

**DIFFERENTIAL GENE EXPRESSION IN
GERMINATING AND THERMOINHIBITED
ACHENES OF *Tagetes minuta* L.**

PAUL NORMAN HILLS

Submitted in fulfilment of the academic requirements
for the degree of Doctor of Philosophy in the
Research Centre for Plant Growth and Development,
School of Botany and Zoology,
University of Natal

Pietermaritzburg

2003

DECLARATION

The experimental work described in this thesis was carried out in the Research Centre for Plant Growth and Development in the School of Botany and Zoology at the University of Natal, Pietermaritzburg, from January 1997 to April 2003, under the supervision of Professor J. van Staden.

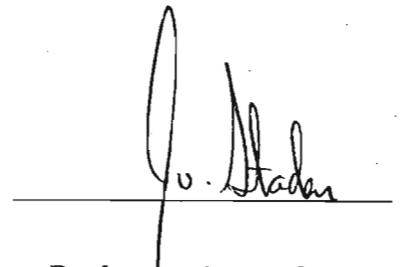
These studies represent original work done by the author and have not otherwise been submitted in any form for any degree or diploma to any other university. Where use was made of the work of others, it has been duly acknowledged in the text.



P.N. Hills

April 2003

I declare that the above statement is correct.



Professor J. van Staden

Supervisor

April 2003



**What is a weed? A plant
whose virtues have not yet been discovered.**

Ralph Waldo Emerson

**To see a World in a Grain of Sand
And Heaven in a Wild Flower
Hold Infinity in the palm of your hand
And eternity in an hour**

William Blake (In *Auguries of Innocence* c. 1800)

ACKNOWLEDGEMENTS

My most sincere thanks are due to many individuals who have contributed toward this study:

- To my supervisor, Prof van Staden, for his support, guidance and encouragement throughout the duration of this study. Thank you for your constant enthusiasm and for all that you have taught me about Science.
- To my family, for their love and support over the years, as well as to my extended family of Ferreiras and Aucamps.
- To my lab colleagues, especially Dr Peter Hare, Nokwanda Makunga, Sara Nigro and Dr Lindsey Podd. Your support, encouragement and help have been invaluable.
- To all of my friends in the Research Centre for Plant Growth and Development, as well as in the School of Botany and Zoology, who have made this such an exciting and enjoyable experience. Thank you all for making this such a great place to work in.
- To all who have been members of my Research Committee over the years, for their invaluable advice.
- To the technical staff of the School of Botany and Zoology, for their efficient assistance and for making sure that everything worked when I needed it to.
- To Prof E. Balazs, for the sequencing of the cDNA clones.
- To the University of Natal and the NRF (as well as the former FRD) for financial support.
- And finally, a special thanks to my partner Mariaan, for her love, support and encouragement, and for making me finish this thesis.

ABSTRACT

When imbibed at their optimum germination temperature of 25°C, achenes of *Tagetes minuta* L. germinate over a period of approximately 48 h. At temperatures of between 35°C and 39°C, the achenes do not germinate but enter into a state of thermoinhibition. These supra-optimal conditions do not harm the achenes, however, and when the temperature is reduced below 35°C radicle emergence may be observed within 4 h. Achenes which have been thermoinhibited for periods of 24 h or more show “accelerated germination” which takes only 24 h, although the actual germination curve is identical to that of normally germinated achenes. This suggests that the achenes are metabolically active at thermoinhibitory temperatures and undergo most of the processes of normal germination, but that at some point any further development is halted, preventing radicle emergence. When the temperature is reduced, this block on germination is removed and since the achenes are already primed for germination, this occurs within a short time.

An analysis of the proteins produced by germinating and thermoinhibited achenes was conducted using SDS-polyacrylamide gel electrophoresis (SDS-PAGE). This procedure was able to resolve approximately 40 different protein bands, but no differences were observed between thermoinhibited and germinating achenes. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) was able to resolve approximately 200 individual polypeptides. The vast majority of polypeptides in *T. minuta* achenes are acidic, although the number of neutral to basic polypeptides increases as germination progresses. Ten polypeptides were identified which were specific to thermoinhibited achenes. These formed two distinct groups on the two-dimensional gels. The larger group contained seven proteins, ranging in size from 22 kDa to 26.7 kDa and with isoelectric points of between 3.0 and 4.0. The smaller group contained three polypeptides with molecular weights of about 14 kDa and a pI of approximately 3.0. These polypeptides were all extremely specific to thermoinhibited achenes and declined rapidly when the incubation temperature was reduced, in a manner which correlated with an increase in the germinability of the achenes. Several characteristics of the expression of these polypeptides were similar

to characteristics of embryo-dormancy in seeds where dormancy is thought to be actively imposed by the expression of specific dormancy-associated genes. This, along with the very tightly-regulated nature of these 10 polypeptides, suggests that thermoinhibition in *T. minuta* may be regulated through gene expression and that these ten polypeptides may represent the products of genes responsible for preventing radicle emergence at unfavourable temperatures.

