAS:GI, pAS:BLH3, pAS:BLH10), encoding the GI, BLH3 and BLH10 peptides respectively fused to the GAL4 BD, were transformed into the yeast strain PJ69 4A. The prey plasmids (pACT:BLH3, pACT:BLH10, pACT:ATH1, pACT:BELL1) were transformed into the opposite mating strain PJ69 4α. Following mating (Chapter 2.3.4), diploid yeast containing both bait and prey plasmids were selected on SC media lacking tryptophan and leucine (SC-trp-leu). Clones exhibiting adenine or histidine prototrophy were identified by replica plating onto SC media lacking adenine (SC-ade) or histidine (SC-his) respectively. Empty bait and prey vectors pAS and pACT were included as negative controls.

Table 4.2 Summary of GI and BELL yeast 2-hybrid interactions

<table>
<thead>
<tr>
<th></th>
<th>pACT</th>
<th>BLH10</th>
<th>BLH3</th>
<th>ATH1</th>
<th>BELL1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(171-322)</td>
<td>(180-524)</td>
<td>(255-355)</td>
<td>(242-380)</td>
</tr>
<tr>
<td>pAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GI (740-1173)</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Del4 (740-810)</td>
<td></td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Del5 (811-1173)</td>
<td></td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>Del7 (109-215)</td>
<td></td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>BLH3 (180-534)</td>
<td></td>
<td>-</td>
<td>*</td>
<td>+</td>
<td>*</td>
</tr>
<tr>
<td>BLH10 (88-466)</td>
<td></td>
<td>-</td>
<td>*</td>
<td>+</td>
<td>*</td>
</tr>
</tbody>
</table>

Table 4.2 Summary of GI and BELL yeast 2-hybrid interactions

A summary of the results of interactions with GI and BELL proteins. Amino acids encoded by the clones are in brackets. Transformants were scored for growth on SC-ade, except where * indicates that clone exhibits auto-activation when selected on SC-ade, these results represent selection on SC-his + 2 mM 3-AT. The + indicates interaction and - indicates no interaction with corresponding protein fusion; nt, not tested. Each interaction was tested using two independent clones and repeated at least 3 times.
The BELL domain of BLH10 did not interact with the larger BLH10 clone in pACT:BLH10 (Figure 4.13b). This is in contrast to the BLH3 2-hybrid results which demonstrated that BLH3 was able to homodimerise via the BELL domain (Chapter 4.4.2). However, pAS:BLH10 did interact with pACT:BLH3 and pACT:Dom1 (Table 4.2). Therefore, BLH10 is able to heterodimerise with BLH3 in a 2-hybrid assay via the BLH3 BELL domain. The ability of BLH3 and BLH10 to dimerise with two other BELL proteins was also tested by the 2-hybrid assay. Neither pAS:BLH3 nor pAS:BLH10 interacted with the BELL domains of ATH1 and BELL1 cloned into pACT (Table 4.2) (Figure 4.13a).

4.6 DISCUSSION

4.6.1 THE BELL AND HOMEODOMAIN REGIONS ARE IMPORTANT FOR MEDIATING BELL::GI INTERACTIONS

The BELL proteins have three conserved domains (the SKY, BELL and homeodomain) and it has been established that the SKY, BELL and homeodomain regions can mediate protein interactions (Muller et al., 2001, Smith et al., 2002, Chen et al., 2003; Hackbusch et al., 2005). GI interacted with the BLH3 deletions Dom1 and Dom4, both of which contain the BELL domain, demonstrating that BLH3 is able to interact with GI via the BELL domain. The hypothesis that the BLH3 BELL domain is important for protein interactions is reinforced by recent interaction data that show the BELL domain alone is sufficient for interaction with the KNOX protein STM (Cole et al., 2006). Nonetheless, only Dom4 and BLH3 interacted with the identical subset of GI clones, indicating that an important region that regulates interaction with GI is contained within the BELL and homeodomain regions together.

The Dom2 deletion does not interact with any GI clone tested, verifying that the homeodomain alone is not involved in protein interactions between BLH3 and GI. However, the entire BLH3 protein C-terminal to the BELL domain which includes the homeodomain (Dom3) interacts with the Del5 deletion (Table 4.1). In contrast, Dom3 does not interact with pAS:GI, which encompasses all of Del5. Although the C-terminal region of the BLH3 protein is able to mediate protein interactions with GI, this does not occur with the same subset of GI deletions as BELL domain interactions. This additional C-terminal interaction domain in BLH3 may improve specificity or binding to the C-terminal
of GI when present in the full length BLH3 protein. Thus, there are at least three regions within BLH3 that are involved in the interaction with GI.

Experimental work by Bellaoui et al. (2001) confirms that two regions N-terminal and C-terminal of the BELL domain are required for interaction with KNOX proteins to occur. Furthermore, the interaction of all BELL1 deletions with KNOX proteins is consistently weaker than interactions with the full length BELL1 protein, demonstrating that individual domains must work in unison to mediate strong protein interactions (Bellaoui et al., 2001). Together, these results suggest that the three conserved domains within the BELL proteins are not mutually exclusive, but that the functions of these regions overlap and may act collectively to mediate protein interactions.

### 4.6.2 GI INTERACTS WITH BLH10 IN YEAST 2-HYBRID

Further 2-hybrid assays were undertaken to determine if GI interacted with other BELL proteins. Due to the high similarity of BLH3 and BLH10 over the three conserved domains it was not unexpected that the BELL domain region of BLH10 was able to interact with GI. Surprisingly, the interaction patterns of the BELL domains from BLH3 and BLH10 were identical, both interacted with pAS:GI, yet neither interacted with the deletions Del4 and Del5. The BELL domain contains amphipathic α-helices, motifs know to be present in ‘sticky’ proteins such as myosin. It appears that even the small region encompassing the BELL domain of these two proteins has some specificity for interaction with GI despite the presence of these motifs. Even though BLH10 and GI interact in a yeast 2-hybrid assay, it remains that BLH10 was not identified in the original yeast 2-hybrid screen. This could be attributed to a non-saturating screen of the cDNA library or the absence of BLH10 from the library.

The BELL domains of BELL1 and ATH1 share 47% and 29% identity with BLH3 respectively, compared to 83% identity shared by BLH3 and BLH10. As ATH1 and BELL1 did not interact with GI nor BLH3, this supports the suggestion that the BELL domain is not simply a region that is able to interact non-specifically. It would be interesting to determine if other BELL proteins interact with GI, or if this interaction is limited to BLH3 and BLH10. However, GI did not interact with any of the other BELL proteins in the original yeast 2-hybrid screen and it is unlikely that all BELL proteins
(apart from BLH3) were absent from the cDNA library used for the original yeast 2-hybrid screen.

4.6.3 TWO REGIONS OF GI CAN MEDIATE PROTEIN INTERACTION

The GI protein contains no evident domains or regions of homology with proteins of known function. Comparison of GI proteins from multiple species illustrate that these proteins are well conserved and contain a significant level of identity over most of their length (Chapter 1.5 and Chapter 4.3). In a previous yeast 2-hybrid screen using the Arabidopsis SPINDLY as bait, Tseng et al. (2004) isolated a truncated C-terminal GI clone (aa788-1173), which is 48 aa shorter than the pAS:GI clone. This supports the results presented in Chapter 3, which demonstrated that the C-terminal half of the GI protein mediates protein interaction, yet does not help to identify specific domains within GI that are required for interaction. The alignment of GI sequences reveals that the ~70 aa at the C-terminus of these proteins are not highly conserved, with the exception of 10 conserved amino acids (Figure 4.4). The GI deletions Del1 and Del2 lack part and all of this C-terminus end of the protein respectively (Figure 4.5). As these truncated GI proteins interact with BLH3, the C-terminus of GI is not required for BELL protein interaction. The mutually exclusive deletions Del4 and Del5 also encompass the C-terminal half of GI and both interact with BLH3 (Table 4.2). By comparing these results it can be concluded that GI contains at least two regions that can mediate interactions with BLH3; these are located between aa740-810 and aa811-1024.

4.6.4 PROTEIN INTERACTIONS WITHIN THE BELL FAMILY

4.6.4.1 BELL::BELL INTERACTIONS ARE A FEATURE OF THE BELL FAMILY

Recently Nagasaki et al. (2005) demonstrated that the rice BELL protein OsBEL homodimerised, which was in contrast to earlier work with two BELL genes from barley that established that the JUBEL1 and JUBEL2 proteins did not homodimerise in yeast 2-hybrid assays (Muller et al., 2001). At the time of these assays, no data were available on the interaction of Arabidopsis BELL proteins with themselves or with other BELL proteins. It was found that BLH3 homodimerised and that this interaction is mediated by the BELL domain. Unlike BLH3, BLH10 did not homodimerise, which reveals that despite significant similarities, these proteins function differently in protein interaction assays. This work is supported by recent large scale interaction assays performed using full length
BELL proteins (Hackbusch et al., 2005). The results obtained confirm that BLH3 is able to homodimerise, while BLH10 does not.

Further findings of this yeast 2-hybrid work concluded that BLH3 and BLH10 heterodimerise in yeast 2-hybrid and that neither protein interacted with BELL proteins ATH1 and BELL1. Since the time of these assays, results have been published that show that full length BLH3 and BLH10 proteins do not interact with full length ATH1 protein (Hackbusch et al., 2005), which supports the results presented here. However, Hackbusch et al. (2005) have demonstrated that the full length BLH3 and BELL1 proteins interact in a 2-hybrid assay. As these results were gained using full length BELL1, it is probable that sequences outside of the BELL domain are necessary for the interaction with BLH3 and these were not included in the BELL1 deletion used in this work.

Apart from BLH3, the only members of the Arabidopsis BELL family that are able to homodimerise in a yeast 2-hybrid assay are BLH1 and PNY (BLH9) (Hackbusch et al., 2005). The BLH1, BLH3 and PNY (BLH9) proteins are not particularly closely related; these proteins share more similarity to other BELL proteins than to each other and phylogenetic analyses show that they are grouped in distinct clades (Chapter 3.3). All BELL proteins, with the exception of PNF (BLH8) and BLH11, interact with other BELL proteins in yeast 2-hybrid (Hackbusch et al., 2005). This facility for interaction is not confined to groups of similar proteins, but occurs between even the most distantly related BELL proteins. This signifies that dimerisation is an ancient and well conserved function within the BELL gene family.

4.6.4.2 Functions of BELL protein interactions in plants

It is well documented that BELL proteins are involved in protein::protein interactions (Bellaoui et al., 2001; Muller et al., 2001; Smith et al., 2002). The BELL proteins in Arabidopsis interact not only with each other but with members of the KNOX and AtOFP families of proteins. Specifically, the BLH3 and BLH10 proteins were found to interact with the same subset of BELL, KNOX and AtOFP proteins, with BLH3 interacting with an additional seven proteins compared to BLH10. Some redundancy in function between BLH3 and BLH10 may be indicated by the number of shared interactions (Hackbusch et al., 2005).
Recent *in planta* experiments have revealed a number of functions of the BELL::KNOX interaction in plants. It is proposed that the BELL::KNOX interaction is important for regulating the DNA binding of these two transcription factors to DNA. The StBEL5-POTH1 dimer binds directly to a TGAC-motif within the promoter of a gibberellin biosynthetic enzyme where it acts to down-regulate expression (Chen *et al.*, 2004). This cooperative interaction is a characteristic common within the TALE family of homeodomain proteins and is proposed to improve DNA binding ability (Chapter 1.5.6). For example, experimental evidence shows that a BELL::KNOX complex has increased DNA binding affinity compared to either protein alone (Smith *et al.*, 2002; Chen *et al.*, 2004). The BELL and KNOX proteins ATH1 and STM target similar DNA sequences, although with different binding affinity and specificity. The putative DNA-binding amino acids within the ATH1 homeodomain are present in all BELL proteins, suggesting that the recognition of target sequences by the BELL proteins may be conserved in this family (Viola and Gonzalez, 2006).

The physical interaction of BELL and KNOX proteins is implicated in the regulation of the subcellular location of the proteins. It was proposed by Bhatt *et al.* (2003) that nuclear localisation of PNY (VAN) depended on interaction with KNOX proteins STM and BP. Cole *et al.* (2006) have found that GFP tagged STM is not nuclear by default and that nuclear import depends on interaction with BELL proteins BLH3, PNY (BLH9) and ATH1. Taken together, these results imply that interaction between BELL and KNOX proteins are important for the nuclear localisation of protein interaction partners. Members of the plant specific AtOFP family are proposed to regulate the subcellular location of BELL proteins via protein interaction. This is based on experiments where an exclusively nuclear GFP-tagged BLH1 protein was relocated to the cytoplasmic space when co-expressed with the AtOFP1 protein (Hackbusch *et al.*, 2005).

Examination of the interaction between the KNOX protein STM and three BELL protein partners demonstrates that these three different STM::BELL interactions have separate functions in plants (Cole *et al.*, 2006). This reveals that the BELL proteins ATH1, BLH3 and PNY (BLH9) have distinct roles, despite being intricately linked by interaction networks involving themselves and other BELL proteins. In this case, what is the role of BELL::BELL interactions in plants? As BELL::KNOX dimers are required for the efficient binding of DNA, it is possible that the BELL::BELL dimers have a negative role in the regulation of this protein family. BELL::BELL interactions may effectively deplete
BELL proteins, thereby preventing interaction with other target proteins. In conclusion, the
BELL proteins, including BLH3 and BLH10, are likely to function within a complex
network via interactions with subsets of varied protein families in *Arabidopsis*. How this
intricate interaction network is related to the function of BLH3, BLH10 and GI is yet to be
elucidated.

4.6.5 SUMMARY

The results presented in this chapter confirm not only the BLH3 and GI protein::protein
interaction, but also the specific interaction of GI with the closely related BELL protein
BLH10. Additionally, it was shown that the interaction between the GI and BLH3 proteins
is regulated by regions spanning the BELL and homeodomain of BLH3. The BELL
domain of the less closely related BELL1 and ATH1 proteins was not sufficient for
interaction with the GI protein. Therefore GI interacts with a subset of BELL proteins,
which includes BLH3 and BLH10.

Comparison of the protein interaction results presented here to recently published data on
BELL proteins reveals that the BELL proteins function within complex interaction
networks; which signifies that these proteins may have overlapping functions in plants.
The experiments described in the subsequent chapters follow on from the protein
interaction assays and are designed to discover the *in planta* role of BLH3 and BLH10.
Expression analyses are described (Chapter 5) and mutant phenotypes are assessed
(Chapter 6) to determine the regulation and function of the BLH3 and BLH10 genes.
Comparisons and interactions with *GI* could resolve a common or over-lapping function of
these genes in plants.
CHAPTER FIVE: ANALYSIS OF BLH3 AND BLH10 EXPRESSION

5.1 INTRODUCTION

The interaction of BLH3 and BLH10 with GI and investigation of the protein domains involved in these interactions were presented in Chapters 3 and 4. This chapter describes the initial steps toward the characterisation of the BLH3 and BLH10 genes. First, the expression patterns of BLH3 and BLH10 in particular tissues of the plant and throughout development were examined. Flowering in plants is regulated not only by environmental cues, but internal signals which determine when the plant is developmentally competent to flower (Chapter 1.2). Transcription factors have significant roles in diverse areas of plant development, such as light signalling, leaf patterning and flower development (reviewed by Hake et al., 2004; Duek and Fankhauser, 2005; Robles and Pelaz, 2005). Therefore, the presence of the BLH3 and BLH10 transcripts at particular times during plant development was assessed. Fowler (2000) found that the GI transcript was detected in all plant tissues tested, with higher transcript levels detected with increasing age of the plant. To determine where BLH3 and BLH10 expression overlaps with GI, the expression of these genes in different plant tissues was also tested.
The *GI* transcript is circadian regulated with a peak in the evening and a trough at dawn. The cyclical pattern of *GI* expression is slightly altered in LD compared to SD photoperiods (Fowler et al., 1999) (Chapter 1.5.3). Hence, the expression of *BLH3* and *BLH10* at different time points during the day in LD and SD was investigated. The circadian clock is central in the measurement of day length and consequently the regulation of flowering in response to photoperiod (Chapter 1.4.3.5). The expression of *GI* is disrupted in circadian clock mutants such as *lhy* and *cca1*, which is associated with altered flowering time in these mutants. In what is proposed to be a regulatory feedback loop, *gi* mutations also affect the expression of other clock controlled genes (Fowler et al., 1999; Park et al., 1999). Therefore, an investigation of the effect of the *gi*, *cca1* and *lhy* mutations on *BLH3* levels was performed. These experiments are described in Chapter 5.2.4.

Finally, the location of the BLH3 protein at the sub-cellular level was determined and compared to that of GI. Despite computer-based predictions that GI contains multiple transmembrane domains, the GI protein has been shown to be localised to the nucleus in expression assays. Using a 35S:GI:Green Fluorescent Protein (GFP) reporter fusion, GI was localised to the nucleus of onion cells in transient expression assays (Huq et al., 2000). Mizoguchi et al. (2005) demonstrated that a functional GI:GFP fusion was nuclear localised and promoted flowering in transgenic *Arabidopsis*. The first identified BELL protein, BELL1, was shown to be compartmentalised within the nucleus in protoplasts (Reiser et al., 1995). Recent transient assays have revealed that two further BELL proteins, BLH1 and BLH7, are nuclear localised (Hackbusch et al., 2005). As *BLH3* encodes a putative transcription factor, it is probable that BLH3 is also targeted to the nucleus. To establish where BLH3 was localised within the cell, a translational fusion of GFP and BLH3 was used in a transient expression assay. The details of these experiments are illustrated in Chapter 5.3

### 5.2 WHEN AND WHERE ARE BLH3 AND BLH10 EXPRESSED IN PLANTS?

This section describes the analysis of the expression of the *BLH3* and *BLH10* genes. To investigate *BLH3* and *BLH10* transcript levels reverse-transcriptase polymerase chain reaction (RT-PCR) and Northern hybridisation analyses were performed.
5.2.1 TIME-COURSE OF BLH3 AND BLH10 EXPRESSION

GI transcript levels cycle through the day and decrease to undetectable levels during the night. Furthermore, the expression pattern of GI shows subtle differences in LD and SD (Fowler et al., 1999). A time-course was undertaken to resolve the BLH3 and BLH10 expression patterns during a daily cycle. Initial attempts at detecting BLH3 by Northern analysis proved difficult as this gene seems to be expressed at relatively low levels in plants (C. Pullen, this laboratory). For this reason semi-quantitative RT-PCR was employed to assess the presence of BLH3 and BLH10 at different times of the day. BLH3 and BLH10 have a high degree of similarity and are members of a multi-gene family in Arabidopsis (Chapter 3.3). Fortuitously, the identity between BLH3 and BLH10 (and BELL genes in general) is detected primarily at the protein level and is mainly restricted to the region of the proteins spanning the three conserved domains, thus minimising the likelihood of cross-amplification of other BELL genes. Nevertheless, the primers designed to amplify BLH3 and BLH10 were compared to the sequence of BLH10 and BLH3 respectively using the Amplify program (Chapter 2.6.1), which confirmed that each primer pair was specific for the appropriate gene.

Wild type Col plants were grown in LD and SD conditions (Chapter 2.4.1.2) and harvested every 4 h for 24 h. Total RNA was extracted and purified (Chapter 2.5.2.4 and 2.5.2.5) and used as a template in two independent reverse transcription (RT) reactions (Chapter 2.5.4.2). The first strand reaction was diluted ten fold and used as a template for at least two independent PCR amplification reactions (Chapter 2.5.4.1). RNA samples were also subjected to PCR in the absence of an RT step to ensure that the products detected were not the result of amplification from genomic DNA contamination. Trial PCR amplification reactions were performed using BLH3 and BLH10 specific primers, where samples were removed for analysis of products after 15, 20, 25 and 30 cycles. Comparison of the products after gel electrophoresis indicated that amplification of the products had not yet reached saturation at 25 cycles but had at 30 cycles. Therefore PCR amplification following RT was carried out for 25 cycles in all subsequent experiments. PCR amplification reactions with ubiquitin (UBQ) specific primers were carried out as an internal loading control and trial PCR amplification reactions indicated that 15 cycles was sufficient for viewing products before saturation (data not shown). All RT-PCR reactions were repeated at least twice with independent RNA samples and representative experiments are shown.
Initially, an examination of BLH3 and BLH10 expression in plants grown in SD was performed. The results demonstrated that the BLH3 and BLH10 transcripts were present at all time points during the day, although the BLH3 and BLH10 transcripts did vary at different times during the day and night (Figure 5.1). This experiment was repeated using RNA extracted from plants grown in LD conditions. The results demonstrated that the BLH3 and BLH10 transcripts were present at the five time points tested (data not shown). Consequently, it is unlikely that the BLH3 and BLH10 genes are circadian regulated. However, this experiment is not absolutely quantitative and it is not possible to conclude that BLH3 and/or BLH10 transcripts do not cycle at all during a 24 h period.

At 0 h and 12 h after dawn two BLH10 products were detected, 800 bp and 890 bp (Figure 5.1), although the target product of 800 bp was dominant in these PCR reactions. These PCR products were purified (Chapter 2.5.5.3) and directly sequenced (Chapter 2.5.3.8). The 800 bp and 890 bp products contained the expected BLH10 sequence; the difference between the two was the presence of an un-spliced intron (see Chapter 3.5). This was unlikely to be due to DNA contamination, as RNA samples were previously subject to PCR in the absence of the reverse transcription step and no PCR products were obtained. The larger BLH10 transcript was not seen in later RT-PCR assays or Northern analyses probed with the BLH10 cDNA (Figures 5.2 and 5.3).
5.2.2 DEVELOPMENTAL REGULATION OF BLH3 AND BLH10 EXPRESSION

5.2.2.1 BLH3 AND BLH10 ARE DETECTED BY RT-PCR IN ALL PLANT TISSUES TESTED

In order to determine where BLH3 and BLH10 expressed in plants and where the transcripts overlap with GI, the spatial and temporal expression in plants was examined by RT-PCR. Plants were grown in LD conditions in growth cabinets and harvested at ages from 2-leaf to mature plants. The various plant tissues were harvested from mature plants. Total RNA was extracted and used as a template for RT reactions as described previously (Chapter 5.2.1). BLH3 and BLH10 transcripts were detected at all plant ages tested, from 2-leaf seedlings to mature plants (Figure 5.2). Transcripts were also present in all plant tissues selected, including rosette leaves, flowers, roots and stem/cauline leaves. The RT-PCR was repeated three times, with a similar result obtained for each.

![Figure 5.2](image)

**Figure 5.2 Analysis of BLH3 and BLH10 expression in plants by RT-PCR**
Total RNA (4µg) was used as a template for reverse transcription, followed by amplification with gene specific primers (Chapter 2.5.4.1, Chapter 2.5.4.2). Amplification products of BLH3 and BLH10 of the expected size were visualised on an agarose gel. The ubiquitin (UBQ) RT-PCR was used as an internal control. Aerial tissue was harvested from plants of various ages at ~ 0 h after dawn; 2 leaf, 4 leaf, bolted (visible onset of flowering) and mature plants (with developed siliques). Plant material was harvested from four different tissue types; stem and cauline leaves, rosette leaves, flowers, roots and flowers.

5.2.2.2 OPTIMISING THE NORTHERN HYBRIDISATION PROTOCOL TO DETECT BLH3 AND BLH10

Initially, RT-PCR was utilised to examine the expression of BLH3 and BLH10 transcripts, as presumably these genes are not expressed at sufficient levels to be detected consistently by Northern analysis. Adjustment of the Northern hybridisation protocol, particularly by increasing the total RNA and utilising new BLH3 and BLH10 probes with increased sensitivity, improved the detection of the BLH3 and BLH10 transcripts (data not shown).
As the RT-PCR assays of BLH3 and BLH10 expression were not absolutely quantitative, Northern analysis was also employed to establish if the transcript levels of these genes were different in various tissues or at different times of development.

The BLH3 and BLH10 genes share 68% identity at the nucleotide level and belong to a gene family that has thirteen members in Arabidopsis (Chapter 3.3). Therefore, the BLH3 and BLH10 cDNA probes that were used for Northern analyses were tested for cross reactivity. A partial cDNA fragment of BLH10 (1.2 kb) was released from pBS:BLH10 by restriction enzyme digestion with KpnI and PstI and purified. A BLH3 probe of 940 bp, which encompassed the 5' end of the gene, was amplified from the BLH3 cDNA. Three dilutions of control samples of the BLH3 and BLH10 cDNAs (10 ng, 1 ng, 100 pg) were transferred to a nylon membrane and hybridised separately with radio-labelled BLH3 and BLH10 fragments. The BLH10 probe hybridised to BLH3 at high concentrations, but was specific for BLH10 at 100 pg. Similarly, the BLH3 probe cross-hybridised to BLH10 at 10 ng concentration, but was specific for BLH3 at 1 ng concentration (data not shown).

The BLH3 and BLH10 cDNA probes described above were used for Northern hybridisation experiments, as described in Chapter 2.5.6. BLH3 and BLH10 hybridising bands were quantified and background hybridisation levels were subtracted using MacBAS version 2.5 software (Chapter 2.5.6.3). Subsequent hybridisation of a 25s/26s rDNA probe from Asapragus officinalis (pTIP6; King and Davies, 1992) was used to calculate the relative loadings of RNA in each lane. Normalised relative expression levels were expressed as a proportion of the highest value for graphing purposes.

5.2.2.3 Northern hybridisation analysis of BLH3 and BLH10

The expression of BLH3 and BLH10 at particular times of development and in a number of tissues was analysed by Northern hybridisation. Plants were grown in LD conditions in growth cabinets and aerial tissue was harvested at dawn. Plants of three ages were harvested; 2 leaf, 4 leaf and mature plants. As expected from the RT-PCR results, BLH3 was expressed in both immature 2 leaf plants and mature plants, though expression was higher in older plants (Figure 5.3a). In these Northern assays, the BLH10 transcript was detected in 2 leaf, 4 leaf and mature plants. In contrast to BLH3, the BLH10 transcripts were higher in younger plants compared to mature plants (Figure 5.3b).
Plants were grown in LD conditions and tissue was harvested from four tissue types of mature plants; siliques, rosette leaves, stem and cauline leaves and flowers. BLH3 transcripts were detected in all tissues tested, supporting the RT-PCR results. The BLH3 transcripts were highest in rosette leaves compared to flowers, siliques and stem and cauline leaves (Figure 5.3a). The BLH10 transcript was also detected in all tissues tested (Figure 5.3b). Two independent expression assays were performed with similar results and a representative experiment is shown in Figure 5.3.

Figure 5.3 Examination of BLH3 and BLH10 expression in plants by Northern analysis

Total RNA (25 µg) was extracted from tissues and plants of various ages as shown and analysed by Northern hybridisation (Chapter 2.5.6). Plant material was harvested from four different tissue types (flowers, siliques, stem and cauline leaves, rosette leaves) and three ages (2 leaf, 4 leaf and mature plants with developed siliques) at ~ 0 h after dawn. The rRNA probe was used as a loading control.

a Northern hybridisation using BLH3 cDNA and rRNA probes (Chapter 2.5.6.3).

b Northern hybridisation using BLH10 cDNA and rRNA probes (Chapter 2.5.6.3).

c Relative expression of BLH3 in different tissues in Col and gi-2 mutant plants. The results are presented as the proportion of the highest value after normalisation with respect to 25S/26S rRNA levels.

Interestingly, BLH3 expression was lower in rosette leaves from gi-2 mutant plants compared to wild type. In contrast, there was little change in BLH3 expression in flowers.
from gi-2 mutants compared to wild type (Figure 5.3a and c). This experiment was repeated with independent samples and the same result was obtained (data not shown). Like BLH3, BLH10 transcripts were also slightly lower in rosette leaves from the gi-2 mutant compared to wild type and these results were consistent in two independent experiments (Figure 5.3b and data not shown). It is unlikely that the variation in BLH3 and BLH10 expression seen in these assays is due to a difference in the developmental ages of the plants, as rosette leaves were harvested from mature flowering Col and gi-2 plants. These results imply that the levels of the BLH3 and BLH10 transcripts are affected by the loss of GI. The expression of BLH3 and BLH10 in gi mutants was examined further in Chapter 5.2.3.

5.2.3 BLH3 AND BLH10 EXPRESSION DURING THE DAY RE-EXAMINED BY NORTHERN ANALYSIS

As the RT-PCR results obtained previously were not absolutely quantitative (Chapter 5.2.1), the levels of the BLH3 and BLH10 transcripts at different times during the day were assayed by Northern analysis. Col and gi-2 plants were grown in LD conditions until approximately 5-6 leaves had been produced. Tissue was harvested at two time points; at 0 h (dawn) and 8 h after dawn. As indicated by the RT-PCR results, the relative levels of the BLH3 transcript were not significantly different between 0 h and 8 h after dawn in LD (Figure 5.4a).

BLH10 expression was analysed at 0 h and 8 h in LD and comparable levels of the BLH10 transcript were detected at both time points. Furthermore, BLH3 and BLH10 were expressed at similar levels at 0 h and 4 h in SD (Figure 5.4c and data not shown). These results support the outcome of the RT-PCR experiments (Chapter 5.2.2), which indicate that the BLH3 and BLH10 transcripts do not vary significantly during the day.

The results of previous Northern experiments suggested that BLH3 and BLH10 expression may be altered in gi-2 mutants (Chapter 5.2.2.2). Consequently, the expression levels of BLH3 in wild type and gi-2 mutant plants were investigated further. Intriguingly, in this experiment BLH3 levels were lower in gi-2 mutants compared to Col at 0h, yet showed little difference at 8 h in plants grown in LD (Figure 5.4a and b). In a similar experiment, the expression of BLH10 was also reduced in gi-2 mutants at 0 h after dawn in LD (data not shown).
Figure 5.4  Comparison of BLH3 expression in Col and gi-2 plants

a Plants were grown in LD and aerial tissue was harvested from 4-6 leaf plants at 0 h and 8 h after dawn. Total RNA (25 µg) was extracted and expression was analysed by Northern hybridisation using a partial BLH3 cDNA probe (Chapter 2.5.6). The rRNA probe was used as a loading control.

b Relative expression of BLH3 in Col and gi-2 mutant plants at 0 h and 8 h after dawn. The results are presented as the proportion of the highest value after normalisation with respect to 25S/26S rRNA levels.

c Plants were grown in SD and total RNA (25 µg) was extracted from aerial tissue at 0 h and 4 h after dawn. Expression was analysed by Northern hybridisation using partial BLH3 and BLH10 cDNA probes (Chapter 2.5.6). The rRNA probe was used as a loading control.

This experiment was repeated with plants grown in SD, with plants harvested at 0 h and 4 h after dawn. The results obtained were comparable, with BLH3 expression levels in gi-2 plants approximately 60% of BLH3 levels in Col plants at 0 h (Figure 5.4c). The levels of the BLH10 transcript in Col and gi-2 plants grown in SD were also investigated and the results indicated that BLH10 was slightly reduced in gi-2 mutants (Figure 5.4c). There was little difference in BLH3 and BLH10 expression between Col and gi-2 plants at 4 h after dawn in SD (Figure 5.4c and data not shown).

5.2.4 ANALYSIS OF BLH3 EXPRESSION IN CIRCADIAN CLOCK MUTANTS

The expression of GI is altered in circadian clock mutants. In the early flowering mutant elf3 GI transcripts are damped at higher than trough levels at all times of the day. Over-expression of the central clock genes LHY and CCA1 in the lhy and cca1 mutants results in reduction in the cycling and overall transcript levels of GI (Fowler et al. 1999). The GI transcript continues to cycle in the lhy cca1 double mutant, however the peak of GI
expression is earlier in the day and this is associated with the extreme early flowering phenotype of these plants (Mizoguchi et al., 2005). These results suggest that the circadian regulated GI output is closely linked to the function of this gene.

It is likely that neither BLH3 nor BLH10 are directly clock regulated (Chapter 5.2.1), yet the products of both these genes interact with GI. Therefore BLH3 and BLH10 may also act indirectly downstream of the clock via association with a clock regulated gene such as GI. This was tested by examining if BLH3 and BLH10 transcript levels were affected by mutations in central clock genes. Wild type Col and the gi-2, cca1 and elf3 mutants were grown in LD conditions. Plants were shifted to constant light conditions at dawn and tissue was harvested every 8 h for 24 h. The expression of BLH3 in gi-2, elf3, and cca1 mutants was tested by Northern analysis.

![Figure 5.5](image)

**Figure 5.5 Comparison of BLH3 expression in circadian clock mutants gi-2, cca1 and elf3**

**a** Plants were grown in LD and aerial tissue was harvested from 4-6 leaf plants at 0 h after dawn. Total RNA (25 µg) was extracted and expression was analysed by Northern hybridisation analysis (Chapter 2.5.6). The rRNA probe was used as a loading control. Graph of relative expression of BLH3 in Col and gi-2 mutant plants at 0 h after dawn. The proportion of the highest value after normalisation with respect to 25S/26S rRNA levels was calculated and the mean of the two samples ± SE was plotted.

**b** Total RNA (25 µg) was extracted from aerial tissue at 8 h, 16 h and 24 h after moving the above plants into LL conditions at dawn (0 h). BLH3 expression was analysed by Northern hybridisation (chapter 2.5.6). The rRNA probe was used as a loading control.

Supporting previous RT-PCR and Northern analyses, the BLH3 transcript was present at all time points analysed (Figure 5.5). As indicated by previous Northern assays, BLH3
levels were reduced in two independent samples of the late flowering gi-2 mutant at 0 h (Figure 5.5a). BLH3 levels were unchanged in gi-2 mutants at two time points after dawn (8h and 16 h, Figure 5.5b). Furthermore, 24 h after the shift to constant light (apparent dawn) there was no longer a significant difference in BLH3 transcript levels between Col and gi-2 mutants (Figure 5.5b, 24 h). The comparison of BLH3 expression in Col and gi-2 plants was repeated with triplicate samples at 0 h, 8 h and 24 h after a switch to constant light conditions. Average transcript levels of BLH3 were consistently lower in gi-2 mutants compared to Col at the 0 h (dawn) time point only (data not shown).

It was evident that BLH3 levels were also altered in the cca1 mutant compared to wild type Col at dawn (0 h) (Figure 5.5a). In comparison, BLH3 levels were not significantly affected by the elf3 mutation at any time tested (Figure 5.5a and b). In an independent experiment, BLH3 expression in the lhy-1 over-expression mutant (Ler background) was also compared to expression in the Ler wild type. This revealed that BLH3 transcript levels were not significantly altered in the lhy-1 circadian clock mutant at any time during the day tested (data not shown).

**5.2.5 BLH3 EXPRESSION IN RESPONSE TO RED LIGHT**

The expression levels of BLH3 are not greatly affected by light/dark cycles (Chapter 5.2.1.1), yet it is possible that BLH3 expression is affected by light or dark directly. Mining of microarray data provided a new perspective for the investigation of BLH3 and BLH10 expression. The BLH3 and BLH10 transcripts were present at higher levels in light grown seedlings compared to dark grown seedlings (Jiao et al., 2003). Recently, GI expression has been found to be affected by various light conditions, including red and far red enriched light (Devlin et al., 2003; Tepperman et al., 2004; Paltiel et al., 2006).

To determine if BLH3 expression levels in seedlings were affected by red light, seeds were exposed to 1 h white light to induce germination and were grown in complete darkness for 4 d. Seedlings were harvested before and after exposure to 1 h of red light (Rc). A preliminary Northern assay verified that the BLH3 transcript was present in dark grown wild type seedlings (Figure 5.6). This was expected, as the BLH3 cDNA was originally isolated from a cDNA library constructed from etiolated seedlings. BLH3 levels increased in seedlings exposed to only 1 h red light, suggestive of a rapid response to red light. The levels of the BLH3 transcript increased in both Col and gi-2 mutants (Figure 5.6).
Figure 5.6 Analysis of BLH3 mRNA levels in response to red light

Seedlings were grown in tissue culture for 4 d (Chapter 2.4.1.6). Northern hybridisation experiments (Chapter 2.5.6) were carried out using total RNA (25 µg) extracted from wild type Col and gi-2 mutant plants grown in DD and at 1 h after the shift to Rc. Partial cDNA sequence for BLH3 was used as a hybridisation probe and the hybridisation of the 25S/26S rRNA probe was used as a loading control.

Attempts to repeat this Northern on two separate occasions failed to give any result. Unfortunately, the amount and quality of RNA obtained from the seedlings was not sufficient to detect BLH3. The BLH3 transcript is not particularly highly expressed and at least 25 µg total RNA is required to detect BLH3 by Northern hybridisation. In addition, poor germination and growth of the seedlings in tissue culture contributed to the limited amount of plant tissue from which RNA was able to be obtained.

5.3 LOCALISATION OF BLH3 USING THE GFP REPORTER GENE

This section of work seeks to determine if the BLH3 protein is localised to a particular subcellular compartment. The GFP protein from the jellyfish *Aquorea victoria* was used as a reporter gene for the localisation of BLH3 within the cell. GFP has proven to be a particularly useful reporter gene for identifying the intracellular location of proteins in plants. GFP requires no cofactors for fluorescence and does not disrupt the native folding or function of the tagged protein of interest. Importantly, BLH3 can be localised in living cells using the GFP reporter gene (reviewed by de Ruitjer et al., 2003). However, the detection of the GFP reporter in cells is not as sensitive as reporter genes such as *GUS*, which have the advantage of enzymatic amplification of the signal. BLH3 is difficult to detect by Northern analysis, which implies that the level of expression obtained from the BLH3 promoter may not be sufficient to induce detectable levels of GFP. For these reasons, the constitutive 35S promoter was selected to express GFP:BLH3 in plants.
5.3.1 THE GFP REPORTER CONSTRUCTS

The 35S:GFP expression cassette used in this work was obtained from the vector pAVA-393 (von Arnim et al. 1998). This expression cassette contains a double CaMV 35S promoter and the modified GFP cDNA mGFP4 for expression in plants (Haseloff et al. 1997). BLH3 was alternately fused to the C-terminus and N-terminus of a GFP reporter construct driven by the CaMV 35S promoter. The 35S:GFP, 35SGFP:BLH3 and 35SBLH3:GFP expression vectors used in this work were constructed by A. Goldshmidt (this laboratory).

5.3.2 LOCALISING BLH3 USING TRANSIENT EXPRESSION ASSAYS

The p35S:GFP, 35SBLH3:GFP and p35SGFP:BLH3 constructs were introduced into the abaxial surface of *Arabidopsis* leaves by biolistics for transient expression assays (Chapter 2.4.5). The PDS-1000 He Biolistic Particle Delivery System (Bio-Rad Laboratories Ltd) at HORT Research (Auckland) was used to accelerate DNA coated gold particles into *Arabidopsis* leaf tissue. GFP fluorescence was visualised using blue light (450-490 nm illumination, 525 nm short pass filter for GFP emission).

5.3.2.1 OPTIMISING GFP EXPRESSION ASSAYS

Preliminary expression assays were performed using the three expression constructs described above. Initially a limited number of leaves shot with 35S:GFP control exhibited GFP fluorescence, yet this was at a much lower frequency than expected (data not shown). No fluorescence due to GFP was detected in *Arabidopsis* leaves shot with 35SGFP:BLH3 or 35SBLH3:GFP. There were several possible reasons for poor transient expression of GFP in these assays. Firstly, very low transformation efficiency was obtained, as indicated by the limited expression seen from the 35S:GFP control. Secondly, damage was severe in these leaves, which resulted in visible auto-fluorescence. Leaves also had fungus on the surface which exhibited fluorescence under blue light excitation. These factors made the detection of GFP fluorescence difficult. Finally, leaves were not stored in the dark after biolistics in the initial assays. This would not affect transformation efficiency, but would result in a less intense fluorescent signal, making cells expressing the BLH3 fusion much harder to detect, especially if the BLH3 fusion was not expressed to particularly high levels.
In response to the initial results, alterations were made to the biolistics protocol to improve the transformation efficiencies and detection of GFP. Leaves were sterilised in a weak bleach solution to reduce fungus on the surface and were placed on MS media to recover before use in biolistics. Examination of the leaves under blue light demonstrated that there was little or no auto-fluorescence due to surface fungus after the sterilisation treatment (data not shown). DNA for use in biolistics was prepared using the Qiagen Midi-prep Kit. However, this plasmid DNA was difficult to digest with restriction enzymes, indicating that the DNA was not as pure as expected. As it is essential that DNA used in biolistics is ultra-pure, this may have been a contributing factor to the poor transformation efficiencies observed. DNA was further purified by phenol:chloroform and chloroform extractions and tested by restriction enzyme digestion. This purified DNA was used for all subsequent transient expression assays. The shooting distance for the biolistic procedure was decreased to 6 cm, which is the standard for leaf tissue. Finally, to reduce the auto-fluorescence due to damage to the leaf tissue, two acceleration rates tested (400 psi and 800 psi). It is expected that 400 psi is sufficient for Arabidopsis leaf tissue and there may be less damage to the leaf as the gold particles travel at lower speed. Finally, samples were stored in complete darkness for 48-76 hours after bombardment.

5.3.2.2 GFP:BLH3 IS NUCLEAR LOCALISED

The optimised Arabidopsis transient expression assays were carried out on three independent occasions. The transformation efficiency was greatly improved and the analysis of the 35S:GFP controls showed that the majority of leaves (>70%) displayed GFP fluorescence in the expected pattern. Auto-fluorescence due to leaf damage was still visible, but these areas of leaf tissue were easily identified and avoided (Appendix 4.1a). The number of fluorescent cells per leaf was extremely variable and ranged approximately ten fold from ~10-100+. This is due to the uneven density of gold beads on the rupture disc, which tend to clump together and are unevenly distributed to the leaf tissue.

As expected, entire cells exhibited a fluorescent expression pattern in leaves treated with the 35S:GFP control constructs (Figure 5.7b). In contrast, leaves containing the 35SGFP:BLH3 constructs exhibited fluorescence in small round compartments (Figure 5.7a). It is not possible to conclude that the observed fluorescence is nuclear and not located in another subcellular organelle. However, the size observed does support nuclear localisation. Fifteen 35S:GFP and eighteen 35SBLH3:GFP images were taken. The
35SBLH3:GFP construct failed to show any GFP expression when introduced into *Arabidopsis* leaves. Control assays using water coated gold beads failed to give any fluorescence patterns like those of the 35S:GFP and 35SGFP:BLH3 constructs (Appendix 4.1b, c and d).

![Figure 5.7 Subcellular localisation of BLH3 in *Arabidopsis* and onion epidermal cells](image)

**Figure 5.7 Subcellular localisation of BLH3 in *Arabidopsis* and onion epidermal cells**

Transient expression assays were carried out with constructs encoding the GFP reporter gene (35S:GFP) and GFP fused to the N-terminal of BLH3 (35SGFP:BLH3). These constructs were introduced into the abaxial surface of *Arabidopsis* leaves and into onion epidermal peels by particle bombardment. Samples were viewed using blue light excitation to detect GFP fluorescence.

**a, b** *Arabidopsis* epidermal cells exhibiting expression of 35SGFP:BLH3 and 35S:GFP respectively. This experiment was repeated and at least 15 cells exhibiting GFP expression were photographed for each construct.

**c, d** Onion epidermal cells expressing 35SGFP:BLH3 and 35S:GFP control. At least 10 cells displaying GFP expression were photographed for each construct.

**e, f** Onion epidermal cells from **c** and **d** stained with DAPI and viewed using UV excitation (340-380 nm illumination, >425 nm emission) to visualise nuclear DNA.

Transient expression of the GFP constructs was also observed using a Leica TCS 4D confocal scanning laser microscope. The examination of cells using the scanning microscope give the impression of viewing cells in 3D, resulting in greater depth perception. These observations also suggested nuclear localisation of BLH3:GFP compared to the 35S:GFP control (Appendix 4.1e and f).

Onion epidermal cells have been shown to be particularly useful for transient expression assays. Onion cells are large and it is simple to prepare a peel containing a single cell layer,
thereby making it easier to analyse single cells expressing GFP. For these reasons, onion epidermal cells were also used to transiently express GFP and GFP:BLH3. Epidermal peels were prepared from the concave surface of the inner layer of an onion bulb and placed on MS plates not more than 4 h before use. Gold beads were introduced into the inward facing surface of onion epidermal peels by biolistics. Initially, three acceleration speeds were tested (400 psi, 800 psi and 1100 psi). Only at 1100 psi was any GFP fluorescence detected in peels shot by 35S:GFP controls (data not shown). This acceleration speed was used for all further transient assays.

A 35S:GFP control plasmid was introduced into onion epidermal cells by biolistics. Cells bombarded with 35S:GFP displayed GFP expression throughout the cell (Figure 5.7d). In contrast, 35SGFP:BLH3 fluorescence was much less intense and more localised, apparently to the nucleus (Figure 5.7c). To visualise nuclei, onion epidermal tissue was mounted in DAPI stain (1/2 MS, 20% glycerol, 2 μg/mL DAPI) (adapted from (Staiger et al., 2003) and viewed using UV light. The nuclear location of 35SGFP:BLH3 in onion epidermal cells was confirmed as GFP:BLH3 fluorescence corresponded to DAPI stained nuclei (Figure 5.7e). At least ten images representing each construct were taken. These results confirm that BLH3 is likely to be a nuclear protein.

5.3.3 Stable Expression of GFP Fusion Proteins in Transgenic Arabidopsis

The transient expression assays revealed that BLH3 is located in the nucleus in Arabidopsis and onion cells (Chapter 5.3.2). The expression of the GFP:BLH3 fusion protein in planta is expected to imitate the expression of the native BLH3 protein in plants, so may be useful to support transient expression assays. For this purpose transgenic Arabidopsis plants were generated. Constructs for the over expression of BLH3:GFP and GFP:BLH3 were created and introduced into Arabidopsis plants. The 35SBLH3:GFP and 35SGFP:BLH3 expression cassettes were subcloned from pGEMT into the T-DNA region of the binary vector pART27, creating pBIN:35SBLH3:GFP and pBIN:35SGFP:BLH3 (A. Goldshmidt, this laboratory). The T-DNA region contains the 35S:nptII gene construct, which when constitutively expressed in transgenic plants provides resistance to kanamycin.
5.3.3.1 INTRODUCTION OF EXPRESSION VECTORS INTO *AGROBACTERIUM*

As part of this thesis work, the binary vectors pBIN:35SBLH3:GFP and pBIN:35SGFP:BLH3 were transformed into *Agrobacterium* by triparental mating. The plasmids were transferred from *E. coli* DH10B to *Agrobacterium* GV3101 using a third strain, *E. coli* pRK2013 (Chapter 2.2.3.2). Briefly, each parental strain was grown on the appropriate selective plate before being mixed with other strains in pairwise and triplet combinations on non-selective plates. The bacteria were incubated overnight at 28°C to allow mating to occur. Bacteria were then plated onto YN media containing rifampicin, gentamycin and kanamycin and incubated at 28°C for 2 days. Transformed *Agrobacterium*, resulting only from the mating of the three strains, grew well on selective media. The control pair-wise combinations did not result in any transformants. Furthermore, transformants were unable to grow at 37°C, confirming that transformants were not *E. coli* but *Agrobacterium*.

The identity of the two binary vectors, carried by independent *Agrobacterium* transformants, were confirmed by restriction enzyme digestion. Transformants were used to inoculate over-night cultures, from which plasmid DNA was extracted (Chapter 2.5.1.3) and retransformed into *E. coli* (Chapter 2.2.3.1). Plasmid DNA was isolated from *E. coli* (Chapter 2.5.1.1) and analysed by restriction enzyme digestion to confirm the identity of the plasmids. The *Agrobacterium* clones containing pBIN:35SGFP:BLH3 and pBIN:35SBLH3:GFP respectively produced the expected restriction fragments (Figure 5.8) and were used to transform *Arabidopsis* (Chapter 5.3.3.2).

![Figure 5.8](image)

**Figure 5.8** Confirmation of pBIN:35SBLH3:GFP and pBIN:35SGFP:BLH3 in *Agrobacterium*

Restriction enzyme digests used to verify pBIN:35SGFP:BLH3 (**GFP:BLH3**) and pBIN:35SBLH3:GFP (**BLH3:GFP**) in two independent *Agrobacterium* clones. The gel contains *Bgl*II and *Pst*I digests and the approximate fragment sizes expected in kb are: GFP:BLH3 (clones 1 and 2) *Bgl*II, 0.5, 14.5 and *Pst*I, 0.5, 2.5, 3.4, 8.5; BLH3:GFP (clones 3 and 4) *Bgl*II, 1.7, 13.3 and *Pst*I, 0.5, 2.5, 3.4, 8.5. 1kb; 1 kb plus DNA size marker, 1, 2 and 8 kb fragments indicated beside the gel.
5.3.3.2 TRANSFORMATION OF ARABIDOPSIS AND THE SELECTION OF TRANSFORMANTS

The *Agrobacterium* clones containing the pBIN:35SGFP:BLH3 and pBIN:35SBLH3:GFP binary vectors were used to transform *Arabidopsis*, via the *Agrobacterium*-mediated floral dip method (Clough and Bent, 1998). This protocol is modified from the vacuum infiltration method of Bechtold *et al.* (1993) and results in higher transformation efficiencies (Clough and Bent, 1998) (Chapter 2.4.2). In summary, the aerial tissue of *Arabidopsis* plants are submerged into a suspension of *Agrobacterium* cells which carry the binary plasmid containing the transgene of interest. A vacuum (-80 kPa) is applied briefly to facilitate the infiltration of bacteria into the intracellular spaces within the floral tissue. Plants are grown to maturity and allowed to self fertilise.

*Agrobacterium* containing the expression constructs (Chapter 5.3.3.1) were used to inoculate cultures for plant transformation. Bacterial cells were harvested from large overnight cultures by centrifugation and resuspended in IM. This suspension was used to transform *Arabidopsis* as described in Chapter 2.4.2.2. Wild type Col and gi-2 mutant plants were transformed using this floral dip method. Plants were allowed to recover and set seed which was harvested in bulk. Transformation of plants with the two constructs was repeated on two separate occasions.

Antibiotic selection was used to identify plants that carried the T-DNA insertions. Primary transformants (T1 plants) were identified as seedlings resistant to kanamycin. Seed was collected from transformed plants, sterilised and spread at a density of 4000 seed/plate onto GM plates containing kanamycin (Chapter 2.4.4.1). Transformants were identified 2-3 weeks after germination as green healthy seedlings that had produced roots and true leaves. No gi-2 transformants containing the GFP expression constructs were obtained. Twenty healthy transformants representing 35SBLH3:GFP/Col and 35SGFP:BLH3/Col were selected for transfer to soil and grown to maturity. The plants were allowed to self fertilise and the T2 seed was collected from each plant individually.

5.3.3.3 EXAMINATION OF GFP EXPRESSION IN TRANSGENIC PLANTS

The transgenic plants containing 35SBLH3:GFP and 35SGFP:BLH3 generated in the previous section of work were used in GFP expression assays. Twenty plant lines were selected and T2 generation plants were grown on GM plates containing kanamycin to select for plants containing the T-DNA insertions. Transgenic plants containing 35S:GFP and 35SGFP:GI were obtained from Erwin Krauskopf (this laboratory) and Karen Lee (John
Innes Centre, Norwich) respectively for use as GFP controls. These plants were also grown on GM plates before use in expression assays. Plants were dissected and mounted in water on standard microscope slides to view GFP expression.

To determine the level of autofluorescence in *Arabidopsis* plants and assess if this would impede the detection of GFP specific fluorescence, wild type Col plants were examined using fluorescent microscopy. Very low levels of green fluorescence were detected in green leaf tissue (Appendix 4.2b), signifying that the autofluorescence in these plants was not significant and would not interfere with GFP detection. Fluorescence due to GFP was detected throughout the leaf tissue in plants containing the 35S:GFP construct (Appendix 4.2a). The entire leaf was bright green in comparison to wild type plants. The pattern of GFP expression appeared to be concentrated around the edges of cells, as the clear outline of the epidermal cells could be seen.

The 35SBLH3:GFP plants lines tested exhibited no GFP fluorescence. This was not unexpected, due to the lack of GFP expression from this construct in previous transient expression assays (Chapter 5.3.2). The 35SGFP:BLH3 and 35SGFP:GI plants appeared brighter green than wild type plants, indicating that these plants may be expressing GFP, although the levels of fluorescence were not considerably higher than background levels (Appendix 4.2c and d). No obvious nuclear localisation of GFP expression was seen in either the 35S:GFP:BLH3 nor the 35S:GFP:GI control plants. This experiment was repeated on two more independent occasions and similar results were obtained.

### 5.3.3.4 Why was GFP not detected in transgenic plants?

There are several possible reasons for the lack of expression from the 35SGFP:BLH3 construct in transgenic *Arabidopsis* plants. Firstly, lack of expression could be contributed to a mutation in the expression construct. However, the 35SGFP:BLH3 expression cassette successfully expressed GFP:BLH3 protein fusions in transient assays, so it is unlikely that a mutation in this construct arose during subcloning into the binary vector. The most likely explanation is that the sensitivity of detection of the microscope used was simply not sensitive enough to detect the levels of GFP:BLH3 protein generated by this construct. This is supported by the fact that no GFP fluorescence was detected in the 35SGFP:GI control plants. Previous work with the 35SGFP:GI transgenic plants has shown nuclear localised GFP:GI expression, although this was difficult to detect (Karen Lee, John Innes Centre).
This analysis may be further complicated by protein stability and/or protein turnover. The expression of the BLH3 protein during the day is unknown, yet the GI protein cycles and is not detectable at all times of the day (David et al., 2005). It is possible that the fusion proteins were not present at high levels during the middle of the day when transgenic plants were examined.

Interestingly, the 35SBLH3:GFP expression construct failed to express the BLH3:GFP fusion protein both in transient assays and transgenic plants. It is unlikely that mutations within the construct are to blame for the lack of expression, as identical constructs were used to generate both the 35SGFP:BLH3 and 35SBLH3:GFP expression cassettes. However, as the constructs were not sequenced, the presence of mutations cannot be discounted. One may speculate that the fusion of GFP to the C-terminal of BLH3 affects either the folding or stability of the BLH3 protein, whereas fusion of GFP to the N-terminal does not affect the BLH3 protein. In summary, this experiment was inconclusive for locating the BLH3 protein within cells in transgenic Arabidopsis plants.

5.4 Discussion

5.4.1 Comparison of the expression of GI, BLH3 and BLH10 in plants

Northern and RT-PCR assays have demonstrated that like GI, the BLH3 and BLH10 genes are expressed in all tissues and at all stages of development tested. Similarly to GI, BLH3 expression is increases in older plants compared to seedlings. In contrast to BLH3, the BLH10 transcript was highest in younger plants. Further distinctions between these genes were evident at the tissue level, as the expression of BLH3 and BLH10 in rosette leaves was high and low respectively. These results demonstrate that despite the significant similarities between these genes, the expression patterns of BLH3 and BLH10 in plants are not alike.

As the expression patterns of GI, BLH3 and BLH10 overlap it is plausible that BLH3 and/or BLH10 to interact with GI in planta. However, the broad expression patterns observed for BLH3 and BLH10 do not provide any additional information on where or when in plant development that the GI::BELL interaction may occur. Cole et al. (2006) have utilised an in situ hybridisation technique to demonstrate that BLH3 is expressed in
peripheral zones within the indeterminate inflorescence and determinate floral meristems. Unfortunately, the location of GI within meristematic tissue remains unknown, as previous *in situ* hybridisation assays have failed to detect GI (Fowler, 2000). As GI is expressed to sufficient levels to be detected by Northern analysis, it was proposed that GI would not be identified by *in situ* hybridisation if the GI transcript was present at lower levels throughout the tissues tested as opposed to concentrated expression in a distinct zone of cells (Fowler, 2000). If this is the case, it is possible that the GI::BLH3 interaction in plants is limited to the regions where BLH3 is expressed. Contrasting results have been obtained in wheat, where *in situ* hybridisation has verified that TaGI is localised within specific leaf epidermal cells, where the expression levels vary during daily cycles. In addition, TaGI was also detected at comparatively constant levels at the SAM (Zhao et al., 2005). Further *in situ* hybridisation analyses would be required to determine how this relates to the location of GI in *Arabidopsis*.

Additional testimony towards the biological relevance of the GI:BLH3 protein interaction comes from the subcellular localisation of these proteins. Previously, Huq et al. (2000) revealed that despite encoding putative transmembrane domains, GI was nuclear localised in transient expression assays. Assays utilising transgenic *Arabidopsis* expressing 35S:GI:GFP demonstrate that GI:GFP is nuclear located and is functional in promoting flowering, providing strong evidence that GI functions in the nucleus (Mizoguchi et al., 2005). Transient expression assays in *Arabidopsis* and onion tissue have confirmed that BLH3 too is a nuclear protein. Both GI and BLH3 localised to the nucleus in onion tissue, which implies that these proteins are nuclear by default and do not require other protein partners from *Arabidopsis* to facilitate this localisation. This is supported by recent evidence which verifies that BLH3 is nuclear localised by default. Furthermore, the interaction of BLH3 with the KNOX protein STM is necessary for the nuclear import of STM to the nucleus (Cole et al., 2006). This raises the possibility that BLH3 interacts with different proteins outside and inside the nuclear environment.

Proof of the GI:BLH3 interaction *in planta* could be obtained using Fluorescence Resonance Energy Transfer (FRET) to monitor complex formation between these two proteins in plants. An attempt was made to co-localise BLH3 and GI in plants using this method, yet the results were inconclusive (Jo Putterill, this laboratory). Antibodies were raised to BLH3 (Appendix 5) that may be utilised to analyse the expression of BLH3 at the protein level. Where and when the BLH3 protein is expressed in plants could be
determined. A co-immunolocalisation assay is one method that may prove useful in identifying the precise domains in which the GI and BLH3 proteins are expressed and where these overlap in plants.

5.4.2 BLH3 EXPRESSION IN CIRCADIAN CLOCK MUTANTS

The BLH3 and BLH10 transcripts are not directly regulated by the circadian clock as they do not fluctuate significantly through day/night cycles. Yet, the transcript levels of the BLH3 and BLH10 genes were notably lower in the gi-2 mutant background. Interestingly, this reduction in expression occurred only at dawn in plants maintained in daily cycles, whether LD and SD conditions. These results were quite unexpected as both the GI transcript and the GI protein are undetectable at dawn (Fowler et al., 1999; David et al., 2006). For this reason, the lower expression of BLH3 and BLH10 at dawn in gi-2 plants is not likely to be a direct consequence of the loss of GI.

BLH3 expression is not significantly affected by the elf3 and lhy-1 mutations. Both the cyclical expression of GI and circadian clock function are altered in these mutant plants. Therefore, neither the wild type cycling of the GI transcript, nor an intact circadian clock is necessary to maintain BLH3 expression. Yet, the expression of BLH3 in the cca1 mutant background is significantly lower than wild type, again at the dawn time point. CCA1 is normally expressed with a peak at dawn and the cyclical pattern of CCA1 expression is disrupted in both cca1 and gi mutants. In addition, cca1 and gi-2 mutants have impaired circadian clock function and thus a reduced ability to detect day length. The circadian clock is required by plants to anticipate the transition from dark to light at dawn, in consequence BLH3 and BLH10 may be indirectly affected by the clock, particularly at dawn. The lower expression levels of BLH3 observed at dawn in gi-2 mutants is overcome when plants are grown in LL conditions, which lack a dark/light transition. This is suggestive that light and/or dark could have a role in this phenotype.

The BLH3 transcripts are scarcely detectable by Northern hybridisation in dark grown seedlings, yet are measurable one hour after the transfer of seedlings to red light. This increase of expression in response to light occurs even in the gi-2 mutant background. However, as attempts to repeat this experiment were not successful, it can not be concluded that BLH3 expression is affected by light without further expression assays. Detecting BLH3 by Northern assays is challenging; expression levels of this gene are
relatively low and **BLH3** is only consistently detected by this method with an increase in the amount of total RNA are used. Due to the subtle phenotypes observed in this expression work and the difficulty in obtaining sufficient RNA from seedlings for multiple Northern assays, the response of **BLH3** to light would be more accurately quantified by real time RT-PCR.

Light of almost any quality has an affect on **GI**. The expression of **GI** is low in dark grown plants and the **GI** transcript increases in response to white, red, blue and far red light or shade conditions (Devlin *et al*., 2003; Hudson *et al*., 2003; Tepperman *et al*., 2004; Paltiel *et al*., 2006). **GI** is likely to act downstream of **PHYB**, as microarray analyses by Tepperman *et al*. (2004) identified **GI** as a gene up regulated in reaction to red light and that **PHYB** is required for this response. However, the induction of **GI** expression in response to shade and blue light is mediated by **PHYA** and cryptochromes **CRY1** and **CRY2** respectively (Devlin *et al*., 2003; Paltiel *et al*., 2006). It would be interesting to compare the affect of various light quantities and qualities on the expression of **BLH3** and **BLH10** to determine if the expression of these genes is affected by light. The **GI** protein cycles, exhibiting a peak late in the day and is actively degraded during the dark period (David *et al*., 2006). It would be useful to establish if the expression of the **BLH3** protein mimics mRNA expression and is present at all times during the day using the antibodies raised to **BLH3** (Appendix 5). The effect of light/dark cycles and circadian clock mutations on **BLH3** protein levels may prove more informative than the analysis of the **BLH3** transcript.

### 5.4.3 Summary

The **BLH3** and **BLH10** genes are widely expressed in *Arabidopsis*; transcripts are detectable throughout development and at all times examined during day/night cycles. Although the levels of the **BLH3** and **BLH10** transcripts are not regulated directly by the circadian clock, **BLH3** transcript levels are lower in clock mutants **gi-2** and **cca1** at dawn when plants are maintained in driven light/dark cycles. As expected for a putative transcription factor, the **BLH3** protein is localised in the nucleus, a subcellular compartment also occupied by **GI**. In the following chapter, reverse genetics is employed to provide more information on the function of the **BLH3** and **BLH10** genes and how this relates to the role of **GI** in plants.
6.1 Introduction

The analyses of the expression profiles of \(BLH3\) and \(BLH10\) were presented in Chapter 5. This penultimate chapter describes the characterisation of \(BLH3\) and \(BLH10\) using common plant molecular biology techniques.

Mis-expression of genes, either by increasing or decreasing transcript levels, has proven to be a useful technique for examining gene function. Plants with reduced expression of \(BLH3\) and \(BLH10\) were identified and used for further investigation into the role of these genes in plants. The \(GI\) gene is proposed to function in flowering and light signalling pathways, as \(gi\) mutant plants are late flowering and display elongated hypocotyls when grown in red light (Fowler et al., 1999; Huq et al., 2000). In consequence, the \(blh3\) and \(blh10\) mutants were characterised with respect to flowering time and hypocotyl elongation to determine if the roles of these genes \(in planta\) had aspects in common with \(GI\).

The expression of \(GI\) is controlled by the circadian clock. The loss of \(GI\) expression affects the expression of clock component genes such as \(CCA1\) and \(LHY\), indicating that \(GI\) also acts as part of a feedback loop to regulate the clock (Fowler et al., 1999). Therefore, the effect of reduced \(BLH3\) expression on the transcript levels of circadian regulated \(LHY\) was also addressed in this work. The over-expression of \(GI\) induces early flowering. Finally, transgenic plants containing 35S:GI in a \(blh3\) mutant background were generated to determine if \(BLH3\) is required downstream of \(GI\) to promote flowering.
6.2 OVER-EXPRESSION OF BLH3 AND BLH10 IN TRANSGENIC PLANTS

Increased expression of GI promotes flowering in transgenic Arabidopsis (Milich, 2001; Mizoguchi et al., 2005). To determine the role of the BLH3 and BLH10 genes the effects of over-expressing these genes in transgenic plants was investigated. For this purpose, the expression vectors p35S:BLH3 and p35S:BLH10-R were constructed and introduced into Arabidopsis wild type Col and gi-2 mutant plants. The generation of transgenic lines and the examination of these lines with respect to flowering phenotypes will be described.

6.2.1 GENERATION OF BINARY VECTORS FOR USE IN PLANT TRANSFORMATION

Two expression vectors were generated in this section of work. The BLH3 and BLH10 cDNA sequences were cloned downstream of the CaMV 35S viral promoter, which is expected to drive constitutive expression of these genes in plants (references in Gleave, 1992). The expression cassettes were inserted within the T-DNA region of the binary vector pVK. The pVK plasmid is a pART27 based plasmid with the 35S promoter and ocs 3’ sequences inserted within the NotI site (Chapter 2.1.3). The T-DNA contains a pnos:nptII gene construct, which is constitutively expressed in transgenic plants and confers resistance to kanamycin.

6.2.1.1 CONSTRUCTION OF P35S:BLH3

The full length BLH3 cDNA was excised from pZL-BLH3 (Chapter 2.1.3) by digestion with EcoRI and XbaI and the 2.1 kb BLH3 fragment was isolated. The pVK vector was digested with EcoRI and XbaI and ligated with the BLH3 fragment to create p35S:BLH3 (see Figure 6.2). The ligation mixture was used to transform E. coli and plasmid DNA was isolated from selected spectinomycin resistant colonies. Clones containing the 2.1 kb BLH3 insert were identified by restriction enzyme digestion (data not shown) and one was selected for further experiments (Chapter 6.2.2).

6.2.1.2 CONSTRUCTION OF P35S:BLH10-R

At the time of this work a cDNA encoding a full length BLH10 protein had not yet been isolated. A RAFL BLH10 cDNA was obtained that encoded a truncated protein (BLH10-R) (discussed in Chapter 3.4.2). It was unlikely that this cDNA encoded a functional
transcriptional factor, as the putative protein lacked the homeodomain region. Yet it was possible that the truncated BLH10 protein played some other role in plants, so the BLH10-R cDNA was selected for over expression in plants. The BLH10-R cDNA was released from pBS-BLH10 in two fragments (0.85 kb and 1.5 kb) by restriction enzyme digestion with EcoRI and KpnI. The restriction fragments were separated by electrophoresis through an agarose gel and purified. The pVK plasmid was digested with EcoRI and KpnI and ligated with the smaller BLH10-R fragment (0.85 kb) to create pR26.8. The resulting constructs were transformed into E. coli and plated onto selective media containing spectinomycin. Plasmid DNA was extracted from selected resistant colonies and examined by restriction enzyme digestion to confirm the presence of the 0.85 kb BLH10 insertion (data not shown).

A single pR26.8 clone was digested with KpnI and phosphatase treated to prevent self ligation. The remaining 1.5 kb fragment of BLH10, flanked by KpnI overhanging ends, was ligated with KpnI digested pR26.8 to create p35S:BLH10-R. Plasmid DNA was extracted from twelve spectinomycin resistant colonies and analysed by restriction enzyme digestion to confirm the presence of the 1.5 kb insertion. Five clones were further analysed by digestion with BglII to determine the orientation of the 1.5 kb insert (data not shown). Two clones contained the 1.5 kb BLH10 insert in the correct orientation and one was selected for further experiments (Chapter 6.2.2).

6.2.2 INTRODUCTION OF P35S:BLH3 AND P35S:BLH10-R INTO ARABIDOPSIS

6.2.2.1 TRANSFORMATION OF AGROBACTERIUM WITH TWO BINARY CONSTRUCTS

The p35S:BLH3 and p35S:BLH10-R binary constructs generated in Chapter 6.2.1 were transformed into Agrobacterium by triparental mating, as described previously (Chapter

---

**Figure 6.1** Simplified restriction enzyme map of p35S:BLH10-R

The restriction map is not drawn to scale. The restriction enzyme sites for BglII (B), HindIII (H), EcoRI, KpnI, NotI and PstI are indicated. The direction of open reading frames are indicated by the arrows. 35S, CaMV 35S promoter; 3', ocs 3' terminator; T-DNA left and right borders.
Independent *Agrobacterium* transformants were analysed by restriction enzyme digestion to confirm the identity of the plasmids. Briefly, plasmid DNA was extracted and retransformed into *E. coli*. Subsequently, plasmid DNA was isolated from *E. coli* and analysed by restriction enzyme digestion to confirm the identity of the plasmid. *Agrobacterium* clones containing either p35S:BLH3 or p35S:BLH10-R produced the expected restriction fragments (Figure 6.2 and data not shown). These clones were used to transform *Arabidopsis* (Chapter 6.2.2.2).

**Figure 6.2** Restriction enzyme digests used to verify p35S:BLH3 in *Agrobacterium*

a Restriction enzyme map of p35S:BLH3 (not drawn to scale). The restriction enzyme sites for HindIII (H), NotI, PstI, EcoRI, EcoRV and XbaI are indicated. The direction of open reading frames are indicated by the arrows. 35S, CaMV 35S promoter; 3', ocs 3' terminator; T-DNA left and right borders.

b Plasmid DNA was extracted from *E. coli* (Chapter 2.5.1.1) and subject to restriction enzyme digestion (Chapter 2.5.3.6). The gel contains HindIII/EcoRV and PstI digests of three independent clones (1, 2, 3). The approximate fragment sizes expected in kb are: HindIII/EcoRV; 0.5, 0.7, 1.6, 5.2, 7.4 and PstI; 2.5, 4.6, 8.5, 1 kb; 1 kb + DNA size marker.