Since these polypeptides were only resolved using silver-staining and could not therefore be used for amino acid sequence analysis, this hypothesis was further investigated using differential display of mRNA to isolate genes which are expressed specifically in thermoinhibited achenes. A large number of cDNA fragments which were specific to either germinating or thermoinhibited achenes were identified and extracted from the differential display gels. Those cDNAs specific to the thermoinhibited achenes were taken for further analysis. Of the 62 fragments excised from the gels, 25 could be reamplified to generate single bands of the correct size on agarose gels. A further 22 cDNAs produced multiple bands, where one band was much brighter than the others and correlated with the size of the original fragment. Thirteen of the cDNAs which generated single bands were cloned into the plasmid vector pGEM[®]-T Easy and transformed into either *Escherichia coli* JM109 or *E. coli* XL1-Blue. Recombinant colonies were identified using blue-white colour selection and the presence of the insert confirmed by colony blotting and restriction analysis. Three clones were chosen for each of the cDNAs. Reverse northern analysis confirmed that all 39 clones were specific to the mRNA pool of thermoinhibited achenes. High quality sequence data were obtained for 27 of the cDNA samples, the remainder appeared to have been degraded in transit. Alignment of the various sequences revealed that a total of 14 different sequences had been cloned, indicating that several of the bands isolated from the differential display gels contained multiple sequences. Electronic homology searches tentatively identified three of the sequences, whilst the remainder did not show significant homology to any known sequences. Of the cDNAs identified in this way, one may encode a plant transcription factor-like or nuclear RNA-binding protein whilst the other two may encode an RNase-L Inhibitor-like protein and a miraculin homologue. The potential

roles of such genes in the imposition or maintenance of the thermoinhibited state are discussed. Although further research needs to be conducted to isolate full length cDNA sequences and to determine their exact expression patterns in germinating and thermoinhibited achenes, these results are consistent with the hypothesis that thermoinhibition in *T. minuta* achenes is under positive genetic control in a manner analogous to embryo dormancy. This thesis represents the first molecular study of thermoinhibition as well as the first report of active control over this phenomenon in any species. Since thermoinhibition, unlike dormancy, can be rapidly imposed and released under strictly controlled conditions without the need for any dormancy breaking treatment, *T. minuta* achenes represent an excellent model system for studies on the molecular control of seed germination.

PUBLICATIONS FROM THIS THESIS

Hills, P.N., C.D. Viljoen and J. van Staden. 2001. Differences in polypeptide expression in thermoinhibited and germinating achenes of *Tagetes minuta* L. *Plant Growth Regulation* **34**:187-194

PUBLICATIONS RELATED TO THESIS

Hills, P.N. and J. van Staden. 2002. An improved DNA extraction procedure for plant tissues with a high phenolic content. *South African Journal of Botany* **68**: 549-550

Podd, L.A., P.N. Hills and J. van Staden. 2001. Physiological responses and extension of vase life of cut carnation flowers treated with ethanol and acetaldehyde. II. Protein content and enzyme activity. *Plant Growth Regulation* **38**: 107-117

CONFERENCE CONTRIBUTIONS FROM THIS THESIS

PAPERS AND POSTERS PRESENTED AT INTERNATIONAL CONFERENCES

Hills, P.N., J. van Staden and C.D. Viljoen. Changes in protein profiles associated with thermoinhibition in achenes of *Tagetes minuta* L. IXth International Congress on Plant Tissue and Cell Culture: *Plant Biotechnology and In Vitro Biology in the 21st Century*. Jerusalem, Israel, 14-19 June 1998. (Poster)

Hills, P.N. and J. van Staden. Differential gene expression in germinating and thermoinhibited *Tagetes minuta* L. achenes. Thirteenth Congress of the Federation of European Societies of Plant Physiology (FESPP). Crete, September 2002. (Paper)

Hills, P.N. and J. van Staden. Gene expression in thermoinhibited *Tagetes minuta* L. (Khakibos) achenes. Joint International Conference: South African Association of Botanists and International Society for Ethnopharmacology. Pretoria, South Africa, 7-11 January 2003. (Paper)

PAPERS AND POSTERS PRESENTED AT NATIONAL CONFERENCES

Hills, P.N., J. van Staden and C.D. Viljoen. Changes in protein profiles associated with thermoinhibition in achenes of *Tagetes minuta* L. (Khakibos). South African Association of Botanists 24th Annual Conference: *Plants, People and Progress into the 21st Century*. University of Cape Town, Cape Town, 12-16 January 1998. (Paper)