### 6.2.2.2 *Agrobacterium*-mediated Transformation of *Arabidopsis*

The p35S:BLH3 and p35S:BLH10-R binary constructs were transformed into *Arabidopsis* plants via *Agrobacterium*-mediated transformation, as described in Chapter 5.3.3.2. The transformation results are shown in Table 6.1. The transformation efficiencies obtained were disappointing, as it is possible to achieve transformation efficiencies up to 5% using this method (Milich, 2001). This was most likely due to the wet conditions in the greenhouse at the time of the experiment and the resulting poor condition of the plants. In addition, the general poor condition of the plants contributed to lower seed production than expected, which may have further affected the transformation efficiency.
Table 6.1 Results of Arabidopsis transformation experiments

<table>
<thead>
<tr>
<th>Construct/Genotype</th>
<th>Plants transformed</th>
<th>Transformation efficiency (%)</th>
<th>T1 plants</th>
</tr>
</thead>
<tbody>
<tr>
<td>p35S:BLH3/Col</td>
<td>36</td>
<td>0.55</td>
<td>40</td>
</tr>
<tr>
<td>p35S:BLH3/gi-2</td>
<td>36</td>
<td>0.25</td>
<td>40</td>
</tr>
<tr>
<td>p35S:BLH10-R/Col</td>
<td>18</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>p35S:BLH10-R/gi-2</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

i Construct used and genotype transformed, Col wild type and gi-2 mutant plants in Col background. ii Calculated as number of transformants per 1000 seed. iii T1 plants are the number of plants replanted from selective plates to soil and represent only a small proportion of the total number of transformants.

Despite the low transformation efficiency obtained, T1 transformants representing the p35S:BLH3/Col and p35S:BLH3/gi-2 plant lines were identified. Healthy transformants were selected and transferred from tissue culture to soil. The majority of seedlings from p35S:BLH3 plant lines survived and were grown to maturity. The plants were allowed to self fertilise and the T2 seed was collected from each plant individually.

Unfortunately, no transformants containing 35S:BLH10-R were identified, although all seed collected after transformation (~2 g) was screened. Due to the identification and isolation of a new cDNA which encoded a full length BLH10 protein (Chapter 3.4.2.2), the attempt to generate these transgenic lines was not repeated.

6.2.3 ANALYSIS OF TRANSGENIC PLANTS CONTAINING P35S:BLH3

The 35S:BLH3 expression construct was transformed into two plant types, Col wild type and gi-2 mutants (Chapter 6.2.2). The progeny of surviving transformants were analysed further in this section of work.

6.2.3.1 VERIFYING THE PRESENCE OF THE P35S:BLH3 INSERTION IN TRANSGENIC LINES

PCR amplification was used to establish if the p35S:BLH3 plant lines generated in Chapter 6.2.2 contained the p35S:BLH3 insertion. Twenty p35S:BLH3/Col and twenty-six p35S:BLH3/gi-2 lines were selected and antibiotic resistant T2 seedlings were planted into soil. Genomic DNA was extracted from pooled leaf tissue harvested from these plants (Chapter 2.5.2.1) and used as a template for PCR. The 35S:BLH3 transgene was amplified by PCR using a 35S promoter and BLH3 specific primer combination (Figure 6.3a). Only
seven out of the twenty p35S:BLH3/Col lines were positive for the p35S:BLH3 construct (Figure 6.3b). Moreover, only eighteen of the twenty six p35S:BLH3/ gi-2 lines contained the 35S:BLH3 insertion (data not shown). This was unexpected, as these plants are likely to contain the T-DNA as they are resistant to kanamycin. Only the lines that tested positive for the p35S:BLH3 insertion were used in further assays.

![Figure 6.3](image)

**Figure 6.3  Identification of transgenic plants containing p35S:BLH3**

a  A schematic representation of the primer positions within the p35S:BLH3 insertion.

b  Identification of plants containing the 35S:BLH3 expression construct. Genomic DNA was extracted from T2 plants resistant to kanamycin (Chapter 2.5.2.1) and used as a template for PCR amplification (Chapter 2.5.4). A 35S promoter specific primer (35S.F) and BLH3 gene specific reverse primer (3.1R) were used to amplify the junction between 35S promoter and BLH3. The expected product of 750 bp was visualised on an agarose gel (Chapter 2.5.3.2). 1-14; selected plant lines, +ve; p35S:BLH3 plasmid positive control, H2O; negative water control, 1kb; 1 kb + DNA size marker.

### 6.2.3.2 Assessing the Phenotype of Transgenic Plants

To investigate if the over expression of BLH3 had any effects on plant development, T2 generation plants from seven independent p35S:BLH3/Col lines (Chapter 6.2.3.1) were selected for analysis. The over expression of GI results in early flowering in Arabidopsis, for that reason the effect of BLH3 over expression on flowering was examined. T2 generation plants that were segregating for the T-DNA insert were grown in long day conditions. Col plants transformed with p35S:BLH3 flowered at approximately 9-10 leaves. This was the same number of leaves produced at flowering as that of wild type Col controls (data not shown).

The over-expression of BELL genes have been associated with dwarfism, reduced fertility and altered patterning and phyllotaxy of leaves and flowers (Dong et al., 2000; Muller et al., 2000; Chen et al., 2003; Cole et al., 2006). The 35S:BLH3/Col plants were examined for developmental phenotypes. There were no obvious effects on other areas of plant development examined, such as phyllotaxy of rosette leaves and flowers, stem elongation or floral development. These plant lines visibly resembled Col wild type plants.
Selected p35S:BLH3/gi-2 plant lines that contained the p35S:BLH3 insertion (Chapter 6.2.3.1) were also analysed for flowering time. These plants were late flowering, as were the parent gi-2 plants (data not shown).

6.2.3.3 Confirmation of T-DNA Insertion Copy Number

Due to the lack of obvious phenotype in the 35S:BLH3 transgenic lines, it was possible that the 35S:BLH3 construct was not being expressed. Multiple T-DNA insertions are often associated with silencing or co-suppression of transgene expression in transgenic plants (reviewed by Fagard and Vaucheret, 2000). The number of independent T-DNA insertions within the plant genome can be established by the segregation of the antibiotic resistant marker in the progeny of the transgenic plants (T2 plants) (Chapter 2.4.4.2).

The seven p35S:BLH3/Col transgenic lines that tested positive for the p35S:BLH3 insertion (Chapter 6.2.3) were selected for further analysis. Two lines did not set seed and as a result only five p35S:BLH3/Col lines were subject to segregation analysis. All lines tested exhibited ratios consistent with a single insertion event (Table 6.2). The lack of phenotype seen in these transgenic plants is not due to silencing caused by multiple independent T-DNA insertions. However, it is possible that multiple T-DNAs have inserted at a single locus.

<table>
<thead>
<tr>
<th>Line</th>
<th>Observed</th>
<th>3:1 Expected</th>
<th>n</th>
<th>$\chi^2$</th>
<th># inserts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kan$^R$:Kan$^S$</td>
<td>Kan$^R$:Kan$^S$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K5</td>
<td>78:30</td>
<td>81:27</td>
<td>108</td>
<td>0.444</td>
<td>1</td>
</tr>
<tr>
<td>K12</td>
<td>61:27</td>
<td>66:22</td>
<td>88</td>
<td>1.515</td>
<td>1</td>
</tr>
<tr>
<td>K17</td>
<td>67:17</td>
<td>63:21</td>
<td>84</td>
<td>1.016</td>
<td>1</td>
</tr>
<tr>
<td>K18</td>
<td>27:10</td>
<td>28:9</td>
<td>37</td>
<td>0.147</td>
<td>1</td>
</tr>
<tr>
<td>K19</td>
<td>56:18</td>
<td>55:19</td>
<td>74</td>
<td>0.071</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6.2 Inheritance of the T-DNA in 35S:BLH3/Col transgenic lines

T2 seed was sterilised and plated onto GM containing kanamycin (Chapter 2.4.4.2). Antibiotic resistant plants were identified as healthy plants with good leaf and root growth, whereas antibiotic sensitive plants failed to grow roots and/or true leaves. Chi-squared values were calculated as described in Chapter 2.4.4.2. For chi-squared values of less than 3.841 (p, 0.05, one degree of freedom) the null hypothesis that these plant lines contained a single T-DNA insertion was accepted. Kan$^R$:Kan$^S$; kanamycin resistant versus kanamycin sensitive, n; plant number, this should be ~100, some lines exhibited poor germination.
6.2.3.4 ANALYSIS OF BLH3 EXPRESSION IN TRANSGENIC PLANTS

The p35S:BLH3/Col lines were analysed for flowering time and other developmental phenotypes (Chapter 6.2.3.2). It was unexpected that the over expression of a transcription factor such as BLH3 would have no effect on plant development. This raised the possibility that BLH3 was not being over expressed by the 35S promoter in these plant lines. Five independent lines known to contain the p35S:BLH3 construct (Chapter 6.2.3.1) and a single T-DNA insertion (Chapter 6.2.3.3) were analysed for BLH3 transcript levels.

At the time when this experiment was carried out, BLH3 was unable to be consistently detected by Northern analysis (Chapter 5.2), therefore RT-PCR was employed to visualise BLH3 expression. Plants were grown in LD conditions in growth cabinets and aerial tissue was harvested at the 5-6 leaf stage. Total RNA was extracted and purified and each sample (4 µg) was used as a template in two independent reverse transcription (RT) reactions. The first strand reaction was diluted ten fold and used as a template for at least two independent PCR amplification reactions. This experiment indicated that the levels of BLH3 expression in the transgenic plant lines were approximately equivalent to that in Col wild type plants (Figure 6.4), although it must be noted that these results were not absolutely quantitative. Due to the lack of an obvious phenotype, these plant lines were not analysed further during this thesis work.

Figure 6.4 Examination of BLH3 expression in p35S:BLH3/Col transgenic plants
Analysis of BLH3 expression in transgenic plants containing p35S:BLH3 by RT-PCR. Total RNA (4µg) (Chapter 2.5.2.4) was used as a template for reverse transcription, followed by amplification with gene specific primers (Chapter 2.5.4). Amplification products of BLH3 (480 bp) of the expected size were visualised on an agarose gel (Chapter 2.5.3.2). The ubiquitin (UBQ) RT-PCR was used as an internal control. 1-5: five independent transgenic plant lines, Col; wild type control, N; wild type no RT control, H2O; water control.
Chapter 6

6.3 IDENTIFICATION OF BLH3 AND BLH10 MUTANT PLANTS

This section describes the use of reverse genetics to determine the function of the BLH3 and BLH10 genes in plants. The Arabidopsis Information Resource (TAIR, www.arabidopsis.org) allows researchers to search enormous seed resources that are readily available from the Arabidopsis Biological Resource Centre (ABRC) (Ohio, USA). Populations of Arabidopsis mutants have been generated by a number of methods, including T-DNA insertion and mutagenesis. These were screened using BLH3 and BLH10 sequences to identify genetic mutants, with the goal of using plant mutant phenotypes to work backwards to determine the function of these genes.

6.3.1 SCREENING OF WISCONSIN LINES

Previous to this thesis work, primary screening of the Wisconsin T-DNA insertion lines (Sussman et al., 2000) was carried out to identify possible T-DNA insertions within the BLH3 gene. The first two rounds of PCR screening were performed at the Biotechnology Centre (University of Wisconsin) using a T-DNA and BLH3 gene specific primer pair. The results were analysed by Carly Pullen (this laboratory), which revealed that a T-DNA insertion was present approximately 1.3 kb upstream of the predicted start site of the BLH3 gene. The seed pool containing the putative BLH3 T-DNA insertion line was obtained from the ABRC.

The third and final round of screening was performed as part of this thesis work. The Wisconsin seed pool H86 contained 25 seed pools, each consisting of approximately 250 seed from nine plant lines. The seed was sterilised and each pool was plated onto GM plates containing kanamycin to select for plants containing the T-DNA. DNA was extracted from each pool of plants and used as a template for PCR. As expected, only one pool of the twenty-five produced the expected amplification product after PCR with the T-DNA and BLH3 specific primers (data not shown). Tissue was harvested from individual plants from the positive pool, DNA was extracted and again subject to PCR to identify individuals which contained the T-DNA insertion. Three plants were positive for the BLH3 T-DNA insertion (data not shown). PCR was performed on DNA from the three positive plants using BLH3 specific primers to determine if any of these plants were homozygous for the T-DNA insertion. One plant did not produce a BLH3 specific band, indicating that this individual was homozygous for the insertion (data not shown).
The expression of $BLH3$ in plants containing the T-DNA was examined by Northern analysis. The expression of the $BLH3$ was equal to wild type levels, both in plants thought to be heterozygous and homozygous for the T-DNA insertion (data not shown). The T-DNA insertion upstream of $BLH3$ was not sufficient to knock out $BLH3$ expression. These plant lines were not used for any further work.

### 6.3.2 Screening of Syngenta Lines

An alternate T-DNA insertion population, known as the Syngenta *Arabidopsis* Insertion Library (SAIL), was screened *in silico* for the presence of a T-DNA insertion within the $BLH3$ and/or $BLH10$ genes. The SAIL insertion collection was generated from approximately 100,000 individual T-DNA mutagenised *Arabidopsis* plants (Col ecotype) (Sessions *et al*., 2002). The sequences flanking the T-DNA insertions were screened using the $BLH3$ and $BLH10$ genomic sequences as bait. A line carrying a T-DNA insertion within the 5’ UTR of $BLH3$ was designated the $blh3$ mutant. A line with a T-DNA insertion within the first exon of the $BLH10$ coding region was called $blh10$. The seed harvested from these two lines were obtained from the ABRC for further analyses.

#### 6.3.2.1 Back-crossing $BLH3$ and $BLH10$ Mutants

The genotype of the $blh3$ and $blh10$ mutant plants lines acquired from the ABRC were confirmed by PCR (data not shown). Plants containing the T-DNA insertions were back-crossed to wild type Col. Progeny of the cross were allowed to self fertilise and the F2 progeny were analysed by PCR to identify individuals homozygous for the T-DNA insertions. Plants that were positive for the product representing the T-DNA insertion and negative for the product representing the wild type gene were assumed to be homozygous for the insertions into the $BLH3$ or $BLH10$ genes (Figure 6.5b). These were $blh3-1$, $blh3-3$, $blh10-1$ and $blh10-34$.

The T-DNA contains a selectable marker gene that confers resistance to the herbicide BASTA. The progeny of back-crossed $blh3$ and $blh10$ mutants were 100% resistant to BASTA, confirming that these plants were homozygous for the T-DNA insertion. Southern analyses were also carried out using $BLH3$ and $BLH10$ specific probes and the results verified that the back-crossed $blh3$ and $blh10$ lines respectively were homozygous for the T-DNA insertions (data not shown). The majority of the phenotypic analyses described in this chapter were carried out using the $blh3-1$ and $blh10-1$ T-DNA mutant lines.
Figure 6.5 Identification of blh3 and blh10 insertion mutants

a A schematic representation of the position of the T-DNA within the BLH3 and BLH10 genomic sequence. T-DNAs indicated by black triangle, black and grey bars represent exons and introns respectively, the grey striped bar represents the UTR regions. **LB3**: T-DNA left border primer, 3.1R, 3.2R; **BLH3** gene specific reverse primers, 3.1F; **BLH3** gene specific forward primer, 10.1F; **BLH10** gene specific forward primer, 10.1R, 10.2R, **BLH10** gene specific reverse primers.

b Identification of plants homozygous for the T-DNA insertions. Mutants were crossed to Col (Chapter 2.4.3) and allowed to self-fertilise in the next generation. Genomic DNA was extracted from F2 plants (Chapter 2.5.2.1) and used as a template for PCR (Chapter 2.5.4). A T-DNA specific primer (LB3) and gene specific reverse primer were used to amplify the junction between T-DNA and genomic sequence (LB3+3.1R, LB3+10.2R). Wild type genes were amplified using gene specific primers only (**BLH3**, 3.1F+3.2R; **BLH10**, 10.1F+10.1R). H2O: negative control, +: positive control, 3-1, 3-2, 3-3: **blh3** F2 plants, 10-1, 10-8, 10-34: **blh10** F2 plants.

6.3.2.2 CHARACTERISING THE T-DNA INSERTIONS IN THE BLH3 AND BLH10 MUTANTS

The insertion of T-DNA sequences is often accompanied by a deletion of genomic DNA. To ensure that the T-DNA insertions within the **BLH3** and **BLH10** genes were not associated with a deletion of upstream or downstream genes, the exact position of the T-DNA insertions were confirmed by PCR amplification and sequencing. The sequences flanking the T-DNA insertions were amplified using several primer sets. The number and orientation of T-DNA insertions was deduced by the presence or absence of an amplification product with different primer combinations (Figure 6.6). The PCR results indicated that more than one T-DNA has inserted at both the **BLH3** and **BLH10** loci in the **blh3** and **blh10** mutants respectively (Figure 6.6a and b).

The PCR products obtained were purified and subject to direct sequencing using the LB and RB amplification primers as sequencing primers. Sequencing from the LB of the T-DNA in the **blh3** mutant confirmed that T-DNA is inserted within the 5’UTR of **BLH3**, 555 bp upstream of the predicted translation start site. The **blh3** T-DNA insertion is flanked on both sides with LB sequence, as shown in Figure 6.6a, supporting the assumption that
more than one T-DNA has inserted in the BLH3 locus. The insertion of the T-DNA in BLH3 is accompanied by a 37 bp insert of unknown origin and a deletion of 2 bp of BLH3 5′ UTR sequence. Furthermore, analysis of the sequence from the LB indicated that 257 bp of LB T-DNA sequence had been deleted during the insertion into the Arabidopsis genome. The sequence obtained from the product amplified with the RB primer matches partial T-DNA sequences, suggesting that the RB is located internally within a larger T-DNA insertion.

![Diagram](image)

**Figure 6.6 Characterisation of T-DNA insertions in BLH3 and BLH10**

A schematic representation of the position of the T-DNA within the BLH3 and BLH10 genomic sequence. The position of the left border of the T-DNA, relative to the A of the predicted translation start codon, for blh3 is -555 and BLH10 +545. This was determined by DNA sequence analysis of PCR fragments amplified from the T-DNA insertion sites. Black and grey bars represent exons and introns respectively, the grey striped bars represent 5′ UTR. T-DNAs indicated by red triangles not to scale; blh3 T-DNA 7.5 kb; blh10 T-DNA 4.5 kb. **LB**: T-DNA left border primer; **RB**: T-DNA right border primer; **3.1R**: BLH3 gene specific reverse primer; **3.1F**: BLH3 gene specific forward primer; **10.1F**: BLH10 gene specific forward primer; **10.2R**: BLH10 gene specific reverse primer. The expected and actual products obtained after PCR amplification of the genomic region flanking the T-DNA insertions are listed in the tables.

<table>
<thead>
<tr>
<th>Primer combination</th>
<th>Expected product</th>
<th>Actual product</th>
</tr>
</thead>
<tbody>
<tr>
<td>LB, 3.1R</td>
<td>1200 bp</td>
<td>1200 bp</td>
</tr>
<tr>
<td>RB, 3.1R</td>
<td>-</td>
<td>2600 bp</td>
</tr>
<tr>
<td>LB, RB</td>
<td>-</td>
<td>250, 400 bp</td>
</tr>
<tr>
<td>RB, 3.1F</td>
<td>~1200 bp</td>
<td>~2600 bp</td>
</tr>
<tr>
<td>LB, 3.1F</td>
<td>-</td>
<td>~1100 bp</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primer combination</th>
<th>Expected product</th>
<th>Actual product</th>
</tr>
</thead>
<tbody>
<tr>
<td>LB, 10.2R</td>
<td>550 bp</td>
<td>550 bp</td>
</tr>
<tr>
<td>RB, 10.2R</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LB, RB</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RB, 10.1F</td>
<td>~600 bp</td>
<td>~600 bp</td>
</tr>
<tr>
<td>LB, 10.1F</td>
<td>-</td>
<td>~500 bp</td>
</tr>
</tbody>
</table>

Sequencing of the BLH10 insertion with LB primers confirmed that this T-DNA insertion is also flanked by LB sequences on both ends (Figure 6.6b). The LB sequence is located 545 bp downstream from predicted ATG and a 13 bp deletion within the BLH10 coding sequence accompanies the insertion of the T-DNA. Again, sequence from the LB indicates that 246 bp has been deleted during the insertion of the T-DNA into the Arabidopsis
The sequence of the RB amplification product shows high similarity to sequence from *Arabidopsis* chromosome IV. This was considered to be the site of a second T-DNA insertion in the *blh10* mutant plants.

### 6.3.2.3 Confirmation of T-DNA Number in the *BLH3* and *BLH10* Mutants

T-DNA segregation analysis was carried out and the results confirmed that the original *blh3* mutant contained a single T-DNA insertion, whereas the *blh10* line contained two independent insertions (data not shown). Unfortunately, it appeared unlikely that backcrossing the *blh10-1* line had removed the second T-DNA insertion (Chapter 6.2.3.2). Southern analysis was employed to determine if *blh10-1* plants contained more than one T-DNA insertion. The Southern blot was probed with LB T-DNA sequences and the expected T-DNA specific bands were detected (data not shown). Extra bands were also identified which were consistent with additional T-DNA insertions, although whether these corresponded to multiple T-DNA insertions at one loci or independent T-DNA insertions was unknown (data not shown). Therefore, this experiment was inconclusive in determining if *blh10-1* plants contained more than one T-DNA insertion.

To eliminate a possible second T-DNA insertion in the *blh10* plants, the *blh10-1* mutant was again back-crossed to Col wild type plants. The progeny were allowed to self fertilise and the *F₂* progeny were scored for plants segregating with a 3:1 ratio for BASTA resistance. These plants were assumed to contain a single T-DNA insertion which contained the gene conferring BASTA resistance. A single insert plant line, *blh10-2* was bred to homozygosity for the T-DNA. The progeny of this parent line were 100% resistant to the herbicide BASTA. The presence of the T-DNA within the *BLH10* locus was also confirmed by PCR amplification (data not shown). At this point, the majority of phenotypic analyses of the *blh3* and *blh10* mutants had been completed, however phenotypic assays were carried out to ensure that the *blh10-1* and *blh10-2* mutant lines had comparable phenotypes.

### 6.3.2.4 Analysis of Expression in Mutant Plants

In order to test if the presence of the T-DNA insertion had disrupted gene expression, the levels of the *BLH3* and *BLH10* transcripts were examined by Northern analysis. Plants were grown in LD conditions, tissue was harvested at dawn and RNA was extracted. The expression of *BLH3* was greatly reduced in the *blh3* mutant compared to Col wild type (Figure 6.7a). Polyclonal antibodies were raised to BLH3 (Appendix 5) and preliminary
Western blots verified that levels of the BLH3 protein were barely detectable in \textit{blh3} mutants compared to wild type Col plants (Figure A5.1, Appendix 5). No BLH10 expression was detected in \textit{blh10} mutant (Figure 6.7b). This confirmed that the \textit{blh10} mutants were complete knock out plants. The expression levels of BLH3 and BLH10 were not significantly affected in the \textit{blh10} and \textit{blh3} mutants respectively.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6_7.png}
\caption{Analysis of BLH3 and BLH10 mRNA levels in the T-DNA mutants.}
\end{figure}

\textbf{Figure 6.7 Analysis of BLH3 and BLH10 mRNA levels in the T-DNA mutants.}

Expression assays were carried out using total RNA (25 µg) extracted from wild type Col (Col) and mutant plants (Chapter 2.5.2.4). Northern hybridisation was performed as described in Chapter 2.5.6, partial cDNA sequence for BLH3 and BLH10 were used as hybridisation probes and the hybridisation of the 25s/26s rRNA probe was used as a loading control.

\begin{itemize}
  \item[a] BLH3 expression in \textit{blh3} and \textit{blh10} mutants
  \item[b] BLH10 expression in \textit{blh3} and \textit{blh10} mutants
\end{itemize}

\subsection{6.3.3 Identification of BLH3 and BLH10 Point Mutants}

The goal of the \textit{Arabidopsis} Tilling Project (ATP) (Till et al., 2003) is to generate an allelic series of EMS generated mutations throughout the \textit{Arabidopsis} genome. Available point mutants corresponding to the \textit{blh3} and \textit{blh10} genes were identified through the TAIR website (www.arabidopsis.org). The four mutants selected contained G $\rightarrow$ A transitions in conserved regions within the coding sequence. Three \textit{blh3} mutants were identified (\textit{blh3-66}, \textit{blh3-67} and \textit{blh3-73}), all of which contained a single amino acid substitution within the conserved BELL domain. A single \textit{blh10} mutant was identified (\textit{blh10-114}) containing an amino acid substitution within helix II of the homeodomain. Point mutations located within conserved domains or conserved residues of BLH3 and BLH10 are unlikely to be complete knock out mutants, but may affect the protein function if located within an active site within the protein. These mutants may provide information regarding sites important for BLH3 and BLH10 protein function.

The point mutants were identified later during this thesis project than the T-DNA mutants described in Chapter 6.3.2. Although these mutants were generated by EMS mutagenesis, it has been suggested that further backcrossing is not necessary in most cases.
Therefore, a preliminary assessment of the mutant phenotypes of these plants was carried out (Chapter 6.4).

### 6.3.4 The generation of double and triple mutants

#### 6.3.4.1 BLH3Gi-2 double mutants

To examine the genetic interaction between BLH3 and GI, the blh3 T-DNA mutant was cross fertilised with the gi-2 mutant. The progeny of the cross were allowed to self fertilise and screened for the late flowering gi-2 phenotype. Late flowering plants were identified and screened by PCR for the presence of the 8 bp deletion that is the gi-2 mutation (Figure 6.8a).

![Figure 6.8 PCR screen for blh3gi-2 double mutants](image)

**Figure 6.8 PCR screen for blh3gi-2 double mutants**

Genomic DNA was extracted from F2 plants (Chapter 2.5.2.1) and used as a template for PCR (Chapter 2.5.4).

**a** PCR for the gi-2 mutation. The junction spanning gi-2 deletion was amplified using the olig22 and olig37 primers, resulting in a 148 bp product from Col and 140 bp fragment from gi-2. Products were separated on an 8% acrylamide gel (Chapter 2.5.3.3). 1kb; 1 kb plus DNA size marker, only 100 bp and 200 bp fragments visible on this gel; H2O, water control; C, Col wild type; gi, gi-2 mutant; 1-9, products from nine independent F2 plants.

**b** PCR screen for the presence of the blh3 T-DNA. A T-DNA specific primer (LB3) and BLH3 gene specific reverse primer were used to amplify the junction between T-DNA and genomic sequence (1.2 kb). The wild type BLH3 gene was amplified using gene specific primers only (3.2 kb). C, Col wild type control; 1-9, products from nine independent F2 plants. 1kb; 1 kb+ DNA size marker.

Individuals homozygous for the gi-2 mutation were screened by PCR for the presence of the blh3 T-DNA insertion and absence of the BLH3 wild type gene (Figure 6.8b). This was confirmed by selection of the progeny on GM plates containing BASTA and as the progeny were 100% BASTA resistant these plants were deemed homozygous for the T-DNA within blh3. Two independent lines were identified as putative blh3gi-2 double mutants.

#### 6.3.4.2 BLH3Blh10 double mutants
Due to the high degree of conservation between the BLH3 and BLH10 proteins, it seems logical to assume that these proteins may share overlapping functions. To test this assumption, a blh3blh10 double mutant was generated by crossing the blh3 and blh10 mutants. The progeny of the cross were allowed to self fertilise and were screened by PCR for the presence of both the blh3 and blh10 T-DNA insertions. Individuals containing both T-DNA insertions were screened by PCR for the presence of the wild type BLH3 and BLH10 genes (as described in Chapter 6.3.2.1) (data not shown). The plants that did not produce a wild type product were deemed blh3blh10 double mutants. As expected, the progeny of these plants were 100% BASTA resistant when selected on plates containing GM + BASTA.

6.3.4.3 BLH3BLH10Gi-2 AND BLH10Gi-2 MUTANTS

To create a triple mutant, the blh3blh10 and blh3gi-2 double mutants were crossed. The progeny of the cross were allowed to self fertilise and screened for the late flowering gi-2 phenotype. The expected ratio of early flowering to late flowering of 3:1 was not observed in these plants as only 1 in 5 plants were late flowering, presumably because the BLH3, BLH10 and GI genes are all located on chromosome I. Twenty-nine late flowering plants were identified and screened by PCR for the presence of the blh10 T-DNA insertion. Seed was collected from two plants that were recognized as BLH10/blh10 heterozygotes. Subsequently, plants that were homozygous for blh10 were identified by PCR in the resulting F3 seedlings. PCR was also carried out for the blh3 T-DNA insertion to ensure these two plant lines were homozygous for the blh3 mutation (data not shown). Two independent blh3blh10gi-2 mutant lines were generated.

The blh10gi-2 double mutant was generated by crossing the blh10-1 mutant to gi-2. The F2 plants were screened for late flowering and individuals homozygous for the gi-2 mutation were confirmed by PCR (data not shown). Plant lines containing a single blh10 T-DNA insertion were identified by segregation for BASTA resistance at the expected 3:1 ratio and individuals containing the blh10 T-DNA were confirmed by PCR. Selected plants were bred to homozygosity in the F3 generation to create the blh10gi-2 plant line.

6.3.4.4 VERIFYING THE T-DNA NUMBER IN DOUBLE MUTANTS

As the blh10 mutant was thought to contain more than one T-DNA insertion, several attempts were made to determine if the second T-DNA insertion had been inherited by the blh3blh10 mutants. Initially, Southern analysis was employed to determine the T-DNA
number in the double mutants. However, the probable insertion of more than one T-DNA within the blh3 and blh10 loci resulted in extra bands hybridising to the T-DNA probe (data not shown). It was difficult to verify whether these bands corresponded to multiple T-DNA insertions at one loci or independent T-DNA insertions.

As an alternative, a genome walking protocol for the isolation of DNA flanking the T-DNA borders was used, as modified from Spertini et al. (1999) by M. Yoon (this laboratory). Briefly, genomic DNA was subject to restriction enzyme digestion with enzymes known to cut within the T-DNA borders. Small oligo-adaptors were ligated to the digested DNA and used as a template for amplification of the T-DNA and flanking sequences. The background amplification during the majority of PCR reactions was difficult to overcome, as a T-DNA specific nested primer was not available for the second round of amplification. Only amplification of TaqI digested DNA was successful. As predicted, two T-DNA specific bands were amplified from the blh10-1 mutant, indicating this mutant had more than one T-DNA insertion. A third, less intense band was detected, possibly due to the insertion of more than one T-DNA at the blh10 locus (Figure 6.9).

![Figure 6.9 Isolation of T-DNA flanking genomic DNA](image)

The blh3blh10gi-2 triple mutant contained only a single blh10 specific band and a single blh3 specific band, verifying that the second T-DNA insertion in blh10 plants had not been inherited in this plant line. The results from the blh3blh10 plants were not able to be interpreted as samples constantly degraded (Figure 6.9). It is not known if this plant line contains both T-DNA insertions associated with the blh10 mutation.
6.4 ANALYSIS OF MUTANT PHENOTYPES

The plant lines generated in Chapters 6.3.2 and 6.3.4 were analysed for mutant phenotypes. Mutant phenotypes have not yet been described for the majority of the BELL genes in *Arabidopsis*, however the *bell1* and *pny* mutants have been well characterised. A mutation within the *BELL1* gene results in abnormal ovule development and female sterility (Ray *et al.*, 1994). In contrast, *pny* (also known as *van* or *bhr*) mutants exhibit a striking dwarf phenotype, increased leaves and branching and altered internode patterning (Bhatt *et al.*, 2003; Byrne *et al.*, 2003; Smith and Hake, 2003; Bao *et al.*, 2004). Comparison to *blh3* and *blh10* plants grown in standard conditions verified that mutations within the *BLH3* and *BLH10* genes did not affect these areas of plants development. In general, *blh3* and *blh10* plants resembled wild type plants.

6.4.1 FLOWERERING TIME ASSAYS

Plants containing a *gi* mutation are late flowering compared to wild type, particularly under long day conditions. In addition, the BELL gene double mutant *pnypnf* is unable to flower, even when grown in inductive conditions (Smith *et al.*, 2004). Hence the flowering time of the *blh3* and *blh10* mutants was measured. The flowering time of the original *blh3* and *blh10* mutants were analysed in LD and SD conditions. These results suggested that these plant lines flowered slightly later than wild type Col plants in LD and SD conditions (Figure 6.10a and b). In comparison to the *gi-2* mutant, which flowers at over 30 leaves in LD, the *blh3* and *blh10* mutants were not especially late flowering (data not shown).

![Figure 6.10 Preliminary flowering time assays of the *blh3* and *blh10* mutants.](image)

- **a** Flowering time of *blh3* and *blh10* mutants in LD. Values are mean ± SEM, n ≥ 7
- **b** Flowering time of *blh3* and *blh10* mutants in SD. Values are mean ± SEM, n ≥ 7
The \textit{blh3-66}, \textit{blh3-67}, \textit{blh3-73} and \textit{blh10-114} point mutants (Chapter 6.3.3) were grown in standard LD conditions and preliminary assessment of flowering time suggested that these plants flower at the equivalent time as wild type (data not shown).

### 6.4.1.1 Flowering Time of Back-crossed \textit{BLH3} and \textit{BLH10} Lines

The \textit{blh3} and \textit{blh10} lines were back-crossed to wild type Col to ensure the late flowering phenotype was linked to the T-DNA insertions within the \textit{BLH3} and \textit{BLH10} genes. If the late flowering phenotype is a result of a T-DNA insertion within the \textit{BLH3} and \textit{BLH10} genes, it is predicted that the flowering time of the F\textsubscript{2} progeny will segregate for the late flowering phenotype, with 1 in 4 progeny homozygous for the T-DNA and hence late flowering. This was the case when \textit{blh3} and \textit{blh10} where backcrossed to Col (Figure 6.11), where 1 in 4 plants flowered ~ 2 leaves later than the average Col flowering time of 10.5 leaves.

![Segregation of flowering time in the F\textsubscript{2} generation](image)

**Figure 6.11** Segregation of flowering time in the F\textsubscript{2} generation

The progeny of the backcross of \textit{blh3} and \textit{blh10} mutants to wild type were allowed to self fertilise. The F\textsubscript{2} progeny were grown in LD conditions and the flowering time was measured as leaf number at flowering.

Following the back-cross to Col, two homozygote lines representing \textit{blh3} (\textit{blh3-1}, \textit{blh3-1}) and \textit{blh10} (\textit{blh10-1}, \textit{blh10-34}) were selected and assayed for flowering time. The \textit{blh3} and \textit{blh10} mutants were slightly later than wild type when grown in LD conditions, flowering with ~2 more leaves than the equivalent aged wild type plants (Figure 6.12a). These LD flowering assays were repeated on two separate occasions and similar results were
obtained (data not shown). The difference in flowering time between Col and the two mutants was more pronounced when the plants were grown in SD conditions. The \textit{blh3} and \textit{blh10} lines flowered at approximately 10-15 leaves more than Col (Figure 6.12b). These SD flowering time assays were also repeated and similar results were obtained.

![Figure 6.12 Flowering time of independent \textit{blh3} and \textit{blh10} mutant lines.](image)

### Table 6.3 Flowering time data for \textit{blh3} and \textit{blh10} lines in SD.

<table>
<thead>
<tr>
<th>Line</th>
<th>SD conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rosette</td>
</tr>
<tr>
<td>Col</td>
<td>45.5 ± 1.6</td>
</tr>
<tr>
<td>gi-2</td>
<td>59.3 ± 5.3</td>
</tr>
<tr>
<td>\textit{blh3}-1</td>
<td>54.6 ± 3.2</td>
</tr>
<tr>
<td>\textit{blh3}-3</td>
<td>55.6 ± 2.8</td>
</tr>
<tr>
<td>\textit{blh10}-1</td>
<td>52.7 ± 2.6</td>
</tr>
<tr>
<td>\textit{blh10}-34</td>
<td>53.9 ± 1.7</td>
</tr>
</tbody>
</table>

Plants were grown in SD and the leaf number at flowering counted. Values are mean ± SEM, n ≥ 8.
Flowering time was measured as the total number of leaves at flowering; this was the sum of rosette and cauline leaves. The late flowering phenotype of the \textit{blh3} and \textit{blh10} mutants in SD is due to the production of more rosette and cauline leaves compared to Col wild type plants. In contrast, the \textit{gi-2} mutants did not produce more cauline leaves than Col and the late flowering phenotype observed is due to production of more rosette leaves only (Table 6.3).

### 6.4.1.2 Plastochron Index

In this thesis work, the time to flowering was generally calculated as the total number of leaves produced at flowering. The production of more leaves than wild type was interpreted as a late flowering phenotype. This conclusion may be misleading if plant lines produced leaves at a greater rate than wild type. To establish if the late flowering phenotype of the \textit{blh3} and \textit{blh10} plants was simply due to an increased rate of production of rosette leaves compared to wild type, a plastochron assay was performed. The period between the initiation of successive leaves is termed a plastochron (Erickson and Michelini, 1957). Leaf production and hence growth is expected to be constant rather than exponential with respect to time. Wild type, \textit{blh3} and \textit{blh10} mutants were grown in SD conditions and the rate of leaf production was counted weekly from 10 weeks after planting until plants flowered. This experiment established that all three plants lines produced rosette leaves at a reasonably constant rate over the course of the experiment (Figure 6.13).

![Plastochron Index of \textit{blh3} and \textit{blh10} mutants grown in SD.](181)

Plants were grown in SD conditions and the number of visible rosette leaves was counted weekly from 70 d after planting until flowering. Values are mean ± SEM, \( n \geq 7 \).
The results of this experiment verified that the late flowering phenotype of the \textit{blh3} and \textit{blh10} mutants is not due to an increased rate of leaf production compared to wild type. This is supported by the fact that the \textit{blh3} and \textit{blh10} mutants flowered up to 4 weeks later than Col when days to flowering was measured (as opposed to leaf number at flowering) (data not shown).

\subsection*{6.4.1.3 Flowering time in two SD light conditions}

Due to the discovery of an interesting flowering time phenotype in SD, the \textit{blh3} and \textit{blh10} flowering time assays were repeated. For this experiment, the plants were grown in full light conditions under new fluorescent lighting. In contrast to the previous experiments, when grown in these short day conditions the \textit{blh3} and \textit{blh10} mutants flowered at approximately the same leaf number as wild type (Figure 6.14, 150 \( \mu \text{mol.m}^{-2}.s^{-1} \)). When the experiment was carried out in lower light conditions (~80 \( \mu \text{mol.m}^{-2}.s^{-1} \)), the \textit{blh3} and \textit{blh10} plants flowered at approximately 10 leaves later than Col (Figure 6.14, 80 \( \mu \text{mol.m}^{-2}.s^{-1} \)).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure6_14.png}
\caption{Flowering time of plants grown in two SD light conditions}
\footnotesize{Comparison of the flowering time of plants grown in SD full light (~150 \( \mu \text{mol.m}^{-2}.s^{-1} \)) and SD low light (~80 \( \mu \text{mol.m}^{-2}.s^{-1} \)) conditions. Values are mean ± SEM, \( n \geq 10 \).}
\end{figure}

\subsection*{6.4.1.4 Flowering time of double and triple mutants}

The flowering time of the \textit{blh3blh10} double mutant was assessed. Unlike the \textit{blh3} and \textit{blh10} single mutants, these double mutants flowered at the same leaf number as wild type in both LD and SD conditions (Figure 6.15a). These results demonstrate that \textit{BLH3} and \textit{BLH10} do not function in a dose dependent manner, as there is no additive effect on flowering in the double mutant. It has been established that the BLH3 and BLH10 proteins
interact with each other (Chapter 4) and other BELL proteins (Hackbusch et al., 2005); therefore it is possible that the removal of both interacting partners removes a pathway that is subsequently bypassed or compensated for, perhaps by other BELL genes.

![Figure 6.15 Flowering time of double mutants.](image)

**Figure 6.15 Flowering time of double mutants.**

Plants were grown in LD and SD and flowering time measured. Values are mean ± SEM, n ≥ 8.

- **a** Flowering time of blh3blh10 double mutants.
- **b** Flowering time of blh3gi-2 double mutants.

The flowering time of the blh3gi-2 mutant was the equivalent to the gi-2 mutant alone in LD and SD conditions (Figure 6.15b). As there is no additive effect of the blh3 mutation on flowering time, it is likely that BLH3 and GI function in the same pathway. The blh3blh10gi-2 and blh10gi-2 mutants were also late flowering in LD, flowering with approximately the same leaf number as the gi-2 mutant (data not shown). As observed with the blh3gi-2 mutant, there was no additive effect of the blh3 and blh10 mutation on the late flowering of gi-2 mutant. Again, these results are suggestive that BLH3, BLH10 and gi-2 function in the same genetic pathway.
An interesting effect of the \textit{blh3} mutation on flowering in \textit{GI/gi-2} heterozygotes was observed (Figure 6.16a). The F\textsubscript{3} progeny of the cross between \textit{blh3} and \textit{gi-2} plants were grown in LD conditions. These plants were segregating for the extreme late flowering \textit{gi-2} phenotype (~40-50 leaves), compared to wild type which flowered at 12-14 leaves. The flowering time of a population of moderately late flowering plants was unable to be explained by the presence of the \textit{blh3} mutation alone, as \textit{blh3} plants flower at ~2 leaves more than Col in these LD conditions. The plants with an intermediate flowering time phenotype flowered 4-12 leaves later than wild type (Figure 6.16a).

**Figure 6.16** Segregation of flowering time in \textit{blh3 x gi-2} progeny

The F\textsubscript{3} progeny of \textit{blh3 x gi-2} mutants were grown in LD and the flowering time was assessed.

\textbf{a} Graph of flowering time of all \textit{blh3 x gi-2} progeny

\textbf{b} Graph of flowering time of selected F\textsubscript{3} progeny. The genotype of these plants, as indicated by the colour coding, was confirmed by PCR.
Twenty-six plants representing the range of flowering time phenotypes observed were selected and the genotype of these plants was determined by PCR (data not shown). Wild type Col and gi-2 mutant plants were included as controls. As expected, the latest flowering plants were \textit{blh3}gi-2 mutants and plants flowering slightly later than Col were \textit{blh3} mutants. The flowering plants with an intermediate flowering time were heterozygous \textit{GI}/gi-2 in a \textit{blh3} mutant background (Figure 6.16b).

### 6.4.2 Light Response Assays

Previously, Araki and Komeda (1993b) reported that gi-2 mutants exhibited slightly elongated hypocotyls when grown under constant light conditions. In a more recent study, Huq \textit{et al.} (2000) established that gi mutants also have elongated hypocotyls when grown in constant red light (Rc). Therefore, the effect of various light conditions on hypocotyl elongation in \textit{blh3} and \textit{blh10} mutants was examined.

Initially, populations of seed were subject to germination tests to ensure that the germination rates of the mutants plant lines were equivalent to that of Col. The seeds were prepared and subject to light treatment to induce germination as described in Chapter 2.4.1.6. Seedlings were assayed for radical formation 1 d and 2 d later. Although germination rates were not ideal and there was a considerable variation in germination within a population, on average all plant lines germinated at the same rate (data not shown). In consequence, any differences observed in hypocotyl lengths between plant lines could not be attributed to a rapid or delayed germination rate within that line.

#### 6.4.2.1 Analysis of Hypocotyl Elongation in SD

Due to the flowering phenotype observed in short days, hypocotyl elongation in \textit{blh3} and \textit{blh10} seedlings were assessed in SD. Seedlings were grown on MS plates in low light SD conditions and compared to seedlings grown in constant light (LL) conditions. In LL conditions, hypocotyl elongation was inhibited by light and the hypocotyl lengths of all seedlings were particularly short. As indicated in previous work by Araki and Komeda (1993b), the gi-2 hypocotyls were slightly longer than Col in LL (Figure 6.17a). In SD conditions, gi-2 mutants exhibited hypocotyls almost twice the length of the wild type Col seedlings. Although there was no obvious effect of the \textit{blh3} and \textit{blh10} mutations on
hypocotyl length in SD, there was an additive effect of \textit{blh3} in the \textit{blh3gi-2} double mutant (Figure 6.17a).

The SD hypocotyl assay was repeated in controlled light conditions. Seedlings were grown in high light conditions (Figure 6.17b; SD, 100 \(\mu\text{mol.m}^{-2}.\text{s}^{-2}\)) and compared to seedlings grown in low light conditions (Figure 6.17b; 1/2L SD, 40 \(\mu\text{mol.m}^{-2}.\text{s}^{-2}\)). Interestingly, the \textit{blh3} and \textit{blh10} mutants exhibited slightly longer hypocotyls than wild type only in the low light SD conditions (Figure 6.17b). The \textit{gi-2} and \textit{blh3gi-2} mutants had elongated hypocotyls compared to wild type, with an additive effect of \textit{blh3} in the \textit{gi-2} background in the low light conditions only. This is indicative of a role for \textit{BLH3} in low light conditions that is at least partially independent of \textit{GI}.

**Figure 6.17**  
Hypocotyl length assays in SD

\textbf{a} Seeds were germinated in white light for 2 h and seedlings were grown on growth media under LL for 7 d or SD for 10 d. Graph of mean hypocotyl length of seedlings, values are mean ± SEM, \(n > 15\).

\textbf{b} Seeds were germinated in white light for 2 h and seedlings were grown on growth media under high light SD (100 \(\mu\text{mol.m}^{-2}.\text{s}^{-2}\)) or 1/2L SD (40 \(\mu\text{mol.m}^{-2}.\text{s}^{-2}\)) for 10 d. Graph of mean hypocotyl length of seedlings, values are mean ± SEM, \(n > 15\).
The *gi-2* mutant also had elongated petioles when grown in SD, compared to wild type Col. The *blh3* and *blh10* mutants did not resemble *gi-2* plants and did not have longer petioles than wild type plants in these conditions (data not shown).

### 6.4.2.2 Analysis of Hypocotyl Elongation in Red Light

To test if the *blh3* and *blh10* mutations affect hypocotyl elongation in red light (*Rc*), seedlings were grown in *Rc* and their hypocotyl length measured. Unfortunately, difficulties were encountered in carrying out these hypocotyl assays. Due to the poor germination rates in tissue culture, it was difficult to obtain a reasonable number of seedlings to assay for hypocotyl length. The length of the assay period was increased from 4 d to 10 d in an attempt to increase germination rates and to compensate for any variation in the rates of germination within a population. Preliminary hypocotyl assays demonstrated that like *gi-2* mutants, the *blh3* and *blh10* mutants had longer hypocotyls than wild type seedlings (Figure 6.18). Seedlings were assayed on two separate occasions and the averages were graphed. These results demonstrate that despite poor germination and variation within populations of seed, these assays are relatively constant over two independent experiments (Figure 6.18).

![Mean hypocotyl length (mm)](image)

**Figure 6.18  Preliminary hypocotyl elongation assays in red light**

Graph of mean hypocotyl length of seedlings grown in *Rc*. Seeds were germinated in white light for 2 h and seedlings were grown on growth media under *Rc* (30 μmol.m$^{-2}$.s$^{-1}$) for 10 d. Black and grey bars are comparisons of two independent repeats. Values are mean ± SEM, n >5.

Despite variability these hypocotyl assays appeared relatively robust, so these assays were used to perform large scale hypocotyl elongation experiments. It was found that new seed stocks (~2-4 weeks old) germinated best in tissue culture, consequently only fresh seed
was used in further assays. The hypocotyl length assays in red light were performed several times with similar results and a representative experiment is shown (Figure 6.19a and b). The *gi-2*, *blh3* and *blh10* mutants exhibited reduced inhibition of hypocotyl elongation when grown in red light. The long hypocotyl phenotype was not additive in *blh3*-*gi-2*, *blh10*-*gi-2* or *blh3*-*blh10* double mutants suggesting that the *BLH3*, *BLH10* and *GI* genes act in the same pathway to regulate red light signalling. However, the *blh3*-*blh10*-*gi-2* triple mutant was more severely impaired in red light signalling, as this mutant exhibited longer hypocotyls than the single or double mutants alone (Figure 6.19b). This is suggestive that *BLH3* and *BLH10* together function at least partially independently of *GI* in red light.

![Figure 6.19](image)

**Figure 6.19** Hypocotyl elongation assays in red and low light conditions

- **a** Comparison of hypocotyl lengths of mutants grown in constant red light (Rc) and constant darkness (DD). Seeds were germinated in white light for 2 h and seedlings were grown on growth media under Rc (30 µmol.m\(^{-2}\).s\(^{-1}\)) or DD for 4 days.
- **b** Graph of mean hypocotyl length of seedlings grown in Rc, low light (0.1 µmol.m\(^{-2}\).s\(^{-1}\)) and DD. Values are mean ± 95% confidence intervals, n = 15.

The *gi-2*, *blh3* and *blh10* mutant seedlings exhibited reduced inhibition of hypocotyl elongation under low light conditions. Comparison to seedlings grown in constant darkness (DD) show that only Col and *blh3*-*blh10* seedlings responded to the low light conditions; all
other mutants have elongated hypocotyls of a length equivalent to dark grown control seedlings (Figure 6.19b; ~0.1 \( \mu \text{mol.m}^{-2}.\text{s}^{-1} \) white light). This suggests that the \( \text{gi-2, blh3 and blh10} \) plants are also defective in the response to low light.

Control seedlings were grown in DD to detect any non-light related differences in hypocotyl length. The hypocotyl length of seedlings grown in experimental light conditions were compared to dark grown seedlings to calculate the relative hypocotyl length. These calculations verified that the difference in hypocotyl length seen in experimental light conditions was chiefly a light mediated effect. The \( \text{gi-2, blh3 and blh10} \) mutants exhibited increased relative hypocotyl elongation compared to wild type in both red light and low light conditions (Figure 6.20). The most severe red light phenotype was detected in \( \text{blh3blh10gi-2} \) mutant. Hypocotyl elongation of triple mutant seedlings was approximately the same irrespective of growth in \( \text{Rc and DD} \) conditions (Figure 6.20), which was suggestive that these seedlings are defective in red light signalling.

Due to the long hypocotyl phenotype observed in red and low light conditions, the wavelength specificity of this phenotype was examined. The \( \text{blh3, blh10 and gi-2} \) seedlings were grown under far red and blue light for four days. There was no difference in hypocotyl elongation between mutant and \( \text{Col wild type} \) seedlings (data not shown). These results imply that the long hypocotyl phenotype observed for \( \text{gi-2, blh3 and blh10} \) is specific to red and low light conditions.

Impaired light response can also affect other aspects of seedling development, such as cotyledon expansion and opening (Fankhauser and Casal, 2004). Examination of \( \text{blh3 and} \)
blh10 seedlings grown in red light revealed that apart from the hypocotyl phenotype, these seedlings resembled wild type plants. No obvious affects on cotyledon development was observed (Figure 6.19a).

### 6.4.2.3 An alternative red light hypocotyl assay

An alternative assay was attempted to further determine the effects of light on hypocotyl elongation in the various mutants generated in this work. This assay was adapted from that described by Fankhauser and Casal (2004) and involves subjecting seed to a short light treatment of 15 min followed by 1 d in DD conditions prior to 4 d in experimental light conditions. Seedlings grown in this manner were compared to seedlings grown in the original conditions (Chapter 6.4.2.2) in a preliminary hypocotyl length assay. The results revealed that the difference in hypocotyl elongation seen previously between gi-2, blh3 and Col is no longer detected in seedlings grown under the new regime (data not shown). The experiment was repeated with the appropriate dark grown controls and a similar result was obtained (Figure 6.21). No significant difference in hypocotyl length was seen in any of the plant lines investigated.

![Graph of mean hypocotyl length of seedlings grown in Rc and DD. Seeds were germinated in white light for 15 min, followed by 1 d DD and 4 d Rc or 4 d DD. Values are mean ± 95% CI, n ≥ 20.](image)

**Figure 6.21 Hypocotyl elongation in alternative red light conditions**

6.4.2.4 Hypocotyl elongation in plants containing point mutations

Three blh3 point mutants blh3-66, blh3-67 and blh3-73 were identified that contain a 1 bp change (G to A) that resulted in single amino acid changes within the conserved BELL domain of the predicted protein (Chapter 6.3.3). Preliminary experiments in standard LD conditions suggested that blh3-66 mutants had elongated hypocotyls. In comparison, the
*blh3*-67 and *blh3*-73 mutants appeared equivalent to wild type in LD conditions. A point mutation within *BLH10* was also identified (Chapter 6.3.3). The *blh10*-114 mutant contains a 1 bp change (G to A) within helix II of the homeodomain. This mutant did not have a long hypocotyl phenotype in LD. Perhaps reassuringly, the lack of a hypocotyl phenotype in other point mutant plants indicates that the elongated hypocotyls exhibited by the *blh3*-66 mutant are not simply due to the genetic background of the plant lines.

A standard hypocotyl assay was carried out in Rc light conditions. This result demonstrated that *blh3*-66 had particularly long hypocotyls in red light, almost twice the length of the *gi*-2 mutant (Figure 6.22). Although the *blh3*-66 mutation is present in the Col background, the parent plant also contains the *erecta* (*er*) mutation. It is possible that the long hypocotyl phenotype observed in the *blh3*-66 plants is due to an interaction with the *er* mutation, although this seems unlikely as the *er* mutation is associated with short hypocotyls. The *blh3*-66 plant line must be back-crossed to wild type Col and the hypocotyl length of this line reassessed before any conclusions can be made.

![Figure 6.22](image)

**Figure 6.22** The *blh3*-66 mutant has elongated hypocotyls in Rc
Seeds were germinated in white light for 2 h and seedlings were grown on growth media under Rc for 4 d. Graph of mean hypocotyl length of seedlings, values are mean ± SEM, n >8.

### 6.4.3 OVER-EXPRESSION OF GI IN THE BLH3 MUTANT BACKGROUND

The over-expression of *GI* in transgenic plants results in early flowering in both LD and SD conditions (Milich, 2001; Mizoguchi *et al.*, 2005). To analyse the effect of *GI* over-expression in the *blh3* mutant background, transgenic plants containing a 35S:*GI* expression construct were generated. The expression construct for the over-expression of *GI* was obtained from K. Lee (p:35S:*GI-KM*) (John Innes Research Centre, Norwich). The p:35S:*GI-KM* vector contains a 35S:*GI* expression cassette inserted within the T-DNA region in the pGREEN binary vector.
6.4.3.1 **GENERATION OF 35S:GI TRANSGENIC PLANT LINES**

The 35S:GI-KM construct was introduced in *Arabidopsis gi-2* and *blh3gi-2* plants via *Agrobacterium*-mediated transformation, as described previously (Chapter 5.3.2.2). T₀ seed was harvested and sown onto tissue culture plates and kanamycin resistant transformants were selected (Chapter 2.4.4.1). Very low transformation efficiency and survival rate of plants on transfer from tissue culture to soil resulted in only a single surviving 35S:GI/*gi-2* line and fifteen 35S:GI/*blh3gi-2* lines. This was attributed to unsuitably wet greenhouse conditions and the resulting poor condition of the plants at the time of the transformation.

6.4.3.2 **ANALYSIS OF FLOWERING TIME IN 35S:GI TRANSGENIC PLANTS**

The flowering time of the 35S:GI/*blh3gi-2* lines generated in Chapter 6.4.3.1 was examined. These lines were heterozygous for the 35S:GI insertion, so were predicted to be segregating for the T-DNA and hence early flowering. Eight selected 35S:GI/*blh3gi-2* plant lines were grown in standard LD conditions. Six of the eight lines were segregating for early flowering at the ratio of 3:1, which is expected for plants with a single T-DNA insertion. The presence of the *GI* expression construct in early flowering plant lines was confirmed by PCR amplification (data not shown). These plants flowered slightly early at six to eight leaves, compared to Col plants which flowered at approximately 8 leaves. Late-flowering *gi-2* and *blh3gi-2* mutants flowered at more than 30 leaves (data not shown). These results verified that the 35S:GI insertion was sufficient to induce early flowering in a *blh3gi-2* mutant background.

The single surviving 35S:GI/*gi-2* line was also early flowering, though unexpectedly this plant line was not segregating for flowering time. Analysis of these plants by PCR verified that this plant line did not contain the *gi-2* mutation or the *blh3* T-DNA insertion, suggesting that a contaminant wild type plant had been transformed. Plants were screened by PCR to confirm the presence of the *GI* cDNA insertion (data not shown). In consequence, this plant line was considered to be 35S:GI/Col.

The expression of *GI* in the 35S:GI/Col line and a selected early flowering 35S:GI/*blh3gi-2* line was tested by Northern analysis. Plants were grown in standard LD conditions and aerial tissue was harvested at ~ 1 h after dawn, a time when endogenous *GI* levels are low. The results of Northern hybridisation confirmed that *GI* was expressed at levels
significantly higher than wild type in 35S:GI/Col and 35S:GI/blh3gi-2 transgenic plants (data not shown). The results obtained from these experiments signify that BLH3 is not required for GI to promote flowering in plants over-expressing GI.

6.4.3.3 **Hypocotyl elongation in 35S:GI lines**

The hypocotyl length of the 35S:GI/Col and 35S:GI/blh3gi-2 plant lines was also assessed as part of this work. In preliminary hypocotyl assays, seedlings containing the 35S:GI construct had shorter hypocotyls than wild type when grown in DD control conditions (data not shown). As these seeds were slow to germinate, this response is indicative of the slow germination time of these plant lines compared to Col plants. When grown in Re conditions, 35S:GI seedlings had shorter hypocotyls than gi-2 and blh3 mutants, though not as short as wild type (data not shown). The results obtained here with 35S:GI plants are difficult to compare to Col, as these plant lines germinated at different rates.

Since the time these experiments were carried out Mizoguchi *et al.* (2005) have established that the over-expression of GI results in a short hypocotyl phenotype, indicating these plants are hypersensitive to red light.

6.4.4 **LHY expression in the BLH3 and BLH10 mutants**

The expression of BLH3 is not directly clock-controlled nor is it affected by the clock mutants elf3 and lhy (Chapter 5). In comparison, gi mutations strongly reduce the expression of the circadian clock genes *LHY* and *CCA1* in LD (Fowler *et al.*, 1999). The expression of BLH3 is reduced in gi and cca1 mutants at dawn in plants grown in daily light/dark cycles (Chapter 5.2.4). The reduction of BLH3 or BLH10 expression in the blh3 and blh10 mutants may have a feedback affect on clock controlled genes. To test this hypothesis, the expression of clock component LHY was tested in Col wild type, blh3 and blh10 mutant plants.

Plants were grown in standard LD and SD conditions and aerial tissue was harvested from plants at the 4-6 leaf stage. The results of a Northern hybridisation assay indicated that the reduction of BLH3 and BLH10 did not significantly affect the *LHY* transcription pattern. The *LHY* transcript continues to cycle with a peak at dawn and a trough later during the day (Figure 6.23). A slight increase in the peak of *LHY* expression is apparent in blh3 and blh10 mutants compared to Col at dawn (Figure 6.23). However, as the experiment was not
repeated, the significance of this result is unknown. It is apparent that the effect of blh3 and blh10 on LHY expression is not equivalent to the effect of the gi mutation.

![Figure 6.23](image)

**Figure 6.23 Analysis of LHY mRNA levels in the blh3 and blh10 mutants.**

Wild type (Col) and mutant plants grown in standard LD and SD conditions and aerial tissue was harvested at the time points shown. Northern hybridisation experiments were carried out using total RNA (25 µg) (Chapter 2.5.6); partial cDNA sequence for LHY was used as a hybridisation probe and the hybridisation of the 25S/26S rRNA probe was used as a loading control.

### 6.5 DISCUSSION

**6.5.1 OVER-EXPRESSION OF BLH3 IN TRANSGENIC PLANTS**

Transgenic plants containing a 35S:BLH3 expression construct were created to analyse the effects of over-expression of BLH3 on plant development. Despite containing the correct insertion, these plants failed to over express BLH3. One possible explanation is that there was a fault within the 35S:BLH3 construct. Amplification of the 35S:BLH3 junction indicated that the 35S and BLH3 sequences had been inserted into plants, but sequences outside this region may have been absent in plants or contained mutations that affected expression. On the other hand, it is possible that sequences outside of the coding region were required for the correct control of BLH3 expression. Only BLH3 cDNA sequence was cloned into the 35S:BLH3 construct and it is feasible that upstream sequences, such as those within the leader intron, are important for regulating BLH3 expression in plants.

Similarly, Cole *et al.* (2006) reported that transgenic plants containing a 35S driven BLH3 construct resembled wild type plants, although some independent transgenic lines did flower earlier than wild type. In contrast, plants containing the 35S:BLH3 construct in an STM:GR background flowered rapidly, but only after the release of interaction partner STM by dexamethasone induction (Cole *et al.*, 2006). It is reasonable to suspect that the
expression level of \textit{BLH3} is not as important as the presence and/or absence of interaction partners in the flowering function of \textit{BLH3}.

6.5.2 \textbf{Over-expression of \textit{BLH10-R} in transgenic plants}

Why were no surviving transformants obtained when the 35S:BLH10-R construct was introduced into plants? The 35S:BLH3 and 35S:BLH10-R constructs were generated from the same pART7 and pART27 vectors, therefore it is unlikely that the fault lies in the 35S:BLH10-R vector. Furthermore, the structure of the 35S:BLH10-R vector used to transform plants was verified before and after transformation into \textit{Agrobacterium}. Even though transformation efficiencies for these experiments were not as high as expected, there was no apparent fault in the transformation protocol as transgenic progeny containing the 35S:BLH3 construct were produced. Neither was there a fault in the screening procedure for transgenic progeny, again because the identical procedure resulted in the identification of plants containing the 35S:BLH3 insertion.

As no 35S:BLH10-R transformants survived, it is tempting to speculate that the over-expression of \textit{BLH10-R} may be detrimental to plant development. The protein encoded by \textit{BLH10-R} is truncated and does not contain a homeodomain DNA binding region, thus is unlikely to act as a functional transcription factor. \textit{BLH10-R} encodes the BELL domain, so may still be involved in protein::protein interactions and consequently could have a dominant negative role in plants. Placing \textit{BLH10-R} under the control of a chemically inducible promoter may be more appropriate for determining the role of this alternatively spliced \textit{BLH10} transcript in plants.

6.5.3 \textbf{What is the role of \textit{BLH3} and \textit{BLH10} in flowering?}

The \textit{BLH3} and \textit{BLH10} genes are not classical flowering time genes as they do not exhibit a pronounced late flowering phenotype when mutated. Furthermore, \textit{BLH3} and \textit{BLH10} are not directly regulated by the circadian clock (Chapter 5), nor do mutations in these genes affect the expression of the clock gene \textit{LHY}. These results clearly indicate that although \textit{BLH3} and \textit{BLH10} interact with \textit{GI}, they do not act with \textit{GI} to control the circadian clock nor photoperiodic flowering.

What is the role of \textit{BLH3} and \textit{BLH10} in plants? Interestingly, the \textit{blh3} and \textit{blh10} mutants flower later than wild type in SD, a phenotype that was more pronounced in low light
conditions. This late flowering phenotype in SD was shared by the gi-2 mutant, indicating that GI does not solely function to promote flowering in response to LD photoperiods. Evidence presented by Paltiel et al. (2006) suggests a function for GI in SD that is dependent on temperature, supporting the assumption that GI does not function exclusively in response to day length to control flowering. All plant lines tested, including Col, flowered earlier in the low light SD conditions than in full light SD conditions; this response to light quality could be attributed to a shade avoidance response (Chapter 1.3.1.1). In low light SD conditions, the gi-2, blh3 and blh10 mutants did not have a flowering response as rapid as wild type. It is possible that GI, BLH3 and BLH10 act in response to light quality to affect flowering, particularly in the absence of inductive day lengths.

Intriguingly, blh3 GI/gi-2 plants have a moderately late flowering phenotype in LD which is later than blh3 mutants, though not as late as gi-2 mutants (Chapter 6.4.1). GI is a dominant gene and a single copy is sufficient for early flowering. In consequence, GI/gi heterozygotes normally exhibit flowering times equivalent to wild type in LD conditions. As low levels of BLH3 are still detectable in blh3 mutants, the late flowering phenotype may be due to a reduction in the total levels of BLH3 and GI in blh3 GI/gi-2 heterozygotes. This flowering phenotype is suggestive of an interaction between BLH3 and GI in plants and implies that levels of GI are important in promoting flowering in a blh3 background. Supporting this hypothesis, very high levels of GI in 35S:GI transgenic plants are sufficient to induce early flowering, irrespective of the presence of the blh3 mutation.

Introduction of the 35S:BLH3 construct into gi-2 mutants did not affect the late flowering gi phenotype. However, this experiment is inconclusive in determining the role of BLH3 in flowering, as BLH3 was not over-expressed by the 35S construct. The over expression of GI in a blh3 background suggests that BLH3 is not required for GI to promote flowering. However, as BLH3 belongs to a gene family connected by a network of interactions, it is possible that other BELL genes compensate for the loss of BLH3. In retrospect, it would be informative to determine if the rapid flowering induced by 35S:BLH3 and STM:GR as described by Cole et al. (2006) would be sufficient to complement the late flowering gi mutation. This may elucidate if the BLH3:STM interaction functions epistatically to gi to control flowering time.
The blh3 and blh10 mutants flower later than wild type in SD conditions (Chapter 6.4.1), yet their role in flowering is uncertain. Other BELL genes have been linked to flowering in recent studies. The over-expression of BLH3 and PNY/BLH9 produce plants that flower slightly earlier than wild type. However, PNY/BLH9 over-expression results in a greater rate of leaf production than both wild type and BLH3 over-expressing plants. This is indicative of a subtle difference in the role of these two BELL genes in flowering (Cole et al., 2006). Yet BLH3 and BLH9 interact (Hackbusch et al., 2005), therefore it is possible that these proteins have some overlapping function in plants. The shared function of BELL genes in plants could be testing using the double mutant pny pnf, which is unable to flower, even in inductive conditions (Smith et al., 2004). The effect of increased expression of BLH3 or BLH10 in the pny pnf mutant background could determine if an increase in these BELL genes is able to compensate, partially or fully, for the loss of PNY and PNF. This would contribute to the understanding of how BLH3 and BLH10 act and interact with other BELL genes to affect flowering.

Both the BLH3 and BLH10 transcripts were low in gi-2 and cca1 mutants at a single time point during the day. Furthermore, this phenotype was evident only in plants grown in driven light/dark cycles (Chapter 5). A further connection with clock function is suggested by the results presented in Chapter 6.4.4, which show that LHY expression is slightly increased in blh3 and blh10 mutants at 8 h after dawn, although the cyclical expression pattern of LHY appears unchanged. It is evident that these expression phenotypes are subtle and for this reason the significance, if any, is unclear. To put this into perspective, take the expression of GI as an example. The difference between GI expression in LD and SD is apparently minor, yet only the LD expression pattern is associated with floral induction. Subtle changes in expression may be all that is required to regulate flowering pathways. As discussed in Chapter 5.4, further examination of expression should be undertaken using real time RT-PCR, as the increased sensitivity of this assay is best suited to the detection and quantification of small differences in expression. In this way, the effects of the blh3 and blh10 mutations on clock gene expression, and vice versa, may be more accurately assessed.

6.5.4 What is the role of BLH3 and BLH10 in hypocotyl elongation?
A standard hypocotyl assay, based on protocols described by Moller et al. (2003) and Duek and Fankhauser (2003), was employed to assess the light response of gi-2, blh3 and blh10 mutants. It was found that the gi-2, blh3 and blh10 mutants exhibited longer hypocotyls than Col under red light and low light conditions. The gi-2, blh3 and blh10 mutants do not respond to LL, far red or blue light in the same way, signifying that this developmental phenotype occurs in specific light conditions, which includes red and low light. In general, Arabidopsis accessions have longer hypocotyls when grown in red light compared to white light (Maloof et al., 2001). The elongated hypocotyls exhibited by blh3, blh10 and gi-2 mutants when grown in red light and low light may be better described as an exaggerated developmental response to these light conditions compared to wild type.

It is proposed that mutations which cause insensitivity to red light are within genes involved in the PHYB signalling pathway; although a number of circadian clock mutants also affect light signalling (Fankhauser and Staiger, 2002). It is obvious that the gi and phyB mutants do not have equivalent phenotypes. The phyB mutant is early flowering, whereas gi mutants are late flowering, indicating PHYB and GI have opposite roles in regulating flowering. Furthermore, the elongated hypocotyls displayed by phyB mutants in red light is relative to dark grown plants in all fluence rates tested, suggesting that these plants do not detect red light (Quail et al., 1995). In comparison, hypocotyl length in gi, blh3 and blh10 plants grown in red light are not equal to dark grown plants at any fluence rate tested (Huq et al., 2000). This implies that gi, blh3 and blh10 plants are still able to detect red light. Only the blh3blh10gi-2 mutant has a long hypocotyl phenotype in red light that is equivalent to that seen in dark grown seedlings. This additive phenotype in the triple mutant indicates that BLH3 and BLH10 can act independently to GI in response to red light. An independent function for GI and BLH3 in SD is also indicated by the additive effect of the blh3 mutation on the long hypocotyl phenotype of gi-2 mutants in specific SD conditions.