Hills, P.N., van Staden J. and Viljoen C.D. Changes in gene expression during thermoinhibition of *Tagetes minuta* L. (Khakibos) achenes. South African Association of Botanists 25th Annual Congress. University of Transkei (UNITRA), Umtata, South Africa, 11-15 January 1999. (Paper)

Hills, P.N., van Staden J. and Viljoen C.D. Thermoinhibition of *Tagetes minuta* L. (Khakibos) achenes causes changes in protein profiles. South African Association of Botanists 25th Annual Congress. University of Transkei (UNITRA), Umtata, South Africa, 11-15 January 1999. (Poster)

Hills, P.N. and van Staden, J. Protein involvement in the imposition and maintenance of thermoinhibition in *Tagetes minuta* achenes. South African Association of Botanists 26th Annual Conference: *Plant Biology for the New Millennium*. Potchefstroom, South Africa, 10-14 January 2000.

Hills, P.N. and van Staden, J. Gene transcription in thermoinhibited achenes of *Tagetes minuta* L. (Khakibos). South African Association of Botanists 27th

Annual Conference. Rand Afrikaans University (RAU), Johannesburg, South Africa, 14-17 January 2001.

Hills, P.N. and van Staden, J. Differential expression of genes in germinating and thermoinhibited achenes of *Tagetes minuta* L. (Khakibos). South African Association of Botanists 28th Annual Conference. Rhodes University, Grahamstown, 13-16 January 2002.

| | | |
|----------------------|---|----|
| 1.2.4.8 | Thermoinhibition in <i>Tagetes minuta</i> | 41 |
| 1.3 | AIMS AND OBJECTIVES | 42 |
| CHAPTER 2 | | |
| | GENERAL CHARACTERISTICS OF <i>Tagetes minuta</i> GERMINATION | 43 |
| 2.1 | INTRODUCTION | 43 |
| 2.2 | MATERIALS AND METHODS | 45 |
| 2.2.1 | Collection of plant material | 45 |
| 2.2.2 | Determination of moisture content | 45 |
| 2.2.3 | Germination trials | 46 |
| 2.2.4 | Imbibition curves | 46 |
| 2.2.5 | Freezing Trial | 47 |
| 2.3 | RESULTS AND DISCUSSION | 47 |
| 2.3.1 | Germination trials | 47 |
| 2.3.2 | Imbibition curves | 50 |
| 2.3.3 | Freezing trials | 52 |
| 2.3.4 | Effect of gibberellic acid on germination | 53 |
| 2.4 | CONCLUSION | 54 |
| CHAPTER 3 | | |
| | PROTEIN ANALYSIS | 56 |
| 3.1 | INTRODUCTION | 56 |
| 3.2 | MATERIALS AND METHODS | 58 |
| 3.2.1 | Protein extraction | 58 |
| 3.2.2 | Quantification of protein samples | 58 |
| 3.2.3 | Polyacrylamide gel electrophoresis of proteins | 59 |
| 3.2.3.1 | SDS-polyacrylamide gel electrophoresis (Laemmli system) | 59 |
| 3.2.3.2 | Tricine-SDS-polyacrylamide gel electrophoresis .. | 61 |
| 3.2.3.3 | Two-dimensional-polyacrylamide gel electrophoresis | 62 |
| 3.2.4 | Detection of electrophoresed proteins | 63 |

| | |
|--|----|
| 3.3 RESULTS AND DISCUSSION | 64 |
| 3.3.1 Optimization of protein extraction buffer | 64 |
| 3.3.2 Comparison of standard SDS-PAGE (Laemmli system) and Tricine-SDS-PAGE | 67 |
| 3.3.3 Staining of two-dimensional gels | 67 |
| 3.3.4 Single-dimension SDS-PAGE | 70 |
| 3.3.5 Two-dimensional polyacrylamide gel electrophoresis | 74 |
| 3.4 CONCLUSION | 85 |