A further level of complexity was added by the lack of hypocotyl-length phenotype exhibited by all mutants when grown in alternative conditions. This suggests that varying light/dark treatments before the transfer to red light conditions can influence hypocotyl elongation in these plant lines. Experimental evidence suggests that the pre-germination light treatment can affect hypocotyl elongation of seedlings once they have been shifted to darkness (Alconada Magliano and Casal, 2004). In addition, hypocotyl growth is
controlled by the circadian clock, with hypocotyl elongation occurring primarily in the dark. This rhythmicity is apparent at germination and later coincides with cyclical cotyledon movement (Dowson-Day and Millar, 1999). GI has a role in circadian clock function and cotyledon movements (Chapter 1.4.3; Tseng et al., 2004). Therefore, it cannot be discounted that the effect of the light/dark/light treatment on hypocotyl elongation in this assay is entirely independent of the dark period growth that is mediated by the circadian clock.

The gi-2 mutant had significantly longer hypocotyls and petioles when grown in SD, compared to that of Col. Enhanced petiole elongation is thought to be a symptom of the shade avoidance response (reviewed in Kim et al., 2005b). As gi-2 mutants exhibit significantly longer petioles than wild type in SD this may be a further indication of impaired light signalling in these plants. The exaggerated hypocotyl and petiole response of the gi-2 mutant is not shared by the blh3 mutants, which supports the previous assumption that BLH3 can function independently to GI in response to non-inductive short day photoperiods. Petiole elongation in Arabidopsis is enhanced by an end-of-day FR light treatment in SD in a process is mediated by PHYB and GA biosynthesis genes, including GA20ox1. Incidentally, the GA20ox1 gene is targeted by BELL transcription factors (Chen et al., 2004; Hiramatsu et al., 2005), which may provide an intriguing angle with which to investigate GI, PHYB and BELL function in the shade avoidance response.

6.5.5 IMPROVING LIGHT RESPONSE ASSAYS

The experimental protocol employed during this section of work was useful for identifying light response phenotypes. However, experimental conditions were not ideal for the evaluation of the subtle and sensitive light response phenotypes exhibited by the gi-2, blh3 and blh10 mutants. In the trials carried out in this section of work, filter boxes were placed over plates containing the seedlings. It is possible that this system lead to an increase in temperature under the filter box. Flowering time, circadian clock and hypocotyl length phenotypes of gi mutants have recently been shown to be sensitive to temperature (Gould et al., 2006; Paltiel et al., 2006). The blr mutant, an allele of the BELL gene PNY, also exhibits a temperature sensitive phenotype and it is suggested that the expression of AG, a target of PNY/BLR, is regulated by the thermal clock (Bao et al., 2004). This raises the possibility that blh3 and blh10 are sensitive to temperature and that changes in ambient temperature could affect the results of these hypocotyl assays. Preferably, these assays
should be performed in a temperature controlled cabinet using light emitting diodes designed to produce light of a specific wavelength. Due to the complex nature of seedling de-etiolation and the many variables that interact to affect hypocotyl growth, the precise control of light and temperature are the minimum requirements for future assays. Fluence rate-response curves for hypocotyl growth in red and white light conditions could then be performed. These would be useful for determining the fluence rates at which BLH3 and BLH10 are required. Finally, under light emitting diodes plants could be grown under red and low light regimes in soil instead of tissue culture. This would allow phenotypes such as flowering time and petiole length to be examined and compared to those obtained in standard LD and SD conditions.

Finally, difficulties in seed germination presented problems throughout this piece of work. The 35S:GI transgenic plant lines in particular were consistently slower to germinate than Col wild type. Previously, 35S:GI transgenic lines generated in Ws ecotype plants (Milich, 2001) were also difficult to germinate. During this work, fresh seed collected from blh3 and blh10 mutant plants germinated at approximately the same rate as wild type, although the germination rate of these lines decreased more rapidly with an increase in the age of the seed. If GI acts downstream of PHYA and PHYB, it is plausible that altered GI expression has an effect on seed germination, as these phytochromes play an important role in germination (reviewed by Bentsink and Koornneef, 2002; Wang and Deng, 2004). To determine if the germination of these lines is increased under different light conditions in a way that is mediated by PHYB or PHYA, seeds could be assayed for radical formation after illumination by red or far red light respectively (as described by Oh et al., 2004).

6.5.6 SUMMARY

Transgenic blh3 and blh10 mutants were identified and analysed; it was found that blh3 and blh10 were not classical late flowering mutants and that BLH3 was not required for GI to promote flowering in LD. Yet, BLH3 and BLH10 are likely to influence flowering at some level, as the blh3 and blh10 plants exhibited delayed flowering in specific SD conditions. Furthermore, blh3 GI/gi-2 heterozygotes flowered later than blh3 mutants alone in LD, which was suggestive that the levels of BLH3 and GI together affect flowering. GI acts downstream of photoreceptors such as PHYA and PHYB and the expression of GI is affected by light. Hypocotyl elongation assays implied that BLH3 and BLH10 act in
pathways that are responsive to red and low light conditions, although the phenotypes of these mutants are subtle and sensitive to experimental conditions. The results presented in this section provide the foundation for an exciting new approach with which to investigate the function of GI, BLH3 and BLH10. It is possible that a partially over-lapping role in flowering and seedling de-etiolation exists between BLH3 and GI under specific light conditions. Despite many remaining questions, this work provides a platform for further investigations with GI, BLH3 and BLH10.
CHAPTER SEVEN: CONCLUDING DISCUSSION

7.1 INTRODUCTION

Flowering at the right time is essential to maximising the reproductive success and survival of a plant species. The interaction of environmental cues with a network of genetic pathways co-ordinates the induction or repression of flowering in response to a changing world. Of the environmental signals, one of the most significant is day length. Molecular and genetic studies in *Arabidopsis* have identified genes that induce flowering in response to day length, which have subsequently been placed in the photoperiodic pathway to control flowering (Chapter 1.4.3.5). The *GI* gene is proposed to act in the photoperiodic pathway, yet *GI* also has more general effects on plant development (Chapter 1.5). The *GI* gene has a role in regulating the circadian clock and expression of *GI* itself is an output of the circadian clock. In addition, *GI* has been implicated in both red light signalling and temperature compensation pathways. It has become apparent that this gene has a complex role in response to various environmental signals.

The primary goal of this thesis project was to identify and characterise GI protein interactors. Evidence presented in Chapters 3 and 4 revealed that the BLH3 and BLH10 proteins are closely related members of the plant specific BELL family of transcription factors and demonstrated that both proteins interacted with GI. The results of these experiments are discussed in detail in the summaries of the respective chapters. Expression analyses verified that like *GI*, BLH3 was expressed in all tissues and stages of development and established that the BLH3 protein was localised to the nucleus. Contrasting with *GI*, the expression of BLH3 and BLH10 was not directly regulated by the circadian clock (Chapter 5). Analysis of the *blh3* and *blh10* mutants provided a possible functional link with *GI*, as these mutants had flowering time and light response phenotypes (Chapter 6). This concluding chapter further expands on the summaries presented in the final sections of these chapters.
7.2 REMARKS ON GI AND BELL INTERACTIONS

The yeast 2-hybrid assays employed in Chapters 3 and 4 verified that proteins derived from deletions of the closely related \(BLH3\) and \(BLH10\) genes interacted with GI. The information gained during this work provided an exciting new perspective with which to investigate firstly, the function of GI and secondly, a novel role of BELL proteins. The results obtained during this research were compared to recent publications on BELL proteins, which illustrated that BLH3, BLH10 and GI function within a complex protein interaction network.

7.2.1 THE BELL PROTEIN INTERACTION NETWORK

Reverse yeast 2-hybrid assays established that the BLH3 and BLH10 proteins hetero-dimerised and the BLH3 protein homo-dimerised via the BLH3 BELL domain (Chapter 4). The full length BLH3 and BLH10 proteins have recently been shown to interact with six and three other BELL proteins respectively. In fact all BELL proteins tested, with the exception of PNF (BLH8), interact with at least one other BELL protein (Hackbusch et al., 2005) (Figure 7.1). The BLH3 and BLH10 proteins, with BLH1 and BLH4, form a fully interconnected centre to the BELL interaction network. There is no obvious pattern to these interactions; they occur between closely related proteins and less similar BELL proteins. Only BLH1, BLH3, and PNY (BLH9) homodimerise, yet these proteins are not particularly similar or closely related as they are grouped separately in the phylogenetic analyses carried out as part of this work (Chapter 3.3.4).

![Protein interactions within the BELL family](image)

**Figure 7.1** Protein interactions within the BELL family

A summary of interactions between GI, BLH3 and BLH10. The additional protein interactions between BELL family members were determined by Hackbusch et al. (2005).
Both the BLH3 and BLH10 proteins were shown to interact with GI in yeast 2-hybrid. In contrast, the less closely related BELL proteins, ATH1 and BELL1, did not interact directly with the GI protein (Chapter 4). However, ATH1 and BELL1 are connected to BLH3 and BLH10, both directly and indirectly, through protein interactions between BELL family members. Consequently, GI is connected via BLH3 and BLH10 to the majority of BELL proteins via this interaction network (Figure 7.1). The significance of this interaction network and its affect on BLH3, BLH10 and GI function is unknown.

7.2.2 PROTEIN INTERACTIONS AND BELL FUNCTION

It is uncertain what role the protein interaction network has in BELL protein function. It is unlikely that the functions of the BELL family members are mutually exclusive, as the majority of BELL proteins are inter-connected through protein interactions (Chapter 7.2.1). The BELL interactions have been established using yeast 2-hybrid and it has not been demonstrated that these interactions occur in plants. Unless BELL proteins are constitutively expressed in plants, it is probable that BELL interactions are restricted by the coincidence of protein partners in specific tissues or at particular times. A comparison of independent *in situ* hybridisation assays show that *ATH1, BELL1, BLH3, PNF* and *PNY/BLR* transcripts are all located in indeterminate shoot apex and determinate floral meristem tissue (Bellaoui *et al.*, 2001; Smith and Hake, 2003; Bao *et al.*, 2004; Smith *et al.*, 2004; Cole *et al.*, 2006). Yet, these five genes have distinctly different expression patterns, so the regions where these genes overlap and thus are able to interact, is spatially restricted. *In situ* hybridisation analyses to pinpoint the location of *BLH10* expression and compare this to *BLH3* will determine where these genes overlap and where interaction may occur in plants.

The PNY and PNF proteins do not interact, yet the *PNY* and *PNF* genes act in a dose dependent manner with KNOX protein partner STM to regulate inflorescence development (Kanrar *et al.*, 2006). If the function of BELL proteins is dose dependent, this raises the possibility that BELL::BELL protein interactions have a dominant negative role in plants, sequestering BELL proteins and preventing interactions with other protein partners. Experimental evidence confirms that BELL::BELL dimers bind DNA less effectively than BELL::KNOX dimers (Smith *et al.*, 2002; Chen *et al.*, 2004; Viola and Gonzalez, 2006), which implies that the efficient binding of target genes by BELL proteins requires interaction with KNOX proteins. Interestingly, BLH3 and BLH10 also interact with a
member of the bHLH family of transcription factors (bHLH76) (Hackbusch et al., 2005). Related bHLH proteins bind to G-box DNA motifs as homo- and hetero-dimers (Toledo-Oritz et al., 2003). Could BELL::bHLH76 dimers act as regulatory molecules by preventing binding of the complex to DNA motifs recognised by BELL::KNOX or bHLH::bHLH dimers? DNA binding assays could determine if BELL::BELL, BELL::KNOX or BELL::bHLH protein complexes affect the DNA binding properties of the BLH3 and BLH10 proteins.

The BLH11 protein has not been tested for interaction with other members of the BELL family. This protein is predicted to contain a small 7 amino acid insert within the homeodomain region which is likely to decrease the DNA binding ability of this putative transcription factor. It would be useful to establish if BLH11 interacted with BELL or KNOX proteins, as it would be probable that any interaction would have a negative effect on the DNA binding ability of the protein complex. Alternatively spliced BLH10 transcripts that encode a truncated protein lacking a homeodomain have also been identified (BLH10-R) (Chapter 3.4.2). Truncated BLH10 proteins have the potential to play a negative role in regulating DNA binding, particularly if interaction with BELL or KNOX proteins still occurs. Utilising the BLH10-R cDNA in reverse yeast 2-hybrid assays could establish if the truncated BLH10 protein interacts with BLH3 or KNOX protein partners such as KNAT5, which are known to interact with the full length BLH10 protein. These experiments could resolve if truncated BLH10 and BLH11 proteins have a functional role in BELL protein interactions.

7.2.3 GI AND BELL PROTEIN INTERACTIONS

7.2.3.1 PROTEIN SEQUENCES THAT MEDIATE INTERACTIONS

The deletion assays carried out in Chapter 4 recognised the BELL domain as the smallest independent region within BLH3 and BLH10 that can mediate interaction with GI. The BELL domain within BLH3 also mediates interaction with the KNOX protein STM (Cole et al., 2006). Together, these results support predictions that the amphipathic α-helices within BELL domains are involved in protein:protein interactions.

Three GI deletions interacted with BLH3, the smallest only 70 amino acids long (Del4) (Chapter 4, Table 4.2). Remarkably, 63 amino acids within Del4 are absolutely conserved between three plant species, suggesting conservation due to function in the region.
Nevertheless, comparison of the sequence of this small deletion to proteins of known function reveals no obvious interaction domains. Diverse protein binding domains have been described in the literature and comparisons show that they are typically 35-150 amino acids long. Del4 is well within these margins at 70 amino acids. Proline rich motifs are often associated with protein interaction domains, yet analysis of the BLH3, BLH10 and GI sequences confirms that these proteins do not contain well conserved proline rich regions. The recognition sequences in some proteins targeted for interaction are less than 10 amino acids long (reviewed in Kay et al., 2000); it is possible that although GI contains no sizable domains that mediate protein interaction, this large protein may contain any number of small motifs that are recognised by other proteins. One way to establish which amino acids are directly involved in interaction with BLH3 would be to target particular sequences within Del4 by site directed mutagenesis and identify mutants that no longer interact with BLH3. Alternately, one could work backwards and identify peptide ligands that are targeted specifically by the BLH3 protein and determine if and where these are present in the GI sequence. It would be informative to compare BLH3 and BLH10 targets to motifs within the GI protein and establish if these are the same or if the BLH3 and BLH10 proteins interact with different regions of GI.

7.2.3.2 **WHAT CAN BE LEARNT FROM BELL::GI INTERACTIONS?**

The GI, BLH3 and BLH10 proteins form a tri-angular interaction network in yeast 2-hybrid, but how the proteins interact and act in plants has yet to be elucidated. These three genes are widely expressed throughout plant tissues and time of development (Chapter 5 and Fowler et al., 1999), as a result where and when protein interaction occurs is yet to be determined. Antibodies to BLH3 and GI could be utilised in immuno-localisation assays to identify distinct regions where the BLH3 and GI proteins coincide. BLH3 interacts with multiple proteins, including BELL, KNOX and AtOFP proteins. The interaction of BLH3 with itself, BLH10, STM and GI have been shown to occur via the BELL domain region (Chapter 4 and Cole et al.; 2006). Does BLH3 interact with these proteins at the same time in plants or is each interaction restricted to a particular time or place? Comparisons of immuno-localisation assays to the expression patterns of other BLH3 interacting proteins may provide important information on where specific protein interactions can occur in plants.

What other proteins are involved in the BLH3::GI interaction network? It has been established that GI interacts with SPY and that these genes act in light, flowering and
circadian pathways (Tseng et al., 2004). The BLH3 and BLH10 proteins also interact with a putative bHLH protein (bHLH76) in yeast 2-hybrid (Hackbusch et al., 2005). The bHLH proteins form a large transcription factor family in Arabidopsis and multiple members have been shown to interact and act downstream of the phytochromes in light signal transduction (reviewed by Duek and Fankhauser, 2005). The bHLH76 protein is proposed to recognise the same G-box DNA motif as putative phytochrome interactors (Toledo-Oritz et al., 2003), although forward and reverse genetic approaches are required to determine the role of bHLH76 in plants. As the gi and blh3 mutations have an affect on hypocotyl elongation in specific light conditions, an interaction with a possible phytochrome signalling protein may provide a significant link between light and BLH3::GI function. A co-immunoprecipitation technique could be employed to confirm if the SPY or bHLH proteins form part of a GI::BLH3 complex in plants.

7.3 Perspectives and Future Work

7.3.1 BLH3 and BLH10 Function in Plants

The BELL genes encode putative transcription factors, for this reason it would be useful to establish which genes are the targets of these proteins. The BELL protein BLR (also known as PNY) has been shown to bind directly to cis-regulatory elements within the AG promoter. In addition, transcription of AG was de-repressed in blr mutants, suggesting BLR functions as a transcriptional repressor of AG (Bao et al., 2004). Transcriptional targets of the BLH3 and BLH10 proteins and protein complexes containing these proteins need to be determined. Utilising BLH3 specific antibodies, chromatin immunoprecipitation is one technique that could verify the target of the BLH3 protein in vivo. Genes recognised by interacting BELL and KNOX proteins would be a useful place to start, as is has been suggested that the recognition of target sequences may be conserved within the BELL and KNOX families respectively (Viola and Gonzalez, 2006). Subsequently, the affect of the blh3 mutation on the expression of target genes would determine if BLH3 acts positively or negatively to regulate transcription.

Even though the PNY and PNF proteins do not interact in yeast 2-hybrid, these genes act in a dose dependent manner to regulate inflorescence development (Hackbusch et al., 2005; Kanrar et al., 2006). This implies that a direct interaction between PNY and PNF is not a prerequisite to their function in floral development. Is it possible that BELL proteins
act generally in a dose dependent manner? The \textit{blh3} and \textit{blh10} mutants exhibit a similar phenotype, indicating these genes have a related function in plants. However, the \textit{blh3blh10} double mutant does not have a more severe or additive phenotype, suggesting that the \textit{BLH3} and \textit{BLH10} genes do not function together in a dose dependent manner. The generation and analysis of double, triple and even quadruple mutants containing \textit{blh3} and \textit{blh10} with other BELL genes may provide more information on the dose dependent functions of BELL genes. The \textit{BLH6} and \textit{BLH7} genes are the next most similar BELL genes to \textit{BLH3} and \textit{BLH10} so would be the most obvious place to start. This would establish if these four closely related genes have overlapping functions in \textit{Arabidopsis}.

### 7.3.2 A complex role for \textit{GI} in plants

It is hypothesised that temperature, light and day length cues are integrated by the plant circadian clock (Chapter 1.4.3). Stress induction experiments indicate that the circadian clock can compensate against stress in the common ice plant and stress responsive redox regulation and day length signals interact to minimise oxidative damage when plants are grown in high light LD conditions (Becker \textit{et al.}, 2006; Boxall \textit{et al.}, 2006). \textit{GI} has variously been found to function in pathways that respond to stress, day length, temperature and light (Chapter 1.5). It is tempting to suggest that \textit{GI} functions where these apparently distinct pathways that respond to environmental signals intersect, ultimately to induce flowering at the appropriate time.

#### 7.3.2.1 The light response of \textit{GI}

The results presented in Chapter 6 demonstrated that the \textit{gi}, \textit{blh3} and \textit{blh10} mutants responded to subtle differences in light quality to affect both flowering and hypocotyl length. Although \textit{GI} was proposed to act in a \textit{PHYB} signalling pathway, the accumulation of additional evidence suggests \textit{GI} functions downstream of several different phytochromes and cryptochromes. The \textit{GI} gene is responsive to red light and this is mediated by \textit{PHYB} (Tepperman \textit{et al.}, 2004). \textit{GI} is also affected by light quality, as the \textit{GI} transcript is up regulated in plants grown in shade light conditions in a response that is mediated by \textit{PHYA} and \textit{GI} expression is reduced in \textit{phyA} mutants (Fowler, 2000; Devlin \textit{et al.}, 2003). The cryptochromes also regulate \textit{GI} expression in certain light conditions; \textit{GI} is up regulated in white and blue light in wild type plants, which requires functional \textit{CRY1} and \textit{CRY2} (Paltiel \textit{et al.}, 2006). Establishing the effect of phytochrome and cryptochrome mutations on \textit{GI} expression and conversely, the effect of the \textit{gi} mutation on the circadian
regulation of phytochromes and cryptochromes will help to orientate GI in pathways downstream of the photoreceptors. The relationship between GI and the photoreceptors could be further elucidated with the analysis of flowering and light response phenotypes of gi phy and gi cry double and triple mutants.

The PHYA and PHYB transcripts are regulated by the circadian clock, with a peak during the light part of the daily cycle (Toth et al., 2001). The PHYA and PHYB proteins do not vary in constant conditions, yet nuclear import and hence function depends on both light and the circadian clock (Kozma-Bogna et al., 1999; Kircher et al., 2002). Therefore, in addition to regulating light input to the circadian clock, these photoreceptors must function downstream of the clock. The levels of PHYA and PHYB are not reduced in gi-100 mutants (Huq et al., 2000). However, altered circadian expression patterns or nuclear import of the proteins in gi mutants can not be discounted. Taking this evidence into account, the light phenotypes seen in gi mutants may be indirectly related to an altered circadian function in these plants.

7.3.2.2 THE CIRCADIAN CLOCK AND GI FUNCTION

The expression of GI is regulated by the circadian clock, with peak expression differing in plants grown in promotive LD compared to non-inductive SD (Chapter 1.5.3). Yet the significance of the circadian expression of GI is unknown. Recent evidence indicates that the timing of GI expression during day/night cycles may be important for GI function. The cyclical pattern of GI expression is altered in early flowering tic mutants. The peak of GI expression occurs earlier during the day in tic plants compared to wild type and although the overall expression level of GI is reduced in these mutants, these plants are early flowering (Hall et al., 2003). Controlling GI expression using chimeric or inducible promoter systems could determine when GI is required for floral promotion.

Recent evidence supports distinct circadian and flowering functions for GI (Gould et al., 2006). The flowering phenotype of toc1 mutants was shown to be a direct consequence of the short circadian period caused by toc1. Growth of toc1 plants in 21 h daily cycles that matched the internal short circadian period of the mutant restored wild type flowering in these plants (Mas et al., 2003a). The gi-1 and gi-2 mutations decrease and increase the period length of CAB gene expression respectively, whereas both mutants flower late in LD (Fowler et al., 1999; Park et al., 1999). Growth of the gi mutants in daily cycles that matched the internal circadian period in these plants may resolve if the late flowering
phenotype of either of these mutants is an effect of the altered circadian rhythms in these plants or if the flowering time and circadian regulation of GI are indeed separable.

### 7.3.2.3 Identifying GI Targets

The completion of the *Arabidopsis* genome sequence ([Arabidopsis Genome Initiative, 2000](#)) was the prerequisite to global expression experiments. The use of DNA microarray analyses to analyse the entire transcriptome of *Arabidopsis* is now commonplace. As part of this thesis work, two chemical inducible GI expression systems were analysed (Appendix 7). Preliminary findings indicated that a dexamethasone inducible GI expression system may prove useful for controlled expression of GI in transgenic plants (Appendix 7.3). As GI interacts with BELL transcription factors, it is possible that this interaction affects gene expression in plants. The induction of GI in the presence of a translational inhibitor such as cycloheximide and the global effect on gene activation and repression could be investigated using microarray technology. This experiment may be used to identify and isolate genes that are directly up- or down-regulated in response to GI. Alternatively, GI expression could be induced and the effects on protein levels could be investigated using large scale proteomics with mass spectrometry. Identification and analysis of these proteins will contribute to the model pathways that regulate flowering in *Arabidopsis*.

### 7.3.2.4 Where is GI Required to Promote Flowering?

A variation of the dexamethasone inducible GI expression system (Appendix 7.3) could be used to regulate the expression of GI in a tissue specific manner. This experiment may help to establish where in the plant GI is required. For example, is the induction of GI in the leaves sufficient to induce flowering at the shoot apical meristem? Does exclusive expression of GI at the shoot apical meristem affect flowering? These types of experiments could determine where the GI signal acts to induce flowering. This approach has been successful in locating the tissues in which the broadly expressed floral gene CO acts to induce flowering. An *et al.* (2004) demonstrated that mis-expression of CO from phloem specific promoters, but not meristem specific promoters, can trigger early flowering.

It has been recognized that GI is induced by temperature and light of differing qualities ([Fowler and Thomashow, 2002; Devlin *et al.*, 2003; Tepperman *et al.*, 2004; Cao *et al.*, 2005; Paltiel *et al.*, 2006](#)). The effects of these inductive cues on the expression of a
GI:reporter gene fusion (ie GI:GUS) could also determine where GI is induced in response to environmental signals.

### 7.3.3 APPLICATION OF KNOWLEDGE TO OTHER PLANT SPECIES

An important challenge of plant molecular biology is to take what is learnt in the laboratory, from model species such as Arabidopsis, and apply it to the real world. In life outside plants do not exist in a controlled environment and as a result detect and respond to a myriad of environmental cues. For example, subtle changes in light can variously signal that dusk is near; a neighbouring plant is shading the light; possibly a change in season approaching. Plants must react to unexpected and often opposing abiotic stresses, such as cold and heat, flooding and drought. The ability to respond to the environment is vital to maximising reproductive success and the survival of the species. It is evident that GI has an significant role in the response to environmental cues such as light, day length, temperature, biotic and abiotic stress and this culminates in the promotion of flowering in Arabidopsis. With the identification and characterisation of GI orthologs in rice, barley and wheat, it has become apparent that some of the characteristics of GI have been conserved in other plant species. This leads to an exciting avenue of investigation into the role of GI, not only in other plants but in the different environments in which these plants grow and reproduce. For example, GI functions in both Arabidopsis and rice to promote flowering, although these plants respond to opposite day length cues of long day and short day photoperiods respectively. Further research on the function of GI in different species will ascertain if the role of GI in the response to specific environmental variables is conserved in angiosperms in general.

How relevant the knowledge of flowering time regulation in an annual plant such as Arabidopsis is in explaining seasonality and other aspects of development in long lived tree plants has yet to be elucidated. For example, the central floral integrator FT promotes flowering in aspen trees as it does in Arabidopsis. Yet, FT in aspen has an additional role in the response to day length, by regulating bud set and growth cessation (Bohlenius et al., 2006). Furthermore, increased expression of a Populus relative of FT induces early flowering, partly by reducing the juvenile phase during which trees are not competent to flower (Hsu et al., 2006). Together, these results suggest that FT-like genes in trees have a complex role in day length response and floral promotion.
GI-like genes have been identified in tree plants such as poplar and pine. BELL genes have been identified in the gymnosperm Gnetum and putative BELL-like genes are expressed in poplar and pine. It would be interesting to determine if the functions of GI and BELL-like genes have been conserved in tree plant species or if the evolution of novel roles for these genes has occurred. This would provide more information on the role of GI in regulating flowering in plants and ultimately the function of GI in the real world.

7.4 OVERVIEW

Many research articles have been published in recent times concerning the control of flowering in Arabidopsis. Collectively, this information has contributed to the formation of a hypothetical network of interacting genetic pathways which act to synchronise flowering time to the appropriate environmental conditions. The elaborate role of GI in plants has yet to be fully explained; classically, GI has been placed in the photoperiodic pathway to promote flowering in response to day length. The perception of day length and the integration of these signals are thought to be mediated by the circadian clock. GI expression is controlled by the circadian clock, but GI also regulates clock components in response to both light and temperature cues. In addition, GI has been implicated in light and stress responses in Arabidopsis. In this thesis research the putative GI protein interactors BLH3 and BLH10 were characterised. The analysis of loss-of-function mutants revealed that unlike gi, blh3 and blh10 were not classical flowering time mutants. Further results presented suggest that the BLH3 and BLH10 genes affect flowering and seedling de-etiolation in a manner that is sensitive to light conditions. The results presented here have provided a new perspective for further investigation into the role of GI and how developmental responses such as flowering and de-etiolation are regulated in Arabidopsis.
APPENDIX 1: ANALYSIS OF THE BELL FAMILY PROTEINS

APPENDIX 1.1  PAIRWISE COMPARISON OF AMINO ACID SEQUENCE OF PREDICTED ARABIDOPSIS BELL PROTEINS

The predicted amino acid sequences of the thirteen *Arabidopsis* BELL proteins were compared using the ClustalX algorithm. Percentage amino acid identities are shown, identities of >50% shown in bold type. The BELL proteins numbered 1-10 are annotated according to that of Hackbusch *et al.* (2005). As BLH8 and BLH9 have been previously characterised, these proteins are referred to as PNF and PNY respectively in this thesis (note: PNY has also been characterised as BLR, VAN and RPL).

<table>
<thead>
<tr>
<th></th>
<th>BELL1</th>
<th>BLH1</th>
<th>BLH2</th>
<th>BLH3</th>
<th>BLH4</th>
<th>BLH5</th>
<th>BLH6</th>
<th>BLH7</th>
<th>PNF (BLH8)</th>
<th>PNY (BLH9)</th>
<th>BLH10</th>
<th>BLH11</th>
<th>ATH1</th>
</tr>
</thead>
<tbody>
<tr>
<td>BELL1</td>
<td>28</td>
<td>32</td>
<td>27</td>
<td>31</td>
<td>28</td>
<td>29</td>
<td>31</td>
<td>23</td>
<td>26</td>
<td>25</td>
<td>29</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>BLH1</td>
<td>27</td>
<td>31</td>
<td>39</td>
<td>30</td>
<td>36</td>
<td>22</td>
<td>28</td>
<td>31</td>
<td>36</td>
<td>31</td>
<td>36</td>
<td>36</td>
<td>23</td>
</tr>
<tr>
<td>BLH2</td>
<td>31</td>
<td>66</td>
<td>31</td>
<td>34</td>
<td>31</td>
<td>32</td>
<td>25</td>
<td>26</td>
<td>31</td>
<td>31</td>
<td>33</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>BLH3</td>
<td>29</td>
<td>28</td>
<td>33</td>
<td>35</td>
<td>26</td>
<td>25</td>
<td>61</td>
<td>38</td>
<td>31</td>
<td>35</td>
<td>31</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>BLH4</td>
<td>29</td>
<td>31</td>
<td>33</td>
<td>32</td>
<td>25</td>
<td>27</td>
<td>29</td>
<td>29</td>
<td>35</td>
<td>35</td>
<td>31</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>BLH5</td>
<td>27</td>
<td>29</td>
<td>24</td>
<td>28</td>
<td>30</td>
<td>22</td>
<td>31</td>
<td>22</td>
<td>31</td>
<td>31</td>
<td>32</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>BLH6</td>
<td>26</td>
<td>29</td>
<td>58</td>
<td>24</td>
<td>31</td>
<td>37</td>
<td>21</td>
<td>21</td>
<td>37</td>
<td>31</td>
<td>22</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>BLH7</td>
<td>26</td>
<td>25</td>
<td>34</td>
<td>37</td>
<td>37</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>30</td>
<td>28</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNF</td>
<td>30</td>
<td>22</td>
<td>25</td>
<td>31</td>
<td>22</td>
<td>38</td>
<td>21</td>
<td>21</td>
<td>38</td>
<td>21</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNY</td>
<td>25</td>
<td>31</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLH10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLH11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATH1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX 1.2 THE SEQUENCES OF THE BELL FAMILY PROTEINS USED IN PHYLOGENETIC ANALYSES

Forty-five protein sequences representing the BELL family were used for this work. The protein sequences used in these analyses were derived from cDNA sequences for thirty-seven BELL genes, including thirteen Arabidopsis genes. Another six were predicted protein sequences from the annotated rice genome sequences. Five of these putative proteins were supported by cDNA sequences, however these were not used in sequence analysis as they were either partial or alternatively spliced clones. Sequences were identified in Genbank and the accession numbers are listed below. A further two partial cDNA sequences encoding BELL proteins from wheat and sorghum were derived from EST sequences from the TIGR Gene Indices. These sequences were identified after the phylogenetic analyses were completed, however were included in the pileups as examples of BELL proteins from other plants. These sequences are listed at the end of this section.

Arabidopsis thaliana At; Gnetum gnemon Gg; Hordeum vulgare Hv; Lycopersicon esculentum Le; Malus domestica Md; Oryza sativa Os; Sorghum bicolor Sb; Solanum tuberosum St; Triticum aestivum Ta; Zea mays Zm.

AtATH1 AY096513
MDNNNNTSFDLWMTQNPLLMDIPSPREDSTFTSTMLPWNTIRSDPLQMGFDINSNLTONYSSPR
SDFVQDNVEFMAPPEPPPPPHELHDLRHYDSSSNNMNFQAESEFAQFSGVGPSEPMMSTFGEDFPLI
SNKRNELSLSASVGDSECEIPLCNAATILASEQASCSTKDISNNVCTQGFSLIFGSKYLSHEQISFHA
AYSIDLYSSRGTSESSGAASSAPTFSRENETILGUSSSSNENAEFGSTFGQRRALEAKHMLLLQLMVMDYRSCH
DEIHTVISAFAHAAETLQHRFLQGFQVSEQLRLEQHLREGQKCISSGMSLVERGKDKTQETSMPHQCLQLQL
KKRKNHQIRFRQGRLPEKSVLVRLKNMFQNFLHPYKPDSEKHLLAISGLTRSQVSFNWIFARNVRVLKWPMIEEM
YAEMLKLNHHQSNPRLMPKSVMSQAMHK

AtBELL1 U39944
MARDQFYGHNNHHQEQHQMINIQQGQDETNQPTDHDYHNHQIFGSNNSNNMIDFSKQQIRMSTGSDHH
HHHHTQGSGTQDOGNDQQQLESASSMRLCWNNDPSEVPNEDRPQPSQGQLSLSSSLSSNSPTISLDNQFELRQQQ
QQGYSBNKSTQHQNLQHTQMMNNNMCNQHNNNNHHQIHQQQGISYSLSANEDFQPLQVSEQLVGEKESDE
EVMIMKHHKQKKQKQEEWDTHSNNNHHQDQSAATTSSKHKVPLHLSLFMLQKRKLLSMIIMKLKRYG
YREMQRDAAAFAEVEAVLGGAEIYALIASRAMSHFRCLGQVQIATSQALGEREVDNRAVLSAARGETP
RRRRDQALRQRKSYRQMERTLDAHPRWPRQGLPERAPVTTLRAWLFEHFLHPYPSVDKHLARQTLGSRQVS
NFNIFARLVRKWPMIEETRESEQQMEKQMTNDDMQMNTDPQDQVLQVFEQESLSSLSITNPSHTGNSHSTGTM
SLGSQDFDSLQYGQNVTAYEGGGRGPDGDVLSTTLGLQRNCDGNGVSLASLPVTQAOGQYLFGRDHIEEGPVQYSA
SMLDDQQVLPYRNMLGAQQLLHID

AtBLLH1 AF353094
MAYFYHNGNPESIASDGGGLTLILMNPTTVYQTVQQNDSSNNNSSSSNNNTNTNNNNESSVFLESHA
PQPNASQOFVGIMLGSHEAECSTALADNSLHSGYQYQDSQTPNQAECTFQAPQGQLLSTLSQ
QQQQQQHQHQHPQHVGPSGSHGEDRVQGSGTSGSVTNIGANLVSQSKYKQAEKLIQELDVVDNADDDNNAKSQ
LFSSKKGCGNDKPGVESASAGAGGESGGGAEEAAGRKPVEQLTREQERIQMKKALSNLMEQVQYRQYHQ
MQCVIQSEQAGAGSXIITYSLAKTSRSQFRCLKEIAQIQIQAANKSLGEEEDSVSRGFFGRLKFDVHH
LRQQOLQALQMQHIQPSSNARNRPQGLPERAVSRLRAWLFEHFLHPYPKDSKMLAKQTLQTSRQVSFNWIFIN
ARVRLLKPMVEEYMEEEMEQAKMGNGMSKMLPDQSQLSDASAQSTSNQKSPMATNYHMNPNNHDLGVEVTG
MQGSQPRLRTSTDMQIPASDFDSNNEKLTMKLEEQGIRSDGGGYPMFNMQGQMDRFSRDVSDQELMA
QRYSQSNANGVSLTLSLOGSLSSTHQQGFMQTHHIPRGRVKIGETEYYGPATINGGSSSTTASSSAAAG
AYMGNNMIQNQKRYQAQLLPDFVA

AtBLLH2 AY050459
MGITKTSPNTTLLKTFHNNMSDQYHHHHHQHQGIFNSFNGFDRSDSPNLTTQQKQEHQVRMDDEESSV
AGGRIPYVSAGLSEMPNFGSSGSLGPPPLAQQSNRQLEQHHQIQNPAMTDSSATAAAAMQFLMNP
PPQOPQPOPSPSTFRSHHHNSTLMLPPSLPTNTHQNYTNHSSMHQLPHQHQQISTMQSSPDHHHHHHH
SQTEIGTVSHEGNSGQQLSLSSSLSEEAADAAYPREYYNYGNNSSNASPHQYQNFKTLSSNQHQQV
LRNFQFRSAPASSSSMAANLNSRYYTATQAELLEFFSVCVRGFLKNNKLGNSNPNTCGDGGGGSSSSAGA
NKEHFPSSASDRIEHQRRKVKTMLTMEEVDRYNYCEQMQMVSVNFDIVMNGHAALYPTALQKAMSRRHFRC
Appendices

LeBL1 AF375965
QHDMYQSDTTTATQYPHGLRSLRSIEFVNPDPDTTSDHVNDHSRHMLDGLIAGHADANTNQQARLQLSLGLQHSL
VSTFNNPNSYMNGIDQRNNNNFSSMAAAMQSFSSNVCGTESEFSAIGNLYKLPQSLSEELICGIGKTIDNESS
NEKFRIRLRSNKRSGSMLRRKLGIMPNNLEFNERHELVKMLALLEEVEREYQQHMEETVSTFE
VIAGFGAKAYTALALQAMRFFCRRDISLRSIINQINRMRKPMPDVPXISGSLSLFELKTQLRQLIQSGLQG
IQNSRNQWQPQIRLPESTSVAFLRSWLFEHFLHPYPDSEKLMLSSQTGLNKQNVSNWFINARVLRKPMIEE
MYKEEFAESVDSNLNREAVTDAAE

LeBL2 AF375966
MYOQGTSDNNIADQHAMQOQNHLGGGNNNINIQTLYLMNPNSYMQYGGTIDDTTOOQLQQQQQHQLFLNSAPAG
NALSHANIQHAPLQQHQFVGVPLAPVSLQHDPHINHHLQGLRRMNNQQDSSQVPSVSTPSATSCGGTTDLAS
QALAFQRIIVVSTPPQHRQQQQQQGGLSLSSLQPQQQISSFFFFNNISSSSTPRTNNVTIRGTMDCSSMIMLSKG
LKAQELDLEDSVINVGCNKSNNQDKDNSMNKELIPLVSDVNTNNSGGGGESSQKNEVAIELTARQQL
QMKKAFLAMLAMLEQRYRHYHQQMIIVSSSEQVAGVSGAKSYTQLALHAIKSRFPCLDASIEQKVATSKSL
GEDQELGCGIEGSRKLFVHLDDQLAPMQMMNRPPQRGLPERALSVLARWLFHFLHPYPKDSDKIM
MLAKQTGLTRSQVSNWFINARVLRKPMVEELEYEVKERNQQNSNSTDKNKKEITSNAPNEEKQTISSL
LDQGTQAEITISTSTHISTASLHLNFHNFSGMNFMENTTDTIHHTHNAKNPKMKDHFPPSSLLSVE
EMEKAESTRNKGFNPMLAAAYAMGDQFRDFPHDQQTAMFHNQGVSLTGLPLPSENALMPSQVQNYLSNEL
GSRPEIGSHYNGNYENIDFQSNGKRFQPTQLLPDFVGTNLGT

LeBL3 AF375964
QHETMLATTTTTHSDDWHHDNRRNVLIDDDPMSRVCVFCEGNERPSQGLSLCSLNPSIQLQSFELRHQDLQ
QLISLHDGLKSTSIQQYFPRVLQSYSLQKGAELLELFSCSLGF1KNNDKSSHSSLKLKHDATIITSKQK
LQLSLHLEQKRKTLQLMLEEEDRYKRRKDMQKVVSSVESAVNGAATAYSLASLMRSHFRCLRDGI
VAQIKATKAMAMGQKDTSTITLPEGSTGETPRLLDQTLRQQKAFOQMMHMFWRPPQRGLPERSVSVIARWLFL
EHFLHPYSDVDHKLRQSLGSQVSNRFINARVLRKPMVEELEYEVKERNQQNSNSTDKNKKEITSNAPNEEKQTIS
LDQGTQAEITISTSTHISTASLHLNFHNFSGMNFMENTTDTIHHTHNAKNPKMKDHFPPSSLLSVE
EMEKAESTRNKGFNPMLAAAYAMGDQFRDFPHDQQTAMFHNQGVSLTGLPLPSENALMPSQVQNYLSNEL
GSRPEIGSHYNGNYENIDFQSNGKRFQPTQLLPDFVGTNLGT

LeBL4 AF375967
RQCDSCSGVNTAHAFQRFREDSELTSCSNRNLSSLFGSYKPVYSLQFTGLTGRVLYRMQWEISEIAQLSLQN
HNLNVYRGNCFTENGASTFALSNDGYYAMASSDSPDLMGCEAKKLNLLVALQVVDQYNCLEDBEIMV
SAFHAVTELDPSHARFLAQTISSLYKNLREGOSHNRMMQTQMRMRKSMERLVIKLHKSFFFHQQQOL
RDKHQLWRPQRGLPERSVSVLARWMFQNFHLHPYDKAEQOLLAVKSLGSTRQSVSNWFINARVLRKPMIEEMY
AEMNNRKIRANGHEDREESSHKLLTMK

MdMH1 AF053769
MGVITALPLPPPPPQNKLHHRHHSHSDSENYSNPNSDQYHJGIFTSNGFERSAMTHQEQQQQQQQHQLAJA
QIRKRLVQGFTETPPPPLVGLNEEESGLPLAYETAGLMSFMENPFPPGGPVGAAELLPEHMFPRNMRPP
QOAABSAADAWGRVGGGGGGLGDSDKQHHKHIQSTINTADSAAMQLFLMNPSQRPSPPPSSTTSTTLHM
LLNPSTTTTNLSQFAAASSGGAAGQFTWTEPESHGHEAGGNTAGGGEIIGVVEGQLDLSSLSSLQLHEAAL
KAEERFMSGDSAZSLAVNNQGDDQQQSQNPKNLQSHHQALHLSSQVGGHVGHSFFSGGGFVNNVLIR
NSKYKVAQELLEFCFGVRQLKNKFSTGSTSQRQNTTPNSSPANQSGGGGDDGQSSSSSDKVPLASADRI
EHQRXAVGLSMIDEVRNHYCEQMQVMQFNAFIDLVMGFGAAAVTLYMQAMSORHFRCLDAAIQLQHSC
ELIGEKDGASTQKTQFRRQAFMMEQAAEAPRQRGLPERSVNLIRLWEFHELHPY
SDADKNNLLARQGLSNQVSNFIRARVLRKPMVEEMQYEQANEVEEGGEGGVSATTQERERENSNAG
LAARQTPTTTTTTTTTSNPSASKREDINASEDSLVAINHRQHQHHMMATTTTSTTSAVAPCAAQFAPPAAS
AAMCRSYGTTSANANIAAHHDHQNNNSNDSSTSSLISFGTHTTAAGDVSTLTLHLHGAGGGGNNMDXTSSTSSFSI
RDFGCG

OsBEL AB218895
MAAHNNHHHEHHHLLDMSPPNASAGISSFDHAALGLSLHDVAAADDEHHHHLRRGGGGLQPSWPSQVQVSL
SLYNVNNAGAAGSSPSLLSHQALQQAAAQQPLFQRLGRSGLFYKALLAEFCSDLDEAMDAGKOQPMPNFLPK
KWDVDEGSGSWSGNLSSLRMDDLLEKRLARSLIMSEDDREVRRYRMQAREVSAEAGGQAQYVTKLAM
RMSRHRCCFRDLADLVQVRALRNMAGESQORDAAGVAAAAPGTAKGDTFLRLVLQDCLQRQAFQQSAVGDSF
OsBELL 8 AK070465

MATYYSSPGSERDSQMTDPSSALGLNYLNNPSSPYETFESGILQPPQONCMEMPQPGHASAMQPQDSRSSERMLSHQQSFQSHVHDMKMNLEMNHMMDDAQGSGHGSELDHDAHTGSLQEFGLVNLNNHSSSVPMQSGQLLSLINTQMPAYSIMKPDMLTPSYYDHNLQEDRMRKLNQASEAARINSRLYKAAQELLDELVSVWWSIKQKQAEKVEKESGDADGEKDGPQSGKPSVSNPQEGARAPAELLEAQEKLQMNKMLAMLMLDEVDKRYKHYHMQTVSVTSFDVAGASLVADAASYEEASLSLSCSLSMNAGDDQCSASSASRSLGTMQSRV

OsBELL 9 AK09849

MGSNPSQGSAVDAAMNGYFMAASNGADVLPHFPAFPMDHGSGSGFGYGDADAAAAMDVGAHAFAANNLVLASLATTQFLGFAQAAAPAHGDGDYDLGATTTPPEEMMGYDVAGDVSSSGGAVSCLAGHGHQPGDMAAWCTSAEPCSSNWWSNAGHGSYSFLYLVAGFAEAASALSLSLLSCSLSMNAGDDQCSSASSASRSLGTMQSRV

OsBELL 10 AC133217.2

MGNDGMQVEQRYRQHYQHQQVQVASSFEAVAGGSSARATYALATALRTISQRFCRDLDAIAIQGVRASAALGEAVADGCGCTRVRGLYIDHQLRQRALQLQMLMQSASSWPRQGRPLFRSIVILARLWHEFLHYPDKSDIMLAKQGTLRSTQVSNVFANRRLKMVFEMYELHEETKDDQGDGGAGAGDESKPGSKGGGAGVNGVGVDASAKMDSAKAAMESGQGVHPSLLELADHQRQAGYFLDDDEGADAAALQFLKLKARTEQEEQAAAFHSDVSDTILHAANAAARHDEVSHRELMKFMESGASAGAGAARADDHHEHHHGCVGVSFLAFAAPYGQFATEQFAPAGHGDSGGGGGVSLTLGHPHAQETASFLMTSSLQWSSGSSGAGVAGGVARLHRANFGIMDHPDAYSDDQHQQQHGFVSLTLGPHGGHGGGGVNIFAGAPGSAHGGAGFLLPGATEGQMDQQRFRHEPSLHFLDVF

OsBELL 11 AC03550

MGIAPPCCPOQPQTTTQISPSTSSKSPAIAIVVRRPPATAASSSSMHQSFGHQPSGAVGFGSDFDGRPGSGHDQHEQHQQVQHVASSSRRKRLQVQDFPAAAAGHDLFIEGDHEAGAFGMYDHEAEEAAAGAGSNMLSEMFFPSQFPTGPSATISTLEMQASQALQALQLMQSAASWPRQGRPLFRSIVILARLWHEFLHYPDKSDIMLAKQGTLRSTQVSNVFANRRLKMVFEMYELHEETKDDQGDGGAGAGDESKPGSKGGGAGVNGVGVDASAKMDSAKAAMESGQGVHPSLLELADHQRQAGYFLDDDEGADAAALQFLKLKARTEQEEQAAAFHSDVSDTILHAANAAARHDEVSHRELMKFMESGASAGAGAARADDHHEHHHGCVGVSFLAFAAPYGQFATEQFAPAGHGDSGGGGGVSLTLGHPHAQETASFLMTSSLQWSSGSSGAGVAGGVARLHRANFGIMDHPDAYSDDQHQQQHGFVSLTLGPHGGHGGGGVNIFAGAPGSAHGGAGFLLPGATEGQMDQQRFRHEPSLHFLDVF

OsBELL 12 AC119747.1

MADPNLGFADYFAAADAASSSVTILMPADMEAEAPPAFLGLQAGMELLGVLRGMLGSMMPQAGAKGAALKVAVLADAGDGSGGTMFSLQEGHPQASSPCQQPGLSDLVHLGARGAAPQPIALAPAPWTHDDASSAPQVGHAHLWSSLIFLPTQQLQECFSLPVDSTKRGNGAKAAQTQEDGRDGDDSSSSASNTPSFQPQAMEALELGQLDDKYLMLEDVDRYRRCQPPAPEFAVAGGVARAGMDDQGDSGGGSGNQPIKTSCHASEAREGGQLQVGGDDGQGQPQTRQULHRDAGLSVNVNDVAAAGGVGLHLHQANEFGIMDHDFAADDHSSQHQQGHGFGVGLTLQGISHGAGGGINFAASAGGHDGFGYFVMDAGMGHPGHQVAGMAAQAHERNQILNMAAKTYFKIPAAAQVESGKALFMMKGGRDHHRKDAIVMYAAAPCKKASKFSSCSLSSLYLLFLFLVLLVYLAVLASSFYSSSSCPEELASGSDKVHAAADAGNRRDSPPSDAAPATGLGHLHGFIAASLWKSRREYIRTMWFREPQFGSVWNLDFKYEVFYPSNAIGSTQAPIQKPFYTHGRSGASLRTIVESFRLGPGAEWFMDDDTVFPPDNSLVLVLSRYDHTQFYYYINPSHEIQLNIYFSMAGGFFGFAISRALAAMQLAHMDQGCD1RYPALYGSDDDHACAEVLGFLTRHLGHFCPLDQGVDGLGLGAPHVFVLFLHLHDLQFVVFTTRTSARTALRLEFGPALRDSAGVAGQVSSCVDGDQWTSSWVFQVAVVTRGVLSPREMEMPMTFLNWYRRADYTAFAFTRPVAPQRCP
Below are partial the BELL cDNA sequences identified in wheat and Sorghum, which were derived from EST sequences from the TIGR Gene Indices.

**StBE22 406701**

```
TAESQGNNYVPQHQPREWNGRQIEVGGSSVQVFAGTYPYCFMCCG
ASTSNIPGGVNLQSLHLOVEQIDLRRMNGGLEFFQGESNHNIFGSLGGLNVLR
NSKYYKATQELNECCFGQGKFLKINKVSRNNNSTSPINFSGNNNNSSKAIIPNLESTAEールDQH
RKVLMSLEDVEKRYCEQMQMVVNSFDLVMGFGTAAPYTALAQKAMSRHFRCLKDAIGAQ
```

**StBE29 406702**

```
QGGSLSSQQQGFNGFTAARELVSSPSGASASGIGQQQQQQSISSSPLSSKMKAAEQELDEEVNVGK
MKSTNSTVDDKKKGNMQMDGQDDDLVQGVDAGDKDAPTEGT1QRQOIQMKAKLVMLEDEVQEYRHYH
HMQSFVQHLEQAAGIAGSAXTALALATSQKFRCLDAIQIGRSATQGEGESLSKLGEQFVNDQ
LRRQALGQGMTHQNAWRFQGRLPERASVLRWFELHLHPYKSDKMMLAKQTGLTRSVQSNFWINAR
LRKFWMVEEMEQLGKEQSEDSSTRTSRSITATQDKFSGPSQSNKQSKNDHRQAPSMFVQHRHEFTIGMRNSQAGENLGSIESINSITQGGPSKPKRNMLHSNPSINSNMDVKEEQM
SMFQGSRQDFRDGLMGFMGFNGFGAYGIFIEARSFTEQPSAYSTGVSSTLIFLGHNENLMSATHSH
FLPIPTQNIQIGSEPHFEGSLNTPTSAHSTSSVYETFNIQRKRRAFAPPFLDFV
```

**StBE30 406703**

```
MATYYSSPGSERESQDMYSRDPGGASYPMSSALGNLLYNNPSSGPYTEFSGILQTQQNFMEMPGHGHHSAM
QDSSARESQDMLASHHQRSAVFGHDLQMDKNGMMDGQDDDLVQGVDAGDKDAPTEGT1QRQOIQMKAKLVMLEDEVQEYRHYH
HMQSFVQHLEQAAGIAGSAXTALALATSQKFRCLDAIQIGRSATQGEGESLSKLGEQFVNDQ
LRRQALGQGMTHQNAWRFQGRLPERASVLRWFELHLHPYKSDKMMLAKQTGLTRSVQSNFWINAR
LRKFWMVEEMEQLGKEQSEDSSTRTSRSITATQDKFSGPSQSNKQSKNDHRQAPSMFVQHRHEFTIGMRNSQAGENLGSIESINSITQGGPSKPKRNMLHSNPSINSNMDVKEEQM
SMFQGSRQDFRDGLMGFMGFNGFGAYGIFIEARSFTEQPSAYSTGVSSTLIFLGHNENLMSATHSH
FLPIPTQNIQIGSEPHFEGSLNTPTSAHSTSSVYETFNIQRKRRAFAPPFLDFV
```

**ZmKIP 4082396**

```
MVMAKGHHIDKGIKPKARNERRAPFPRLEQUCCRCDDAHQADESMDAASAGPLHTLGLGSAAGPSCGCVAP
APAPALAAAPATVAVLRGVRMSVLRVADLAAAEEEQIDQERALQRHARAARAKNGNDDG
VQAKLLLYLASESSLRRRYFGELEVRVFSEPAALGGGAAAAAYTLMANGRMHFLRARRLLRRLRQAAAA
RRSRLRQGEREDQDDDDDDGGDSDGDETEVLEVDLRARRTKLAAARAAEQAQWRPGLRLDQVSAVLRALFDHLH
YPNPDEKLRVLATGTSRQISNFWINARVLRKCPMIEEMYKEDFSGASAASVYDASAGSAS
```

Below are partial the BELL cDNA sequences identified in wheat and Sorghum, which were derived from EST sequences from the TIGR Gene Indices.

**TaBEll TIGR Triticum Gene Index**

```
MATYYSSQGEREDSMDPVGAGYYSMPSICALNLLYNNPSSGGYTEFSGILQTQCNFMEMPPGHHAMS
QDSSARESQDMLASHHQRSAVFGHDLQMDKNGMMDGQDDDLVQGVDAGDKDAPTEGT1QRQOIQMKAKLVMLEDEVQEYRHYH
HMQSFVQHLEQAAGIAGSAXTALALATSQKFRCLDAIQIGRSATQGEGESLSKLGEQFVNDQ
LRRQALGQGMTHQNAWRFQGRLPERASVLRWFELHLHPYKSDKMMLAKQTGLTRSVQSNFWINAR
LRKFWMVEEMEQLGKEQSEDSSTRTSRSITATQDKFSGPSQSNKQSKNDHRQAPSMFVQHRHEFTIGMRNSQAGENLGSIESINSITQGGPSKPKRNMLHSNPSINSNMDVKEEQM
SMFQGSRQDFRDGLMGFMGFNGFGAYGIFIEARSFTEQPSAYSTGVSSTLIFLGHNENLMSATHSH
```

Below are partial the BELL cDNA sequences identified in wheat and Sorghum, which were derived from EST sequences from the TIGR Gene Indices.
**APPENDIX 1.3 SEQUENCE ALIGNMENT OF THE BELL FAMILY**

Multiple amino acid alignment of the BELL proteins using Pileup (GCG). Features of the sequence are as indicated: identical residues are shaded black and residues conserved in at least three sequences are grey; the SKY domain is underscored by the green box; the striped box underlines the BELL domain; the homeodomain motif is underlined by the grey box; conserved TALE residues (PYP) are indicated above the homeodomain sequence by the black arrows. The conserved QGLSLSL and VSLTLGL boxes are highlighted by the orange boxes. *Arabidopsis thaliana* At; *Gnetum gemon* Gg; *Hordeum vulgare* Hv; *Lycopersicon esculentum* Lc; *Malus domestica* Md; *Oryza sativa* Os; *Sorghum bicolor* Sb; *Solanum tuberosum* St; *Triticum aestivum* Ta; *Zea mays* Zm.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtATH1</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>LeBL4</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L14</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L9</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L4</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L3</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>HvUBEL2</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L2</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>ZmKIP</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>AtPNY</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>AtBLH6</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>AtBLH7</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L8</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>SbBE1L</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>TaBE1L</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1H1</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L5</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>StBE1L30</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>AtBLH3</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>AtBLH10</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>AtBLH1</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>LeBL1</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L10</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L3</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>LeBL2</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>StBE15</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>StBE11</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>StBE129</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>AtBLH1</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>HvUBEL1</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L11</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>AtBLH2</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>AtBLH4</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>MdMH1</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>StBE13</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>StBE122</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L12</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L7</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L6</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>AtBE1L</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>LeBL3</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>OsBE1L</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>AtBLH5</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>GmE1BEl1</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
<tr>
<td>PNF</td>
<td>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</td>
</tr>
</tbody>
</table>
AtATH1
LeBL4
OsBEL14
OsBEL9
OsBEL4
OsBEL13
HvJUBEL2
OsBEL2
ZmKIP
AtPNY
AtBLH6
AtBLH7
OsBEL8
StBEL7
TaBEL
OsBIHD1
OsBEL5
StBEL30
AtBLH3
AtBLH10
AtBL1
OsBEL10
OsBEL3
LeBL2
StBEL15
StBEL11
StBEL29
AtBLH1
HvJUBEL1
OsBEL11
AtBLH2
AtBLH4
MdMH1
StBEL13
StBEL22
OsBEL12
OsBEL7
AtBEL1
LeBL3
OsBEL
AtBLH5
GmMEBEL1
PNF

AGRELVAKAAPILPPRRAPHAPAPJGGADDYAVGSVPRAAKGGRGLKLGGLGDDAGLAEEGGRRATITAEESAVNRTAAGTRLAISTSSNFHKAASA

GGFHQGSSG.VYGFSSDGFD.RPFGSQL.QQH. QHEDHVAQQSSRDKLRVQGFDPA..GLLPIDGDQH.VEPGAMYD..AAAASNMLAEMFN

GGFHQPSGAVFSGSSDGFDGRFGSCQDHQQHEQQQHQVAQQSSRDKLRVQGFDPAAAAGHGLPLTEGDEHGAEPGMYDHAAEEAAAGSNMLSEMNF

MGITKTSPNTTLLLKTHNMSQDYHHHHHNNHHQGGGIFNSGFDRSDDLNNLTTQQKEQHQRVEMEDEESSVAGGRKIPVYESAGLMSEMNF

MLATTSSMSQDYHHQ......GIFSNSGFRRSSS...TTHQE......EVDESAVVQGQIPYETAGMLSEMFA

HGQFETFSGFRSMARTHQEQQQQQHHIALQIRRALKLRVQGFETPPPLVGLNEESSGLPAYETAGMLSEMNYPPGGPVGAAELLHEHPSANYR

TSYETAGLLLSEMFSNTTS...TAATELLQNLQNYYR

MFNFSAAEILLSFQSKNLSQSSASSEDAVSCRPAAGGFTSFQHTVSQDSVSVNTS6WKNYSAQGSEEWFGVILNSVGYEUGQR QSLSLATPLMLGGSVKE

QGFDPAAAAAGHGLLPIEGDEHGAEPGAMYDHAAEEAAAGSNMLSEMNF
<table>
<thead>
<tr>
<th>protein</th>
<th>Accession</th>
<th>Start</th>
<th>Stop</th>
<th>Name</th>
<th>Description</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtATH1</td>
<td>A0A167833200</td>
<td>398</td>
<td>473</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>LeBL4</td>
<td>A0A167833400</td>
<td>251</td>
<td>323</td>
<td>ARS</td>
<td>Lotus japonicus</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL14</td>
<td>A0A167833900</td>
<td>475</td>
<td>566</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL9</td>
<td>A0A167834200</td>
<td>469</td>
<td>538</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL4</td>
<td>A0A167833500</td>
<td>396</td>
<td>472</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL13</td>
<td>A0A167833600</td>
<td>387</td>
<td>472</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>HvJUBEL2</td>
<td>A0A167833700</td>
<td>370</td>
<td>463</td>
<td>ARS</td>
<td>Hordeum vulgare</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL2</td>
<td>A0A167833800</td>
<td>285</td>
<td>346</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>ZmKIP</td>
<td>A0A167834000</td>
<td>293</td>
<td>359</td>
<td>ARS</td>
<td>Zea mays</td>
<td>ARS</td>
</tr>
<tr>
<td>AtPNY</td>
<td>A0A167834100</td>
<td>374</td>
<td>467</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>StBEL30</td>
<td>A0A167834200</td>
<td>417</td>
<td>508</td>
<td>ARS</td>
<td>Solanum tuberosum</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBLH6</td>
<td>A0A167834300</td>
<td>340</td>
<td>429</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL8</td>
<td>A0A167834400</td>
<td>413</td>
<td>504</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>SbBELL</td>
<td>A0A167834500</td>
<td>414</td>
<td>451</td>
<td>ARS</td>
<td>Silene hispanica</td>
<td>ARS</td>
</tr>
<tr>
<td>TaBELL</td>
<td>A0A167834600</td>
<td>415</td>
<td>506</td>
<td>ARS</td>
<td>Taeniopyrum mongolicum</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL10</td>
<td>A0A167834700</td>
<td>136</td>
<td>228</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL3</td>
<td>A0A167834800</td>
<td>379</td>
<td>468</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>LeBL2</td>
<td>A0A167834900</td>
<td>430</td>
<td>515</td>
<td>ARS</td>
<td>Lotus japonicus</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBLH11</td>
<td>A0A167835000</td>
<td>229</td>
<td>350</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBIHD1</td>
<td>A0A167835100</td>
<td>407</td>
<td>500</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>StBEL29</td>
<td>A0A167835200</td>
<td>262</td>
<td>353</td>
<td>ARS</td>
<td>Solanum tuberosum</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBLH1</td>
<td>A0A167835300</td>
<td>411</td>
<td>501</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL11</td>
<td>A0A167835400</td>
<td>350</td>
<td>442</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL7</td>
<td>A0A167835500</td>
<td>224</td>
<td>316</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>MdMH1</td>
<td>A0A167835600</td>
<td>582</td>
<td>675</td>
<td>ARS</td>
<td>Malus domestica</td>
<td>ARS</td>
</tr>
<tr>
<td>StBEL13</td>
<td>A0A167835700</td>
<td>367</td>
<td>460</td>
<td>ARS</td>
<td>Solanum tuberosum</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBELL1</td>
<td>A0A167835800</td>
<td>416</td>
<td>494</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>LeBL3</td>
<td>A0A167835900</td>
<td>398</td>
<td>482</td>
<td>ARS</td>
<td>Lotus japonicus</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL12</td>
<td>A0A167836000</td>
<td>350</td>
<td>439</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>StBEL11</td>
<td>A0A167836100</td>
<td>267</td>
<td>350</td>
<td>ARS</td>
<td>Solanum tuberosum</td>
<td>ARS</td>
</tr>
<tr>
<td>StBEL13</td>
<td>A0A167836200</td>
<td>350</td>
<td>439</td>
<td>ARS</td>
<td>Solanum tuberosum</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBLH5</td>
<td>A0A167836300</td>
<td>254</td>
<td>347</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>HvJUBEL1</td>
<td>A0A167836400</td>
<td>559</td>
<td>652</td>
<td>ARS</td>
<td>Hordeum vulgare</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL11</td>
<td>A0A167836500</td>
<td>605</td>
<td>705</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBLH2</td>
<td>A0A167836600</td>
<td>524</td>
<td>617</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBLH4</td>
<td>A0A167836700</td>
<td>540</td>
<td>631</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>MdMH1</td>
<td>A0A167836800</td>
<td>582</td>
<td>675</td>
<td>ARS</td>
<td>Malus domestica</td>
<td>ARS</td>
</tr>
<tr>
<td>StBEL13</td>
<td>A0A167836900</td>
<td>367</td>
<td>460</td>
<td>ARS</td>
<td>Solanum tuberosum</td>
<td>ARS</td>
</tr>
<tr>
<td>StBEL12</td>
<td>A0A167837000</td>
<td>432</td>
<td>460</td>
<td>ARS</td>
<td>Solanum tuberosum</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL12</td>
<td>A0A167837100</td>
<td>350</td>
<td>439</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBELL7</td>
<td>A0A167837200</td>
<td>224</td>
<td>316</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL6</td>
<td>A0A167837300</td>
<td>350</td>
<td>439</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBELL6</td>
<td>A0A167837400</td>
<td>416</td>
<td>494</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>LeBL1</td>
<td>A0A167837500</td>
<td>326</td>
<td>393</td>
<td>ARS</td>
<td>Lotus japonicus</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL10</td>
<td>A0A167837600</td>
<td>135</td>
<td>228</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL3</td>
<td>A0A167837700</td>
<td>379</td>
<td>468</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>LeBL2</td>
<td>A0A167837800</td>
<td>430</td>
<td>515</td>
<td>ARS</td>
<td>Lotus japonicus</td>
<td>ARS</td>
</tr>
<tr>
<td>StBEL15</td>
<td>A0A167837900</td>
<td>418</td>
<td>503</td>
<td>ARS</td>
<td>Solanum tuberosum</td>
<td>ARS</td>
</tr>
<tr>
<td>StBEL11</td>
<td>A0A167838000</td>
<td>267</td>
<td>350</td>
<td>ARS</td>
<td>Solanum tuberosum</td>
<td>ARS</td>
</tr>
<tr>
<td>StBEL29</td>
<td>A0A167838100</td>
<td>262</td>
<td>353</td>
<td>ARS</td>
<td>Solanum tuberosum</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBLH1</td>
<td>A0A167838200</td>
<td>411</td>
<td>500</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>HvJUBEL1</td>
<td>A0A167838300</td>
<td>559</td>
<td>652</td>
<td>ARS</td>
<td>Hordeum vulgare</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL11</td>
<td>A0A167838400</td>
<td>605</td>
<td>705</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBLH2</td>
<td>A0A167838500</td>
<td>524</td>
<td>617</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBLH4</td>
<td>A0A167838600</td>
<td>540</td>
<td>631</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>MdMH1</td>
<td>A0A167838700</td>
<td>582</td>
<td>675</td>
<td>ARS</td>
<td>Malus domestica</td>
<td>ARS</td>
</tr>
<tr>
<td>StBEL13</td>
<td>A0A167838800</td>
<td>367</td>
<td>460</td>
<td>ARS</td>
<td>Solanum tuberosum</td>
<td>ARS</td>
</tr>
<tr>
<td>StBEL12</td>
<td>A0A167838900</td>
<td>432</td>
<td>460</td>
<td>ARS</td>
<td>Solanum tuberosum</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL12</td>
<td>A0A167839000</td>
<td>350</td>
<td>439</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL6</td>
<td>A0A167839100</td>
<td>350</td>
<td>439</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBELL6</td>
<td>A0A167839200</td>
<td>416</td>
<td>494</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>LeBL3</td>
<td>A0A167839300</td>
<td>398</td>
<td>482</td>
<td>ARS</td>
<td>Lotus japonicus</td>
<td>ARS</td>
</tr>
<tr>
<td>OsBELL</td>
<td>A0A167839400</td>
<td>320</td>
<td>407</td>
<td>ARS</td>
<td>Oryza sativa</td>
<td>ARS</td>
</tr>
<tr>
<td>AtBLH5</td>
<td>A0A167839500</td>
<td>254</td>
<td>347</td>
<td>ARS</td>
<td>Arabidopsis thaliana</td>
<td>ARS</td>
</tr>
<tr>
<td>GgMELBEL1</td>
<td>A0A167839600</td>
<td>612</td>
<td>705</td>
<td>ARS</td>
<td>Glycine max</td>
<td>ARS</td>
</tr>
<tr>
<td>PNF</td>
<td>A0A167839700</td>
<td>450</td>
<td>543</td>
<td>ARS</td>
<td>Phaseolus vulgaris</td>
<td>ARS</td>
</tr>
<tr>
<td>Species</td>
<td>Accession</td>
<td>Sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtATH1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LeBL4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL4</td>
<td>562</td>
<td>AEPPLAALPLNLNHRFGLED.VSDAVYMSF.GGQDRHFKGIEGHHLHDVFV...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL13</td>
<td>553</td>
<td>AEPPLAALPLNLNHRFGLED.VSDAVYMSF.GGQDRHFTKEIGHHLHDVFV...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HvJUBEL2</td>
<td>563</td>
<td>GEPLPA. .NLNHRFGLEDVSDPYMSFGQDQQRHFAKEIGHHLHDVFV...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZmKIP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtPNY</td>
<td>568</td>
<td>QFLHDFV...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtBLH6</td>
<td>530</td>
<td>FVA--...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtBLH7</td>
<td>480</td>
<td>FVA--...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL8</td>
<td>603</td>
<td>VNEDDDIYNTAPLGVTA.SSDYDSNMQDRQRFEHSPILLHDFV--...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SbBELL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TaBELL</td>
<td>606</td>
<td>VSDDDIYNYGAPLGVSA.SSDYESLNMQDRQRFEHSPILLHDFV--...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBHH1D</td>
<td>597</td>
<td>GAGEEYNSTAGVAASDDYESTNQIDQRQRFEHSPILLHDFV--...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL5</td>
<td>832</td>
<td>[LT]GHCNANGVPAEQQQLYYGSAGDDYINGSDDRQRFGPAQSSLHDFV...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>StBEL30</td>
<td>604</td>
<td>VPLRADDMYSTAPTTMVNLPETEFNCELDSQNRQQPFWLPSAT--...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtBLH3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtBLH10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtBEB1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL10</td>
<td>328</td>
<td>DHHNHEHGGV.GYSLFPAPYG.QFATEQFPAFHGGGGGGGVSILTGLH...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL3</td>
<td>555</td>
<td>DDDDDVPGGGYSGLHTAPAQQHGFQGSDHFPAFH.GGGGGTPSILTGLH...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LeB2</td>
<td>601</td>
<td>.LMAAYMGD.FG.RDFPHQOQ...TANF. .H.GNGVSILTGLHPSLEN...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>StBEL15</td>
<td>590</td>
<td>.LMAAYMGD.FG.RDFPHQOQ...TANF. .H.GNGVSILTGLHPSLEN...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>StBEL11</td>
<td>444</td>
<td>LMTPNYMDDQGTFRNQNHQLATTIT...FCQHNGVSILTGLHPSN...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>StBEL29</td>
<td>452</td>
<td>FSLQNMGNMFFGQAYPEIGARFSDEQPSAPYSTGSSTLV...HLEN...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtBLH11</td>
<td>601</td>
<td>PHCDLS1STDHQQFMTHGIGPRVRKIGETEYFPATINGGSSTTTHAAAAAYNGMQRQKRYVAQQLPFDV...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HvJUBEL1</td>
<td>751</td>
<td>LRQYNCG--...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtBLH2</td>
<td>711</td>
<td>DVSLILGHAGNMPDKDASFCVFREFGGF...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtBLH4</td>
<td>599</td>
<td>DVSLILGHSGNIPDKNTSFVRFDPGD...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MdMH1</td>
<td>775</td>
<td>DVSLILGHAGGGGGNNMPDKTSSSF5IRDGGC...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>StBEL13</td>
<td>546</td>
<td>DVSLILGHAGNLPFENTHFG...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>StBEL12</td>
<td>599</td>
<td>DVSLILGHAGNLPFENTHFG...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL12</td>
<td>540</td>
<td>EMAHPADMPHGGHVGQVFGQAMDGAASHAQEMNI1I1AKMFPKITYPAVQVESGKGALFCMKQGQDRHCKDavnHAAGAKPKCSSFLCSL...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL7</td>
<td>414</td>
<td>EMAHPADMPHGGHVGQVFGQAMDGAASHAQEMNYRSLGPHILLDLAG...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBELL6</td>
<td>539</td>
<td>LPMGQGEEHQ...QLQPTQAQQGAMQDASHQYRLS.ATAAQPHLLLNL...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtBLH1</td>
<td>577</td>
<td>EGMQVYSAEMDDLDDQYQNLQPNLMGAQLHLDIV...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LeB3</td>
<td>489</td>
<td>DHDQDQVYSSLLDSENQNLYRNLMGAQLHLDLAG...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OsBEL</td>
<td>494</td>
<td>QQQQGFMVEAAREEDDVLQPNLMESQOLLHDFAGAS...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtBLH5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GmMEB11</td>
<td>806</td>
<td>IDFSSYMGGRISHESNPRPTGNASVSLILGHIHSGAESKYTGLYLPREDTQCNRSYETHEIHDGHQACVQGFGETHDIOFRKHLITQDLLQ...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
OsBELL10 425 A~A----------------------------------------------- 425
OsBELL3  654 A~A---------------------------------------------------------------------------------- 654
LeBIL2   678 QSGNKRFPTQLLPDFVTGNLGT------------------------------------------ 699
StBIL5   667 QSGNKRFPTQLLPDFVTGNLGT------------------------------------------ 688
StBIL11  520 QSG.KRYATQLQDFVS----------------------------------------------- 535
StBIL29  552 QN.RKRFAAPLLPDFVA----------------------------------------------- 567
OsBIL12  640 LLPVLALYVVALAVSPFYSGSCPEELASGDVAHLAAAGDAGNRRNDSPPDDAAPTGLGHIVF
GIAASSELWKSREYIRTWRPEQMGFVWLDKP 739
OsBIL12  740 VYEFSRNASTGLPGIKISGNTTKFPYTHGGRSRSALRITIVSESFRLG
PGARWFVMDDDTVFFPDNLVDVLSDYDHTQFYIYNPSESIQNLIFS 839
OsBIL12  840 YGMAFGGGGFASRALAAQLAHMQDGCIDRYPALYGSDRRHACVAGLGP
LTPHGLGHQCDLMGDLGLAGPVVPVTLLHDLFQVFPTTTRSRTA 939
OsBIL12  940 ALRRLEFGPARLDSAGVAQQSVCYDGD KKQTVS VSWGFAV
VTGVS PREMEMPRTRFLWNYYRADTAYAFNTRPVARQPCQPQVYMRQSLDRRR 1039
OsBIL12 1040 NTTVEYERRRVAVPCKGWRIDPDAALLDRVIVLKKPDNLWKRSPRRNCCRVLSSPRQ
GDRKMTIDVGCRGGEFARIEVAR------------- 1123
## APPENDIX 1.4  ANALYSIS OF THE GENOMIC STRUCTURE OF THE **ARABIDOPSIS** BELL FAMILY AND FLANKING GENES

TAIR identification numbers and annotation of *Arabidopsis* BELL genes (highlighted) and flanking genes. Predicted amino acid sequences were aligned with proteins flanking other BELL genes using the pairwise BLAST 2 Sequences tool (Tatusova and Madden, 1999), which filters regions of low complexity and is useful for identifying homologous protein domains. If the BLAST alignment detected any similarities between the proteins the percentage amino acid identity is shown.

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Description</th>
<th>BLAST search</th>
</tr>
</thead>
<tbody>
<tr>
<td>At1g19630</td>
<td>cytochrome p450 family</td>
<td>63% At1g75370 44% At4g34580</td>
</tr>
<tr>
<td>At1g19640</td>
<td>putative jasmonic acid Mhyl transferase</td>
<td>81% At1g75380</td>
</tr>
<tr>
<td>At1g19650</td>
<td>SEC14 factor family</td>
<td>63% At1g75370 47% At4g34580</td>
</tr>
<tr>
<td>At1g19660</td>
<td>similar to wound inductive protein</td>
<td>81% At1g75380</td>
</tr>
<tr>
<td>At1g19670</td>
<td>coronatine induced protein</td>
<td>53% At1g75400</td>
</tr>
<tr>
<td>At1g19680</td>
<td>expressed protein</td>
<td>68aa shared At1g19680</td>
</tr>
<tr>
<td>At1g19690</td>
<td>expressed protein</td>
<td>82% At1g75420</td>
</tr>
<tr>
<td>At1g19700</td>
<td>BLH10</td>
<td></td>
</tr>
<tr>
<td>At1g19710</td>
<td>glycosyl transferase family</td>
<td>82% At1g75420</td>
</tr>
<tr>
<td>At1g19715</td>
<td>jacalin lectin family</td>
<td></td>
</tr>
<tr>
<td>At1g19720</td>
<td>penta tricopeptide repeat protein</td>
<td></td>
</tr>
<tr>
<td>At1g19730</td>
<td>theoreodoxin-like protein</td>
<td></td>
</tr>
<tr>
<td>At1g19740</td>
<td>LON domain ATP-dependent protease</td>
<td></td>
</tr>
<tr>
<td>At1g19750</td>
<td>WD-repeat family</td>
<td>25% At2g16405</td>
</tr>
<tr>
<td>At1g19770</td>
<td>purine transporter family</td>
<td>38% At1g75470</td>
</tr>
<tr>
<td>At1g19780</td>
<td>gated channel family</td>
<td></td>
</tr>
<tr>
<td>At1g75340</td>
<td>ZINC finger family</td>
<td></td>
</tr>
<tr>
<td>At1g75350</td>
<td>L31 ribosomal protein family</td>
<td></td>
</tr>
<tr>
<td>At1g75360</td>
<td>expressed protein</td>
<td></td>
</tr>
<tr>
<td>At1g75370</td>
<td>SEC14 factor family</td>
<td>63% At1g19650 44% At2g16380</td>
</tr>
<tr>
<td>At1g75380</td>
<td>similar to wound inductive protein</td>
<td>81% At1g19660</td>
</tr>
<tr>
<td>At1g75390</td>
<td>bZIP family</td>
<td>53% At4g34590</td>
</tr>
<tr>
<td>At1g75400</td>
<td>expressed protein</td>
<td>53% At1g19680</td>
</tr>
<tr>
<td>At1g75410</td>
<td>BLH3</td>
<td></td>
</tr>
<tr>
<td>At1g75420</td>
<td>glycosyl transferase family</td>
<td>82% At1g19710</td>
</tr>
<tr>
<td>At1g75430</td>
<td>BLH11</td>
<td></td>
</tr>
<tr>
<td>At1g75440</td>
<td>ubiquitin conjugating enzyme</td>
<td></td>
</tr>
<tr>
<td>At1g75450</td>
<td>FAD domain cytokinin oxidase</td>
<td></td>
</tr>
<tr>
<td>At1g75460</td>
<td>LON domain ATP-dependent protease</td>
<td></td>
</tr>
<tr>
<td>At1g75470</td>
<td>purine transporter family</td>
<td>38% At1g19770</td>
</tr>
<tr>
<td>At1g75480</td>
<td>Hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>At2g16370</td>
<td>THY-1</td>
<td>87% At4g34570</td>
</tr>
<tr>
<td>At2g16380</td>
<td>SEC14 factor family</td>
<td>44% At1g75370</td>
</tr>
<tr>
<td>At2g16385</td>
<td>expressed protein</td>
<td>60% At4g34600</td>
</tr>
<tr>
<td>At2g16390</td>
<td>helicase domain protein</td>
<td></td>
</tr>
<tr>
<td>At2g16400</td>
<td>BLH7</td>
<td></td>
</tr>
<tr>
<td>At2g16405</td>
<td>WD repeat protein</td>
<td>25% At1g19750</td>
</tr>
<tr>
<td>At2g16410</td>
<td>hypothetical zinc finger protein</td>
<td></td>
</tr>
<tr>
<td>At2g16420</td>
<td>retrotransposon family</td>
<td></td>
</tr>
<tr>
<td>At2g16430</td>
<td>purple acid phosphatase</td>
<td>66% At2g27190</td>
</tr>
<tr>
<td>At2g23730</td>
<td>expressed protein</td>
<td></td>
</tr>
<tr>
<td>At2g23740</td>
<td>ZINC finger family</td>
<td></td>
</tr>
<tr>
<td>At2g23750</td>
<td>SET domain protein</td>
<td></td>
</tr>
<tr>
<td>At2g23755</td>
<td>expressed protein</td>
<td></td>
</tr>
<tr>
<td>At2g23760</td>
<td>BLH4</td>
<td></td>
</tr>
<tr>
<td>At2g23770</td>
<td>protein kinase family</td>
<td></td>
</tr>
<tr>
<td>At2g23780</td>
<td>RING (ZINC) finger family</td>
<td></td>
</tr>
<tr>
<td>Gene ID</td>
<td>Description</td>
<td>Similarity %</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>At2g23790</td>
<td>expressed protein</td>
<td>66%</td>
</tr>
<tr>
<td>At2g27190</td>
<td>purple acid phosphatase family</td>
<td>68%</td>
</tr>
<tr>
<td>At2g27200</td>
<td>GTP binding family</td>
<td></td>
</tr>
<tr>
<td>At2g27210</td>
<td>Kelch repeat protein</td>
<td></td>
</tr>
<tr>
<td>At2g27220</td>
<td>BLH5</td>
<td></td>
</tr>
<tr>
<td>At2g27230</td>
<td>bHLH transcription factor</td>
<td></td>
</tr>
<tr>
<td>At2g27240</td>
<td>expressed protein</td>
<td></td>
</tr>
<tr>
<td>At2g27250</td>
<td>putative CLV1 ligand</td>
<td></td>
</tr>
<tr>
<td>At2g27970</td>
<td>cyclin dependent kinase</td>
<td></td>
</tr>
<tr>
<td>At2g27980</td>
<td>PHD finger-like protein</td>
<td></td>
</tr>
<tr>
<td>At2g27990</td>
<td>PNF</td>
<td></td>
</tr>
<tr>
<td>At2g28000</td>
<td>RUBISCO subunit binding</td>
<td></td>
</tr>
<tr>
<td>At2g28010</td>
<td>aspartyl protease family</td>
<td></td>
</tr>
<tr>
<td>At2g35910</td>
<td>RING finger</td>
<td></td>
</tr>
<tr>
<td>At2g35920</td>
<td>helicase domain protein</td>
<td></td>
</tr>
<tr>
<td>At2g35930</td>
<td>U-box domain protein</td>
<td></td>
</tr>
<tr>
<td>At2g35940</td>
<td>BLH1</td>
<td></td>
</tr>
<tr>
<td>At2g35950</td>
<td>expressed protein</td>
<td></td>
</tr>
<tr>
<td>At2g35960</td>
<td>HIN family protein</td>
<td></td>
</tr>
<tr>
<td>At4g32950</td>
<td>putative protein phosphatase</td>
<td>68%</td>
</tr>
<tr>
<td>At4g32960</td>
<td>expressed protein</td>
<td>68%</td>
</tr>
<tr>
<td>At4g32970</td>
<td>expressed protein</td>
<td></td>
</tr>
<tr>
<td>At4g32980</td>
<td>ATH1</td>
<td></td>
</tr>
<tr>
<td>At4g32990</td>
<td>WD repeat protein</td>
<td></td>
</tr>
<tr>
<td>At4g33000</td>
<td>calcineurin-B like protein</td>
<td></td>
</tr>
<tr>
<td>At4g33010</td>
<td>Putative glycine dehydrogenase</td>
<td></td>
</tr>
<tr>
<td>At4g34530</td>
<td>bHLH protein</td>
<td></td>
</tr>
<tr>
<td>At4g34540</td>
<td>isoflavone reductase family</td>
<td></td>
</tr>
<tr>
<td>At4g34560</td>
<td>expressed protein</td>
<td></td>
</tr>
<tr>
<td>At4g34570</td>
<td>THY-2</td>
<td>87%</td>
</tr>
<tr>
<td>At4g34580</td>
<td>phosphoglyceride transfer protein</td>
<td>47%</td>
</tr>
<tr>
<td>At4g34590</td>
<td>bZIP family</td>
<td>53%</td>
</tr>
<tr>
<td>At4g34600</td>
<td>expressed protein</td>
<td>60%</td>
</tr>
<tr>
<td>At4g34610</td>
<td>BLH6</td>
<td></td>
</tr>
<tr>
<td>At4g34620</td>
<td>ribosomal protein S16</td>
<td></td>
</tr>
<tr>
<td>At4g34630</td>
<td>expressed protein</td>
<td></td>
</tr>
<tr>
<td>At4g34640</td>
<td>squaline synthase</td>
<td></td>
</tr>
<tr>
<td>At4g36820</td>
<td>hypothetical protein</td>
<td>66%</td>
</tr>
<tr>
<td>At4g36830</td>
<td>membrane protein</td>
<td></td>
</tr>
<tr>
<td>At4g36840</td>
<td>Kelch repeat protein</td>
<td></td>
</tr>
<tr>
<td>At4g36850</td>
<td>PQ-Loop repeat protein</td>
<td></td>
</tr>
<tr>
<td>At4g36860</td>
<td>LIM domain protein</td>
<td></td>
</tr>
<tr>
<td>At4g36870</td>
<td>BLH2</td>
<td></td>
</tr>
<tr>
<td>At4g36880</td>
<td>cysteine protease</td>
<td></td>
</tr>
<tr>
<td>At4g36890</td>
<td>glycosyl tranferase family</td>
<td></td>
</tr>
<tr>
<td>At4g36900</td>
<td>AP2 domain containing protein</td>
<td></td>
</tr>
<tr>
<td>At4g36910</td>
<td>CBS domain protein</td>
<td></td>
</tr>
<tr>
<td>At5g02000</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>At5g02010</td>
<td>expressed protein</td>
<td></td>
</tr>
<tr>
<td>At5g02020</td>
<td>expressed protein</td>
<td></td>
</tr>
<tr>
<td>At5g02030</td>
<td>PNY</td>
<td></td>
</tr>
<tr>
<td>At5g02040</td>
<td>PRA1 family</td>
<td></td>
</tr>
<tr>
<td>At5g02050</td>
<td>MAM33 family</td>
<td></td>
</tr>
<tr>
<td>At5g02060</td>
<td>membrane protein</td>
<td></td>
</tr>
<tr>
<td>Gene ID</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------</td>
<td></td>
</tr>
<tr>
<td>At5g41370</td>
<td>XPB1 DNA repair protein</td>
<td></td>
</tr>
<tr>
<td>At5g41380</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>At5g41390</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>At5g41400</td>
<td>RING finger family</td>
<td></td>
</tr>
<tr>
<td>At5g41410</td>
<td>BELL1</td>
<td></td>
</tr>
<tr>
<td>At5g41430</td>
<td>RING finger family</td>
<td></td>
</tr>
<tr>
<td>At5g41440</td>
<td>RING finger family</td>
<td></td>
</tr>
<tr>
<td>At5g41450</td>
<td>RING finger family</td>
<td></td>
</tr>
<tr>
<td>At5g41460</td>
<td>fringe-related protein</td>
<td></td>
</tr>
<tr>
<td>At5g41470</td>
<td>hypothetical protein</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX 1.5 The BLH3 cDNA SEQUENCE AND TRANSLATION

The BLH3 cDNA (AY598452) was sequenced and contiguous sequence was assembled (Carly Pullen, this laboratory). The protein translation was carried out using the Translate program (GCG). The ATG start codon is highlighted with the grey box, the predicted protein sequence is in bold type.

```
1  ccccttttatctctctgttccctgattcagccacacatccacaatctcttttctttttgt
  P  F  L  S  L  C  S  L  I  Q  P  H  P  Q  S  L  F  L  F  C
61  atagtttatcatctagcttttctcagacaaacatcaacggtgaaactcacttaaac
  I  V  Y  H  L  A  L  S  Q  T  K  H  Q  S  D  E  T  H  L  N
121  aacaaacacactttcagatttatattcactttgctggtgattctaccctaataagtgtcggcatg
  N  K  T  L  S  I  Y  Y  P  M  A  V  Y  P  N  S  V  G  M
181  atctcttttaccaagaaatcactttacccgaaacacaaacaacacaacacaacacaacaga
  Q  S  L  Y  Q  E  S  I  Y  L  N  E  Q  Q  Q  Q  Q  Q  Q  A
241  tcttttcctcgctcatcttttctcagagattttctctcggtgatgttcgaacaacacag
  S  S  S  A  A  S  F  S  E  I  V  S  G  D  V  R  N  N  E
301  atggtattatatccccaaacacagcagctgacgtccgctcaacggaacgttaacggtgtaaqt
  M  V  F  I  P  P  T  S  D  V  A  V  N  G  N  V  T  V  S
361  aacgatctaaacctttcaacggtgagacctttattaaggtcttgataacagatccaagtca
  N  D  L  S  F  G  G  G  S  L  S  L  G  N  Q  I  Q  S
421  gctgtctctgtctccgagattttctcttgagtttctcagagatttacac
  A  V  S  V  S  F  Q  Y  H  Y  Q  N  L  S  N  Q  L  S  Y
481  aataatctaatctttcactatgtctgtgagatagaaatctctgagttttcatcag
  N  N  L  N  P  S  T  M  S  D  E  N  G  K  S  L  S  V  H  Q
541  catcactctgtatcaaaattttacctttctcagagattttctcttgagttttcatcag
  H  H  S  D  Q  I  L  P  S  S  V  Y  N  N  G  N  G  N  G  V
601  ggatttcaacaaatactacctttacggtgagacatccgaggtttgtgtagtagtactgatactgatct
  G  F  Y  N  N  Y  E  R  T  S  G  F  V  S  S  V  V  L  R  S
661  cgtacctcaacaaactttacggtgagacatccgaggtttgtgtagtagtactgatactgatct
  R  Y  L  K  P  T  Q  Q  L  L  D  E  V  V  S  V  R  K  D  L
721  aaattgggaatagaataagagatgaagatatgataaaaggttcaacaacttttcacaatggtctagt
  K  L  G  N  K  M  K  N  D  K  G  Q  D  F  H  N  G  S
781  gataacattaacgagatgataaatatctacactcgaggaggtgtgactcttcctcagaacgcgtag
  D  N  I  T  E  D  D  K  S  Q  S  E  L  S  P  S  E  R  Q
841  gagctacagagcaagaagacgaagcattttcataaatgtggagtgtagtagataaaaggtat
  E  L  Q  S  K  S  K  L  L  T  M  V  D  E  V  D  K  R  Y
901  aaccaataccatcatcaaataagcagtgttctagctccttccgagagtggattacagtcttt
  N  Q  Y  H  H  Q  M  E  A  L  L  S  S  F  E  M  V  T  G  L
961  ggagcagctaatagccctatacaacatccgtagctctgaataaatctctctcgcataactctcgt
  G  A  A  K  P  Y  T  S  V  A  L  N  R  I  S  R  H  F  R  C
1021  ttacgcgacgcggataaaacaagagattcaggtgatcagagggagctttggagagagag
  L  R  D  A  I  K  E  Q  I  Q  V  I  R  G  K  L  G  E  R  E
```
1081  acttctgatgaacagagagagagatacccgctttaggtacttagatcaacggttgaga
     T S D E Q G E R I P R L R Y L D Q R L R
1141  caacagagagctttgcatcaacaacttgggaatggttagaccagctttggagaccaacaaga
     Q Q R A L H Q Q L G M V R P A W R P Q R
1201  ggctttacctgaaaactctgtctctatacttgcagctttgacatttctctcat
     G L P E N S V I L R A W L F E H F L H
1261  ccatatctcataagaatcagagaaatcatgttcttcacaagcagacagactatcgaagaac
     P Y P K E S K I M L S K Q T G L S K N
1321  caggttgcaaatgttttataacgcagagctctgcagactatggaacaactagttggaagag
     Q V A N W F I N A R V R L W K P M I E E
1381  atgtataaaaagaagagtttggaatcacgcagagttacctctcataactctaatacaagacacc
     M Y K E E F G E S A E L L S N S N Q D T
1441  aaaaaatgcaggaacacatctcagctcaaacacgagacctctctctctgcaacaacaag
     K K M Q E T S Q L K H E D S S S S S Q Q Q
1501  aatcaggaaaaaccaacacaacaacataccatatatactctgatgcagacacaacacagtct
     N Q G N N N N N N I P Y T S D A E Q N L V
1561  tttcagatccataacacgacccgtgtacactggagttacgacagctttgatgaaactat
     F A D P K P D R A T T G D Y D S L M N Y
1621  catgggttggagtattgagttacacggttttgcgttggttcggtaagaccaacacaagaatggc
     H G F G I D D Y N R Y V G L G N Q Q D G
1681  agatattctaatcccaatatacagcattttggtttgtctaaagattctctgagcttttc
     R Y S N P H Q L H D F V V * K I P E L F
1741  acaagcttaaagactggcaatggttgattgggtgagtatttctttgaagtcgaagcaga
     T S L R T G N A S * L G E Y D S * S A R
1801  aatgcgctctttgcatatcattatagagggattttataaaatatgtatatcttttatatt
     K C A L G I F I D R G Y T N M Y I C N I
1861  tgttgtgctatatataaaggtagattttgtgaaacttttggagtttagatctttactatt
     C C G I Y K G * I L * K L * V L D F T F
1921  gtatattaaagattagcttttggtac
     V Y E N * A L C
APPENDIX 1.6  THE ASSEMBLY AND TRANSLATION OF THE BLH10 cDNA SEQUENCE

The BLH10 cDNA was sequenced using four primers to generate four overlapping blocks of sequence (BLH10_1, BLH10_2, BLH10_3 and BLH10_4 in figure below). The contiguous sequence was assembled by comparison with the predicted cDNA sequence from the NCBI database (BLH10_predicted) using the ContigExpress® software. Numbers after sequence names indicate the positions on consensus sequence.
The BLH10 cDNA (AY570508) was sequenced and the contiguous sequence of 1906 bp was assembled (Appendix 1.6a). The protein translation was carried out using the Translate program (GCG). The ATG start codon is highlighted by the grey box and the predicted protein is in bold type.
Appendices

841  tattaccatcaaatggaagcattagcttcatcatttgagatagtacgagactttgatca
         Y Y H Q M E A L A S S F E I V A G L G S

901  gctaaagcctacacacatcagggtctctcacaacagaattctctcgccatatgtctgctttcgc
         A K P Y T S V A L N R I S R H F R A L R

961  gacgcctaataaggaacaaattcagattgtagaggagataccaggtgggagagacgtgcag
         D A I K E Q I Q I V R E K L G E K G G E

1021  tcgttggatgagcaaaagagtttgcaccaacagcgtgggaattctggtgatcaacgcttaa
         S L D E Q Q E R G E R I P R L R Y L D Q R L

1081  agacagcaagagctttgcataacacagctgggaattctggtggagacctcaaa
         R Q Q R A L H Q Q L G M V R P A W R P Q

1141  agaggtcttcctgaaactctgtctgctcttgggccattttcgtgcctttgacatattcttt
         R G L P E N S V S V L R A W L F E H F L

1201  catcccatatccgaaagaattctgtagaaaatcatgctttgcagacagaggattgtcaaa
         H P Y P K E S E K I M L A K Q T G L S K

1261  aaccaggtgctaatggtgctcataaaacgcagagttcgtctgtgtgaccggatgatcgga
         N Q V A N W F I N A R V R L W K P M I E

1321  gaaatgtatagagaaggtttttggtgtagaatctctgttagttctaatctctcacaattctttccaa
         E M Y K E E F G D E S E L L I S K S S Q

1381  gaacccaacagcacaacaaaccagagctccctatcgcagcagcagcagcagcagcagcagac
         E P N S T N Q E D S S Q Q Q Q Q Q E N

1441  aacacaacagcagctctgtctattcatctctcagacacaaacaaacattgtctttctcatca
         N N N S N L A Y S S A D T T N I V F S S

1501  gaaacccaaccagatcgcgttctacgcaatgataaaccacccagacacacacagataacrac
         Y I S G S N Q E S R F S N S H H L H D

1561  cgctcattccaggtaccagcatacctcggtacatcagctcgtcaggttttgggtgtgtagattaccgt
         E T K P D R V L G N D N D P Q Q Q Q I N

1621  tacattaccggaagcaacacagacaagagagttctcctctcattccctacttacacgcag
         R S S D Y D T L M N Y H G F G V D D Y R

1681  ttggtttgtgtgtatattctgttctttcttctgtacagagonggtgtattcttgagacgca
         F V V * Y Y L V L L G T E E M I L E A Q

1741  gaaatgtcttttccttttttatagttaggcataataacaccctttctctatgattat
         E M C F C I F I V L G I I H P F S M D Y
1801  atagtagatgtgtagttgcttttgtagttagactcttaagccaaatatgatatgt
       I  V  D  V  *  V  A  F  V  Y  S  R  L  L  S  Q  I  L  I  C

1861  gtgtataataaatcactcaggtgtataaaggatacttgcaagtgtg
       V  Y  N  N  I  H  G  G  I  K  D  T  C  K  C
APPENDIX II: Constructs used for yeast 2-hybrid assays

APPENDIX 2.1 PAS:GI plasmid map

A partial GI cDNA fragment was cloned NcoI/PstI into the yeast 2-hybrid vector pAS2.1 by Kim Snowden (this laboratory) (Chapter 3.1.1). This construct was used in yeast 2-hybrid assays during this thesis work.
APPENDIX 2.2  RESTRICTION ENZYME MAPS OF BLH3 DELETION CONSTRUCTS

PACT:DOM2 AND PACT:DOM3

Dom2 and Dom3 (Chapter 4, Figure 4.1) were cloned into the yeast 2-hybrid vector pACT by Carly Pullen (this laboratory). These constructs were used in yeast 2-hybrid assays during this thesis work. Restriction enzyme maps (not to scale) showing the restriction enzyme sites used to confirm the identity of the Dom2 and Dom3 inserts.

APPENDIX 2.3  RESTRICTION ENZYME MAPS OF GI DELETION CONSTRUCTS

PAS:DEL1, PAS:DEL2 AND PAS:DEL4

GI deletions Del1-4 (Chapter 4, Figure 4.5) were cloned into the yeast 2-hybrid vector pAS2.1 by Carly Pullen (this laboratory). These constructs were used in yeast 2-hybrid assays during this thesis work. Restriction enzyme maps (not to scale) showing the restriction enzyme sites used to confirm the identity of the inserts.
APPENDIX III: SEQUENCES OF THE GI PROTEINS

The protein sequences used for computational comparisons were derived from the cDNA sequences of nine GI genes, including one Arabidopsis gene. The Arabidopsis, Hordeum, Oryza and Triticum sequences are full length. These sequences were identified in GenBank and their accession numbers are as follows. The six other partial sequences were derived from EST sequences in the TIGR Gene Indices.

Arabidopsis thaliana At; Hordeum vulgare Hv; Lycopersicon esculentum Le; Oryza sativa Os; Pinus taeda, Pt; Sorghum bicolor Sb; Solanum tuberosum St; Triticum aestivum Ta; Vitus vinifera, Vv.

AtGI AY682088

MASSSSEWERQDQFLQSSLLWPPRDPQHKDQVAVYVEYFGQFSEFQPDIAELVRHQYFPPSTEKRLLDDVLA
MEVLHPEHGHAVLIPIISCLIDGSLVYSEKEAHPPFASISLVCPCSSENDYSEQWALACGEIILRTHYRPI
YKTEQLQGQDTERCSKATTSGSTPEQAKGSPTQHERKPLRPLSWISDILLAAPLGLSRDYFRWSGV2VMGYK
AAEGLKPTTAIRSGSGKHQMPSTPRWAVANGAVLSDCDEVARYETLAVAPALLPPLFFTPLSDEH
LVAGLFLAPEYARFLFHYYAIAITPSATQRLLLLGLLEAEPSPADALDAAVQVTELLRAADEYASGVLRNPW
MLPLIATRAIGMARGAGVAADAAALFRILQPALLFPPLSQEIPVGEIAPGYYRQNKQYPFVAEATTI
EATAQGIASMLCAGHPFENWRICTINEAAYGLFRGPILNSAVDLLPEIVATPLQPFSILWNYLIPILKLYLFR
GSPEACLMKIVFVAVETILSRTPPEPESRELTKARSSFTTSATKNLAMSELMARVHLFESCAVGELAS
RRFVVTLVCVSHEAQSGSSKRPSREYASTTIEANQPVSNQTANKSNGQVQKQFVAAFDSYVLAAVCALEC
AVQFLYPMISGGGNFNSVASVAGTITKPVINGSGSKYAGIDSASHTRIPAILFAESLKFSSVQFTWSY
SEISEAMAAAIAHISERFRSKALTHASLMGRCWDDIKEIHRASSLYNLIDVHKVASIDKEAPEYAFL
KNTPQVQDSTVCLNWKQENTCSTCTCTDATVSTAASRTENMRGGGHKYSYRHDSEGSRFSEGKIDKFDLASSD
ANFLTADRLACFGTTQKLRSVLAEKPSLFSVSVLLLHKLIAPAEIQTPAETSAQQQWQRQVVDACNLNV
ASATAAAAVLQAEPEFAQFFKIDDEEQGKMWNKQNRIQVKVILVEMNNHRDFPESLVLASASDLRATDGM
LVDGEACCLPQLLELELATARAIPQVWAVPGSLVAGLDSNLKCLPQATIRCLHSPEHRSVLRSTVLRIDM
NQSSIP1KVTPKLPTEKNGMNPSYRFNNAAISWDKADIQNCLNWEAHSLSTTMTQFLDTAARELGCITIS
LSQ

HvGI AY740523

MASNGKWDQLDQSSLFWPPHPDQAQQKAQILAYVEYFGQFSDQEQFDVEAQALIQTCPYSEKEKLVDEVL
ATFVLHPEHGHAVLIPIISCLIDGSLVYSEKEAHPPFASISLVCPCSSENDYSEQWALACGEIILRTHYRPI
FKVADCNNTSDQATTSCSAQEKANYSQGNEPERKPLRLPSFWITDILDLPGLSRDYFRWSGV2VMGYKAA
GELKPFPTTATRSGSGKHQMPSTPRWAVANGAVLSDCDEVARYETLAVAPALLPPLFFTPLSDEHLY
AGLFLAPEYARFLFHYYAIAITPSATQRLLLLGLLEAEPSPADALDAAVQVTELLRAADEYASGVLRNPW
MLPLIATRAIGMARGAGVAADAAALFRILQPALLFPPLSQEIPVGEIAPGYYRQNKQYPFVAEATTI
EATAQGIASMLCAGHPFENWRICTINEAAYGLFRGPILNSAVDLLPEIVATPLQPFSILWNYLIPILKLYLFR
GSPEACLMKIVFVAVETILSRTPPEPESRELTKARSSFTTSATKNLAMSELMARVHLFESCAVGELAS
RRFVVTLVCVSHEAQSGSSKRPSREYASTTIEANQPVSNQTANKSNGQVQKQFVAAFDSYVLAAVCALEC
AVQFLYPMISGGGNFNSVASVAGTITKPVINGSGSKYAGIDSASHTRIPAILFAESLKFSSVQFTWSY
SEISEAMAAAIAHISERFRSKALTHASLMGRCWDDIKEIHRASSLYNLIDVHKVASIDKEAPEYAFL
KNTPQVQDSTVCLNWKQENTCSTCTCTDATVSTAASRTENMRGGGHKYSYRHDSEGSRFSEGKIDKFDLASSD
ANFLTADRLACFGTTQKLRSVLAEKPSLFSVSVLLLHKLIAPAEIQTPAETSAQQQWQRQVVDACNLNV
ASATAAAAVLQAEPEFAQFFKIDDEEQGKMWNKQNRIQVKVILVEMNNHRDFPESLVLASASDLRATDGM
LVDGEACCLPQLLELELATARAIPQVWAVPGSLVAGLDSNLKCLPQATIRCLHSPEHRSVLRSTVLRIDM
NQSSIP1KVTPKLPTEKNGMNPSYRFNNAAISWDKADIQNCLNWEAHSLSTTMTQFLDTAARELGCITIS
LSQ

OsGI AK072166
GESGQSVADGLTNLLKCRPATVRCVSPASAHVRALSTVLRDIMYAGSVKPSAKQAADVNGIHNPAYQYLGI
SISDWKADIEKCLMWEANSRLENGMSAQFLDTAARELGCTISV

**VvGI TIGR Grape Gene Index**

AAAMVAAHVSELFRRSKACMHALSVLMRCKWDEEIYTRASSLYNLIDIHSKAVASIVNKAEPELAHILIHATVW
KDSPHGDSEKENDCASTSCFVSNPLLLLHEEDSAYSKSLPFEKAPHLENGTGSNGSLKGIASFPLDASELAN
FLTMDDRHIGFSCSAQVLLRSVLAECQECLCFVSVVLLWKLIAAPETKPSAESSTAQQWRQVVDACLVVSAS
PAKAAATAVVLQAEERLQWPWAKDDLGQKMNRLNQRIVKLVIELMRRNHRPESLVILVSSADLLRLRATDGMV
DGEACTTPQLELLEATARAVQLVLENGESGLAVADGLSNLLKCRVPATIRCLSHPSAHVRALSTVLRDVLQG
GSIKPHIKQGGRGHSQYQVNLGIIDWQADIEKCLTWEAHSRLATGMNQFLDAAAKELGCTISI

**PtGI TIGR Pinus Gene Index**

VLANASDLRATDMLVDGEACTTPQLELLEAMAVAAQLSLGWGPKAMADGLWNLLKYRLPATVQCLSHS
SAHVRALSTVLRDILHAESLNFRYCKNISEKHHSEHLYYGKDMSVQDWHKAVEQCLAWEAHRNGMSVSL
LILAANALGFSANV
APPENDIX IV: GFP EXPRESSION ANALYSES

APPENDIX 4.1  TRANSIENT GFP EXPRESSION IN ARABIDOPSIS LEAF TISSUE

Transient expression assays were carried out with constructs encoding the GFP reporter gene (35S:GFP) and GFP fused to the N-terminal of BLH3 (35S:GFP:BLH3) (Chapter 5.3). These constructs were introduced into the abaxial surface of Arabidopsis leaves by particle bombardment (Chapter 2.4.5). Samples were viewed using blue light excitation to detect GFP fluorescence.

a  Arabidopsis epidermal cells exhibiting auto-fluorescence due to tissue damage after particle bombardment. This auto-fluorescence is not masked by the GFP filter.

b  Arabidopsis leaf after particle bombardment with a water control.

c  Arabidopsis epidermal cell expressing 35S:GFP.

d  Arabidopsis epidermal cell expressing 35S:GFP:BLH3, indicated by the arrow.

e  Arabidopsis epidermal cell expressing 35S:GFP. Viewed under a confocal scanning laser microscope.

f  Arabidopsis epidermal cell expressing 35S:GFP:BLH3, indicated by the arrow. Viewed under a confocal scanning laser microscope.
APPENDIX 4.2  GFP EXPRESSION IN TRANSGENIC Arabidopsis plants

Binary constructs encoding the GFP reporter gene (35S:GFP) and GFP fused to the N-terminal of BLH3 (35S:GFP:BLH3) were introduced into Arabidopsis plants (Col ecotype) (Chapter 5.4). T₂ plants were selected (Chapter 2.4.4) and viewed at 20x magnification, using blue light excitation to detect GFP fluorescence. The brightness/contrast of these pictures was adjusted equally.

a Arabidopsis epidermal cells expressing 35S:GFP.
b Arabidopsis epidermal cells; Col-0 wild type control.
c Arabidopsis epidermal cells expressing 35S:GFP:BLH3.
d Arabidopsis epidermal cells expressing 35S:GFP:GI.
APPENDIX V: GENERATION OF ANTIBODIES TO BLH3

A polyclonal BLH3 antibody was raised in rabbits with the aim of using it to detect the BLH3 protein in plants. The MBP:BLH3 fusion protein encoding BLH3 aa 180-524 (Chapter 3.3), was used for this purpose. The MBP:BLH3 fusion protein was expressed in \textit{E. coli} and purified from crude cellular extract as described in Chapter 2.6.7.5. A factor Xa cleavage site was present at the junction between the MBP and BLH3 proteins. Analysis of the BLH3 protein sequence indicated that the BLH3 protein fragment fused to MBP did not contain any Factor Xa cleavage sites. As a result, the 42 kDa MBP protein was released from the truncated BLH3 protein by enzymatic digestion with 1µg activated Factor Xa (Sigma). However, the ~ 25 kDa BLH3 protein released from MBP was also cleaved into two smaller protein fragments of similar size (~ 12 kDa and ~ 13 kDa) by Factor Xa digestion (data not shown).

It proved difficult to separate the two BLH3 protein fragments from the MBP and the two Factor Xa subunits present in the solution and subsequently purify BLH3 to a sufficiently high concentration. Therefore, the entire MBP:BLH3 fusion protein was used to raise antibodies. A serum containing polyclonal antibodies raised to MBP:BLH3 will include some antibodies specific to MBP epitopes and some to BLH3 epitopes. A search of the \textit{Arabidopsis} database verified that the MBP protein was sufficiently diverged from \textit{Arabidopsis} proteins and not expected to cross-react with \textit{Arabidopsis} proteins.

A total of 2 mg of MBP:BLH3 was purified and supplied to Agresearch (Ruakura), where it was used to inoculate a New Zealand White rabbit on three separate occasions. The serum containing putative BLH3 antibodies was obtained. Western analysis confirmed that polyclonal antibodies present in the serum detected both MBP and BLH3 proteins. The BLH3 antibodies were used to detect the BLH3 fragments released from MBP:BLH3 by Factor Xa digestion, down to 1 ng concentration of total protein (Figure A5.1a). Fortuitously, at 1 ng total protein the polyclonal antibodies had greater affinity to the BLH3 fragments than the uncleaved MBP:BLH3. Due to incomplete digestion of MBP:BLH3, the MBP:BLH3 protein was present in higher concentrations than the smaller BLH3 protein fragments.

<table>
<thead>
<tr>
<th>Total protein concentration</th>
<th>Plant Line</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>Col</td>
<td>Col</td>
</tr>
<tr>
<td>blh3</td>
<td>blh10</td>
</tr>
<tr>
<td>blh3blh10</td>
<td>blh3blh10</td>
</tr>
<tr>
<td>gi-2</td>
<td>gi-2</td>
</tr>
</tbody>
</table>

254
The recombinant MBP:BLH3 fusion protein was expressed in *E. coli* and purified (Chapter 2.6.7.5). Samples were subject to digestion with 1 µg Factor Xa and diluted to the concentrations shown. Protein fragments were separated by SDS-PAGE and subject to Western analysis as described in Chapter 2.5.7.3.

Crude total protein was extracted from plant tissue from four plant lines; Col-0, *blh3*, *blh10* and *blh3blh10gi-2*. Total protein (approximately 100 mg) was separated by SDS-PAGE and subject to Western analysis as described in Chapter 2.5.7.3.

Crude protein extract was prepared from plant seedlings (Chapter 2.5.7.1) and separated by SDS-PAGE (Chapter 2.5.7.2). Preliminary Western analysis demonstrated that a putative BLH3 protein was detected in wild type Col-0 plant tissue (Figure A5.1b). At ~ 45 kDa this was a little smaller than predicted for the BLH3 protein. However, no other proteins were detected by this antibody. Furthermore, only a very slight band was detected in the *blh3*, *blh10* and *blh3blh10gi-2* mutant plants, providing supporting evidence that this polyclonal antibody detects BLH3 in plants. There was no cross reactivity with other *Arabidopsis* proteins, as indicated by the minimal background obtained in this Western analysis (Figure A5.1b).

The relatively weak signal obtained in preliminary Westerns may indicate that BLH3 antibodies need to be purified. Before future experiments, affinity purification of the serum to deplete MBP antibody will increase the concentration of BLH3 antibodies in the serum and thus the sensitivity of the antibody. Alternatively, the concentration of BLH3 in plant extracts may be increased. It has been verified that the BLH3 protein is localized to the nucleus in transient expression assays (Chapter 5.3.2). Purification of protein solely from plant nuclear extracts would be expected to increase the concentration of BLH3 in the protein extract. With further testing of the antibody and optimisation of the protocol, the BLH3 antibody will be extremely useful for analysing the expression pattern of the BLH3 protein in plants.

**APPENDIX VI: SUMMARY OF BLH3 AND BLH10 MUTANTS**
Summary of the $blh3$ and $blh10$ T-DNA mutants generated and the flowering time and hypocotyl elongation assays performed with these plants. This work is described in detail in Chapter 6.

Flowering time and hypocotyl length of mutants were compared to wild type plants. A result equivalent to wild type is expressed as $= wt$, a small difference to wild type as $s$. Light conditions are as follows: LD, long day; SD, short day; DD, constant dark; LL, constant light; Bc, constant blue; Rc, constant red; $\frac{1}{2}$SD, low light short days; 0.1LL, constant very low light.

<table>
<thead>
<tr>
<th>MUTANT</th>
<th>FLOWERING TIME</th>
<th>HYPOCOTYL LENGTH ASSAYS (COMpared TO WILD TYPE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LD</td>
<td>SD</td>
</tr>
<tr>
<td>$blh3-1$</td>
<td>s late</td>
<td>late</td>
</tr>
<tr>
<td>$blh10-1$</td>
<td>s late</td>
<td>late</td>
</tr>
<tr>
<td>$blh3blh10$</td>
<td>= wt</td>
<td>= wt</td>
</tr>
<tr>
<td>$blh3gi$</td>
<td>late</td>
<td>late</td>
</tr>
<tr>
<td>$blh3 Gi/gi$</td>
<td>late</td>
<td>= wt</td>
</tr>
<tr>
<td>$blh10gi$</td>
<td>late</td>
<td>= wt</td>
</tr>
<tr>
<td>$blh3blh10gi$</td>
<td>late</td>
<td>= wt</td>
</tr>
<tr>
<td>35S:GI $blh3gi$</td>
<td>early</td>
<td>short</td>
</tr>
</tbody>
</table>
APPENDIX VII: AN INDUCIBLE GI EXPRESSION SYSTEM

APPENDIX 7.1 INTRODUCTION

Chemical gene induction systems are a powerful tool for regulating the expression of transgenes in plants. Possible applications include controlled expression of transgenes that interfere with growth, reproduction, or cause lethality when inactivated during development, conditional expression of transgenes and correlation of induction of a gene such as GI with a specific altered phenotype (reviewed by Gatz and Lenk, 1998; Zuo and Chua, 2000). For this thesis project, an inducible gene expression system was employed to identify GI target genes that are up- or down-regulated in response to induction of GI expression.

A number of properties are required in an ideal inducible promoter system. In the absence of an inducer, expression levels must be close to zero and conversely must be high when the inducer is present. A chemical inducer should be of low toxicity to the plant, highly specific to the promoter system, easily applied and have a high efficacy at low concentrations (reviewed by Zuo and Chua, 2000; Padidam, 2003). There are a number of systems that have proven successful in the temporal control of gene expression in Arabidopsis, including both steroid and tetracycline based systems (Aoyama and Chua, 1997; Gatz and Lenk, 1998; Bohner et al., 1999). An ethanol inducible promoter system allows reversible induction of the target gene and is able to be modified to allow efficient induction of gene expression in specific plant tissue or domains (Roslan et al., 2001; Deveaux et al., 2003). This system has been used to successfully induce gene expression at different stages in plant development to determine the temporal role of the UFO gene in floral development (Laufs et al., 2003). For this thesis work, both the ethanol inducible promoter and steroid dexamethasone inducible promoter systems were used.
APPENDIX 7.2 THE ETHANOL INDUCIBLE GENE EXPRESSION SYSTEM

An ethanol-inducible promoter was selected to regulate GI expression. This system is based on an alcohol regulation region that is used to control an ethanol utilisation pathway in the fungi Aspergillus nidulans (Fillinger et al., 1995). This system responds specifically to ethanol and involves a transcription factor (AlcR) and promoter region (AlcA) with sufficient evolutionary divergence from plants to avoid activation by endogenous plant proteins. When ethanol is added the AlcR transcription factor binds the AlcA promoter and gene expression is activated (Figure A7.1). The induction of the target gene is rapid, with expression detectable as little as 1 h after ethanol treatment (Roslan et al., 2001; Deveaux et al., 2003).

![Figure A7.1 Activation of the AlcA promoter occurs only in the presence of ethanol](image)

Ethanol is a volatile liquid, so ethanol vapour from a beaker of ethanol is sufficient to induce gene expression (Sweetman et al., 2002). Therefore ethanol is easily applied and easily removed. It has been shown that ethanol is non-toxic to the plant in the concentrations required to activate expression and that a single application can induce high levels of transgene expression (Salter et al., 1998). This process is bi-directional as gene expression rapidly reduces once the ethanol is removed (Caddick et al., 1998; Roslan et al., 2001). However, the rate at which gene expression ceases following ethanol withdrawal is variable and has variously been reported to take hours, days, or weeks (Roslan et al., 2001; Sweetman et al., 2002; Deveaux et al., 2003)
Appendices

**Expression tests for the induction of GI**

Previously, transgenic lines containing the ethanol inducible GI expression construct were generated. Ten lines were selected and tested for induction of GI expression (Milich, 2001). The induction of GI was unsuccessful in these lines and unfortunately GI was present at detectable levels even before treatment with ethanol. Furthermore, these lines were late flowering in a gi-11 background, indicating that induction levels of GI were insufficient to complement the gi mutation.

It has since been demonstrated that the AlcR transcription factor that is constitutively expressed in plants grown in tissue culture is detrimental to plant growth (Roslan et al., 2001; Roberts et al., 2005.). In this case, the plant lines with the highest expression of genes involved in ethanol induction would be selected against. As part of this thesis work new transgenic plant lines containing the ethanol inducible GI expression construct were selected on ½ MS plates containing kanamycin to select for the T-DNA insertion. As these plants do not survive well in tissue culture, seedlings were selected and planted in soil before they had developed true leaves. The majority of the plants displayed unusual developmental phenotypes, such as fasciated stems, stunted bolts and extra leaves. This demonstrated that despite minimal time in tissue culture, the growth and development of these plants had still been affected.

Preliminary ethanol induction tests were carried out on seven selected plant lines. Previous induction tests using a 1% ethanol treatment to induce GI expression did not result in sufficient GI transcript levels (Milich, 2001). In an attempt to increase GI expression, plants were treated with 5% ethanol in water for two successive days and tissue was harvested 24 h later at 4 h after dawn when endogenous GI levels are low. Tissue from plants treated with a water control was also harvested 4 h after dawn as a negative control. The expression of GI was examined using Northern analysis. As expected a 4 kb transcript corresponding to wild type GI was present before ethanol treatment as the expression construct was transformed into a wild type Ws plant. There was no increase in the levels of the 4 kb GI transcript after ethanol treatment, although there was induction of larger transcripts of unknown origin (data not shown).

Five more transgenic plant lines were selected and tested for GI induction. Again, there was no increase in the 4 kb GI transcript after ethanol induction (Figure A7.2). In contrast, larger transcripts were seen both before and after ethanol induction, though at higher levels
after ethanol treatment. It was possible that the ~7 kb inducible gene product was a fusion of GI (4 kb) and AlcR (3 kb). To test this assumption, the Northern blots were re-probed with an AlcR gene fragment. The 3 kb AlcR transcript was present as expected. The larger transcripts were also detected with the AlcR probe, indicating that these transcripts are likely to be variations of an GI:AlcR fusion (Figure A7.2). The expression of the GI:AlcR fusion was not present at high levels or at all times, therefore it is unlikely to be under the control of the 35S promoter. The 7 kb transcript was ethanol inducible in some lines (ie, plant line 3, Figure A7.2), therefore is possibly the result of transcription from the AlcA promoter.

**Figure A7.2 Expression of GI and AlcR in ethanol inducible plant lines**
Plants were grown in LD and harvested at 4 h after dawn (− ethanol control) and 24 h later after ethanol treatment (+ ethanol). RNA was extracted and GI and AlcR expression was examined by Northern hybridisation analysis, by using GI and AlcR probes respectively. 1-5; five arbitrary transgenic lines.

**ANALYSIS THE STRUCTURE OF THE GENE EXPRESSION CONSTRUCT**
The ethanol inducible transgenic plant lines were generated by transformation of Arabidopsis with a T-DNA containing 35S:AlcR and AlcA:GI (Milich, 2001). As the GI:AlcR fusion transcript is present in multiple independent transgenic lines, it is unlikely that the cause of this aberrant transcript is a rearrangement of the T-DNA during plant transformation. Consequently, the pR20 binary plasmid containing this T-DNA was examined further.

The pR20 vector was digested with restriction enzymes and examined by Southern analysis. The restriction enzyme digestion produced fragments of the expected size, suggesting that if any rearrangement had occurred it was too small to be visualised by gel
 electrophoresis (data not shown). The Southern blot was probed with AlcR and nos 3’ sequences. The nos 3’ terminator sequences are located at the 3’ end of both the AlcR and GI genes, hence are present within the T-DNA as an inverted repeat (Figure A6.3). The results of the Southern blot indicated that a rearrangement had occurred in the region spanning the nos 3’ sequences (data not shown). A rearrangement in the termination sequences would be expected to alter the termination signal normally recognised by RNA polymerase. As a result RNA polymerase may read through without disrupting the GI mRNA, thus resulting in a longer transcript than expected.

![Figure A7.3 Structure of the pR20 gene expression construct](image)

**Figure A7.3 Structure of the pR20 gene expression construct**

Schematic diagram of the 35S:AlcR and AlcA:GI expression cassettes in the pR20 binary plasmid. The arrows indicate the direction of transcription. **AlcA**: AlcA promoter sequence, **AlcR**: AlcR transcription factor, **GI**: GI cDNA; 3’; **ocs** 3’ terminator sequences, 35S; CaMV 35S promoter, **PacI**: restriction enzyme site.

**TESTING THE FUNCTIONALITY OF INDUCED GI**

To establish if the GI:AlcR transcripts induced by ethanol are functional in plants, the flowering time of these plant lines was determined. Six transgenic lines shown to produce the GI:AlcR transcript after ethanol induction were selected and flowering time of these plants in SD was tested. Wild type Ws plants typically flower late and plants constitutively expressing GI flower earlier in SD (Milich, 2001). However, the ethanol inducible GI lines did not flower significantly earlier than wild type in these SD conditions. On the contrary, a number of lines flowered later than wild type (data not shown).

Was GI not functional in plants due to the fusion to AlcR sequences or was induction of the GI transcript simply not at sufficient levels to induce early flowering? Regardless of the answer to this question, this system was unable to produce reliably inducible and functional GI in plants. The ethanol inducible lines were not used any further in this thesis work.
APPENDIX 7.3 THE DEXAMETHASONE INDUCIBLE GENE EXPRESSION SYSTEM

The steroid inducible gene expression system used in this thesis work was obtained from I. Moore (University of Oxford). This system comprises two components; a p11OP promoter containing 11 lac operators upstream of a minimal CaMV 35S promoter and the LhGR transcription activator which consists of a lac DNA binding domain, the S. cerevisiae GAL4 transcription activation domain and the ligand binding domain of a rat glucocorticoid receptor (Craft et al., 2005). The LhGR activator is sequestered in an inactive complex in the absence of steroid ligands in plants. After the application of dexamethasone (dex), a synthetic glucocorticoid class hormone, the LhGR protein is able to activate transcription of genes cloned downstream of the p11OP promoter.

The GI coding sequence was cloned downstream of the p11OP promoter in a binary vector (outlined in Figure A7.4). This construct, pBIN:11OP:GI was transferred into Col wild type and a 35S:LhGR-N plant line via Agrobacterium mediated transformation as described (Chapter 5.3.3.2). Thirty 11OP:GI 35S:LhGR and forty 11OP:GI Col lines were created. The 11OP:GI Col plants were generated to be introduced into plants containing LhGR under the control of tissue specific promoters to test the effect of GI induction in different plant tissues. However, due to time constraints these lines were not used further in this thesis work. The 35S:LhGR 11OP:GI lines (R Lines) were used for preliminary GI induction assays.

Figure A7.4 Generation of the dex inducible construct pR34

Strategy for the construction of the binary vector pOP:GI. The 5’ coding sequence of GI, including the ATG, was amplified from the GI cDNA and inserted into pGEMT. The full length GI cDNA was released from PBS:GI by digestion with KpnI/BamHI and inserted into PBS-SstI. The 5’ GI sequence, including the UTR, was released by digestion with KpnI/SstI and was replaced with GI5’ which lacks the UTR. The pR30 construct created contains the entire GI coding sequence but lacks the 5’ UTR. GI was released from pR30 by digestion with KpnI/BamHI and subcloned downstream of the pOP promoter in p11OP to create pR33. Finally, the pOP:GI expression cassette was released from pR33 by digestion with NotI and inserted into the binary vector pR7 (Chapter 2.1.3), creating pOP:GI.
PCR amplification and ligation to pGEMT

ATG
GI cDNA
940 bp product

Digest with KpnI and BamHI
Ligation

ATG
pR29

Digest with KpnI and SstI
Ligation

ATG
pR30

Digest with KpnI and BamHI
Ligation

ATG
pR33

Digest with NotI
Ligation

ATG
pBIN:11OP:GI

4.1 kb

MCS
KpnI/BamHI

pBS:SstI-

1.1 kb

KpnI
SstI

KpnI
BamHI

11 OP
ocs 3'

11 OP
ocs 3'

LB
RB

11 OP
ocs 3'

p11OP

pR30

690 bp

KpnI
SstI

KpnI/BamHI

pGEMT:GI5'

p11OP

pR30

1.1 kb

pR7

3.6 kb

p11OP

3.6 kb

pR7

p11OP

pR33

pR33

pR7

pR7

pR7

pR7

pR7

pR7

pR7

pR7
**BREEDING 35S:LhGR OP:GI gi-2 PLANT LINES**

Fifteen 35S:LhGR 11OP:GI lines were selected and crossed to gi-2 mutants to remove the wild type GI gene. Seedlings were grown in media containing either kanamycin or hygromycin to select for the 35S:LhGR and 11OP:GI insertions respectively. Subsequently, plant lines containing single 35S:LhGR and 11OP:GI insertions were identified by segregation analysis. Late flowering plants were selected in the F2 generation and these were subject to PCR to verify that they were homozygous gi-2 mutants.

Late flowering plants were also tested for the presence of the 35S:LhGR and 11OP:GI constructs by PCR (Figure A7.5b and c). Although these 35S:LhGR 11OP:GI gi-2 plant lines (V lines) were segregating for the 35S:LhGR and 11OP:GI inserts they were also used in preliminary gene expression assays.

![Image](image.png)

**Figure A7.5 Identification of plants containing 35S:LhGR and 11OP:GI**

Genomic DNA was extracted from segregating F2 plants from the V15 plant line and used as a template for PCR amplification. The expected products were visualised on an agarose gel.

- **a** PCR for the gi-2 mutation. The junction spanning gi-2 deletion was amplified using the oli22 and oli37 primers, resulting in a 148 bp product from Col and 140 bp fragment from gi-2. Products were separated on an 8% acrylamide gel, only 100 bp and 200 bp fragments visible on this gel.
- **b** Two GI specific primers (oli33, oli14) were used to amplify the GI gene. The cDNA specific (600 bp) and gDNA specific (820 bp) products were visualised on an agarose gel.
- **c** The GR sequences within LhGR were amplified by the GR1.F and GR1.R primers. The expected product (300 bp) was visualised on an agarose gel.

+; positive control, Col; wild type, 1 kb; 1 kb plus DNA size marker.
**Expression tests for Dex induction of GUS expression**

Control plant lines containing the *E. coli* β-glucuronidase (GUS) reporter gene cloned downstream of the dex inducible operator sequences (35S:LhGR OP:GUS) were obtained from I. Moore (University of Oxford). These control plants were used to trial dex induction of gene expression. Dex was dissolved in ethanol at 10 mM concentration and diluted in water and 0.01% Tween-20 to the required concentration before treatment. The leaves of the control plants were sprayed or painted on one side using a bristle paintbrush with 30 µM dex solution and stained for GUS activity (Chapter 2.4.6). Some GUS staining was detected, although certainly not to the levels equivalent to 35S:GUS, which is expected for these lines (data not shown). The concentration of dex was doubled to 60 µM and the solution was sprayed onto control plants. An increase of GUS expression was seen in plants treated with the higher concentration of dex. As expected, GUS expression was detected throughout the plant, including rosette leaves, stem and flower tissue, however this expression pattern was not uniform (Figure A76.6a and b). This expression pattern could be due to the uneven application of the dex solution by spraying. Alternately, gene silencing could affect expression of the GUS reporter gene in these plants.

![Figure A7.6 Induction of GUS expression in OP:GUS control plants](image)

**Expression tests for Dex induction of GI expression**

Initial GI induction tests were carried out by both spraying and watering plants with the suggested concentration of 30 µM dex in ethanol. Plants segregating for 35S:LhGR 11OP:GI were grown in LD conditions and leaf tissue was harvested before dex treatment and 24 h after dex treatment. Little induction of the GI transcript was seen and leaf tissue was particularly unhealthy after dex treatment (data not shown). One positive from these
experiments was the low level of background GI expression in these plant lines (Figure A7.7). As these plant lines are in the Col-0 wild type, the background levels of GI expression are likely due to the endogenous GI transcript, which is low at this time of day. Although GI induction was not detected, in the absence of dex these lines did not exhibit increased levels of GI expression, verifying that this gene expression system tightly regulates GI expression.

![Figure A7.7](image)

**Figure A7.7  Little GI expression is detected in the absence of dex**

Total RNA (20 µg) was extracted from leaf tissue at 4 h after dawn from dex inducible R-lines. Expression of GI was analysed by Northern hybridisation using a GI cDNA probe. Ethidium bromide stained rRNA was used as a loading control.

A repeat of the induction assay was performed in light of the tests performed with the OP:GUS control plants. Plants were again grown in LD, this time treated with 60 µM dex by both spraying and painting onto the leaves. Leaf tissue was harvested 24 h later at 4 h after dawn, a time when the endogenous GI transcript is low. This induction was more successful, with several plants lines exhibiting an increase in GI expression after treatment with dex (Figure A7.8).

Both spraying and painting the dex containing solution onto leaves effectively induced GI expression. Painting the dex solution resulted in greater induction of GI, although this was also associated with greater background (Figure A7.8). One plant line R13 had no induction of GI under any conditions and this line was discarded. This is likely due to the genome position of the T-DNA containing 11OP:GI, as this can affect transgene expression. The different levels of GI induction seen between the plant lines may be partly attributed to genome positional effects, although this assay was not quantitative. Furthermore, the V lines were segregating for the LhGR and OP:GI insertions. The lack of induction often seen in these lines compared to the corresponding R lines was likely due to the lack of one of these insertions in the individual plants tested.
### Figure A7.8 Examination of GI expression in plants by Northern analysis

**a** Total RNA (20 µg) was extracted from leaf tissue at 4 h after dawn from plants treated with dex by spraying or painting. Expression of GI was analysed by Northern hybridisation using a GI cDNA probe. The rRNA probe was used as a loading control.

**b** Summary of dex induction of GI in 35S:LhGR OP:GI (R) and 35S:LhGR OP:GI gi-2 (V) lines. +, some induction; ++ good induction; -, no induction; nt, not tested

### APPENDIX 7.4 SUMMARY

The two chemically inducible plant expression systems selected for this thesis work were used to induce the GI transcript in transgenic plants. Since this thesis project was started, more publications outlining the strengths and weakness of the chemical expression systems selected have become available.

The ethanol inducible alc gene based expression system has been shown to successfully induce gene expression not only in Arabidopsis but in a range of plant species. As this is a two component system, it is particularly useful for driving gene expression within specific
domains in the plants by restricting AlcR expression to specific tissues (Deveaux et al., 2003). However, the activation of the alc expression system in tissue culture was a serious flaw in this system that may have affected the selection of transgenic lines in this work. Since this time an alc based expression system has been modified for use in tissue culture (Roberts et al., 2005). Alternately, selection of transgenic plants in tissue culture could be avoided by the insertion of a gene that confers herbicide resistance into the T-DNA sequences. Recently presented evidence also shows that ethanol can have an effect on the development of potato plants that have not been transformed with the alc system (Claassens et al., 2005). Junker et al. (2003) suggest that induction of the alc system with acetaldehyde is more efficient and treatment causes fewer changes in metabolite levels than ethanol. Taking into account this new evidence, it is not surprising that the ethanol inducible GI expression system tested as part of this thesis work was not effective at inducing GI expression in plants.

Greater success at inducing GI expression was obtained using a two component dexamethasone inducible gene expression system. This GR-based transcription system has been used effectively to induce the expression of genes in plants. Unfortunately, like the ethanol inducible promoter, difficulties with reliable induction were encountered. Furthermore, plants were particularly unhealthy after dex treatment, thus making it difficult to maintain plants and harvest tissue for expression assays. Recent work has demonstrated that the p11OP promoter used to generate the dex inducible GI construct is unstable due to the multiple repeated lac operator sequences. In fact, it has now been established that four to six copies of the lac operator is stable and sufficient for gene induction (Craft et al., 2005). Furthermore, the inhibition of plant growth following dexamethasone induction has been shown to be due to the use of ethanol as a solvent for this steroid. Dexamethasone dissolved in DMSO is now recommended for activation of the pOP/LhGR gene expression system (Craft et al., 2005). This explains most of the problems encountered with the dex inducible GI system in this section of work.

It should be noted that despite the setbacks, dex induction of GI in these plants lines was relatively successful. The dex inducible system proved far more reliable than the ethanol induction system and background GI expression in the absence of an inducer was significantly less. Of course, the plant lines containing inducible GI in a gi-2 mutant background must be tested further and due to time constraints further testing of GI induction was not able to be carried out as part of this thesis work. A logical starting point
would be to test *GI* induction using DMSO instead of ethanol as a solvent for dex and assess the health and survival of the plants following treatment. The growth of plant lines on GM plates containing increasing concentrations of dex would be useful to see if an increase in *GI* expression is dose responsive. This would also provide quantitative information on the plant lines that best induce the *GI* transcript.

According to recently presented evidence, maximum induction of the LhGR system is observed over 100 h after dex treatment and that plants irrigated with 20 µM dex respond best to repeated treatments 2-3 d apart (Craft *et al.*, 2005). It is likely that testing for *GI* expression 24 h after a single dex treatment was simply not long enough to detect maximal *GI* expression. A single dex treatment can be sufficient to induce a stable protein (reviewed in Moore *et al.*, 2006), however it may be assumed that as the *GI* transcript cycles daily and decreases to undetectable levels each night this transcript it is not particularly stable. In retrospect, repeated irrigation of plants with dex followed by assessment of *GI* expression 3-4 d later may be a more effective *GI* induction protocol. By recognising the shortfalls of the protocols used in this section of work and avoiding them in future assays it is likely that reliable, stable and high levels of *GI* expression may be obtained from this system.

The GR sequence within LhGR contains two sites at which premature polyadenylation occurs, which can lead to the creation of two premature stop codons (I. Moore, pers. comm.). This has not proven to be a problem when the LhGR construct is driven by the constitutive 35S promoter, as it was in the system used in this thesis work. However, this is likely to be problematic if LhGR is cloned downstream of a weaker promoter (I. Moore, pers. comm.). Therefore, expression of the LhGR activator in a tissue specific manner and the subsequent use of the 11OP:GI lines to test *GI* induction in specific tissues may require an alternative or modified dex induction system, depending on the strength of the tissue specific promoters selected.


References


Chan SK, Jaffe L, Capovilla M, Botas J, Mann RS (1994) The DNA binding specificity of ULTRABITHORAX is modulated by cooperative interactions with EXTRADENTICLE, another homeoprotein. Cell 78: 603-615


Devlin PF, Kay SA (2000) Cryptochromes are required for phytochrome signalling to the circadian clock but not for rhythmicity. Plant Cell 12: 2499-2509


References


Duck PD, Fankhauser C (2003) HFR1, a putative bHLH transcription factor, mediates both PhytochromeA and cryptochrome signalling. Plant Journal 34: 827-836


References


Fowler S, Thomashow MF (2002) *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. Plant Cell 14: 1675-1690


Gleave AP (1992) A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. Plant Mol Biol 20: 1203-1207


Hudson ME, Lisch DR, Quail PH (2003) The FHY3 and FAR1 genes encode transposase-related proteins involved in regulation of gene expression by the Phytochrome A signalling pathway. Plant Journal 34: 453-471


Huq E, Tepperman JM, Quail PH (2000) GIGANTEA is a nuclear protein involved in phytochrome signaling in Arabidopsis. Proc Nat Acad Sci 97: 9789-9794


Kikis EA, Khanna R, Quail PH (2005) ELF4 is a phytochrome-regulated component of a negative-feedback loop involving the central oscillator components CCA1 and LHY. Plant Journal 44: 300-313.


Michael TP, Salome PA, McClung CR (2003b) Two Arabidopsis circadian oscillators can be distinguished by differential temperature sensitivity. Proc Nat Acad Sci 100: 6878-6883


Mizoguchi T, Wheatley K, Hanzawa Y, Wright L, Mizoguchi M, Song HR, Carre IA, Coupland G (2002) LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in Arabidopsis. Dev Cell 2: 629-641


Ni M, Tepperman JM, Quail PH (1999) Binding of Phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light. Nature 400: 781-784
References


References


Salome PA, McClung CR (2005b) PSEUDO-RESPONSE REGULATOR 7 and 9 are partially redundant genes essential for the temperature responsiveness of the Arabidopsis circadian clock. Plant Cell 17: 791-803


References


Sothern RB, Tseng T-S, Orcutt SL, Olszewski NE, Koukkari WL (2002) **GIGANTEA** and **SPINDLY** genes linked to the clock pathway that controls circadian characteristics of transpiration in *Arabidopsis*. Chronobiology Int 19: 1005-1022


Yoo SK, Chung KS, Kim J, Lee JH, Hong SM, Yoo SJ, Yoo SY, Lee JS, Ahn JH (2005) \textit{CONSTANS} activates \textit{SUPPRESSOR OF OVEREXPRESSION OF CONSTANS} to promote flowering in \textit{Arabidopsis}. Plant Physiology 139: 770-778


Zhao XY, Liu MS, Li JR, Guan CM, Zhang XS (2005) The wheat \( \textit{TaGI1} \), involved in photoperiodic flowering, encodes an \textit{Arabidopsis} \textit{GI} ortholog. Plant Mol Biol 58: 53-64.