CHAPTER 4

| | |
|---|-----------|
| NUCLEIC ACID ANALYSIS | 86 |
| 4.1 INTRODUCTION | 86 |
| 4.2 MATERIALS AND METHODS | 89 |
| 4.2.1 Isolation of RNA | 89 |
| 4.2.2 Analysis of RNA | 90 |
| 4.2.3 Differential display of mRNA | 91 |
| 4.2.3.1 Reverse transcription | 91 |
| 4.2.3.2 Amplification of cDNA by the polymerase chain reaction | 92 |
| 4.2.3.3 Labeling of molecular weight markers for use in differential display gels | 93 |
| 4.2.3.4 Electrophoresis of PCR products from differential display reactions | 94 |
| 4.2.4 Recovery and reamplification of differentially expressed cDNAs | 94 |
| 4.3 RESULTS AND DISCUSSION | 95 |
| 4.3.1 Isolation of total RNA | 95 |
| 4.3.2 Isolation of polyadenylated RNA | 101 |
| 4.3.3 Optimization of differential display reactions | 102 |
| 4.3.4 Analysis of changes in gene expression in thermoinhibited achenes of <i>Tagetes minuta</i> using differential display of mRNA | 105 |

| | |
|---|-----|
| 4.3.5 Reamplification of differentially expressed cDNAs | 109 |
| 4.4 CONCLUSION | 113 |

CHAPTER 5

| | |
|--|------------|
| CLONING AND ANALYSIS OF DIFFERENTIALLY EXPRESSED cDNAs | 115 |
| 5.1 INTRODUCTION | 115 |
| 5.2 MATERIALS AND METHODS | 122 |
| 5.2.1 Bacterial strains | 122 |
| 5.2.2 Preparation of competent <i>Escherichia coli</i> cultures | 122 |
| 5.2.3 Preparation of PCR-amplified cDNAs for cloning | 123 |
| 5.2.4 Ligation of cDNA fragments into the pGEM [®] -T EASY vector plasmid | 125 |
| 5.2.5 Transformation of <i>Escherichia coli</i> with the pGEM [®] -T EASY vector | 127 |
| 5.2.6 Colony hybridisation | 128 |
| 5.2.6.1 Preparation of replica filters | 128 |
| 5.2.6.2 Preparation of filters for hybridisation | 129 |
| 5.2.6.3 Preparation of radioactive probe DNA | 130 |
| 5.2.6.4 Hybridisation of filters with radioactive probe DNA | 130 |
| 5.2.6.5 Detection of the hybridisation signal and recombinant colonies | 131 |
| 5.2.7 Isolation of plasmid DNA | 131 |
| 5.2.8 Restriction analysis | 133 |
| 5.2.9 Preparation of freezer stocks | 133 |
| 5.2.10 Reverse-northern hybridisation analysis | 134 |
| 5.2.10.1 Dot-blotting of recombinant plasmids | 134 |
| 5.2.10.2 Preparation of probes for reverse-northern analysis | 135 |
| 5.2.10.3 Hybridisation | 136 |
| 5.2.11 Sequencing of cDNA clones | 136 |
| 5.3 RESULTS AND DISCUSSION | 137 |

| | |
|--|-----|
| 5.3.1 Transformation of competent cultures with the pGEM [®] -T | |
| EASY vector plasmid | 137 |
| 5.3.2 Reverse-northern blotting of cDNAs | 141 |
| 5.3.3 Sequencing of cDNAs and analysis of sequence data | 141 |
| 5.4 CONCLUSION | 163 |
| 5.5 GENERAL CONCLUSION | 164 |
| | |
| LITERATURE CITED | 167 |
| | |
| APPENDIX A | 180 |
| APPENDIX B | 188 |

LIST OF FIGURES

- Figure 1.1: External morphology of *Tagetes minuta* L. Mature plants may reach approximately 2-3 m in height in dense stands (A). All aerial parts of the plant, including the leaves (B) and the bracts enclosing the capitulae (C) are covered with orange oil glands which produce an abundance of essential oils. 2
- Figure 1.2: Morphology of inflorescence, capitulae and achenes of *Tagetes minuta* L. Densely packed terminal inflorescence (A). Detail of a pair of capitulae with oil glands visible on the fused bracts (B). Ray (top) and disc (bottom) florets (C). Enlargement of pappus, showing characteristic single long scale and several shorter scales. Note the golden hairs on the achene (D). Awl-shaped mature achenes (E). 3
- Figure 1.3: Model of genetic regulation of late embryogenesis in *Arabidopsis* based on mutant phenotypes *leafy cotyledon* and *ABA-insensitive*. Major events occurring during normal development are shown in the rectangles. These are not presented in sequential order. Arrows indicate which wild type genes are proposed to be required to complete each event (redrawn from MEINKE, FRANZMANN, NICKLE and YEUNG, 1994). 21
- Figure 1.4: Modified model of genetic regulation of late embryogenesis in *Arabidopsis*, encompassing data from all germination-related mutants. Major events occurring during normal development are shown in the rectangles. Black arrows indicate which wild type genes are proposed to be required to complete each event. *DAG1* (green arrow) is expressed in the mother plant and causes the induction of dormancy during late embryogenesis. *CTS* reduces the expression of the *LEC*-genes and *ABI3* and activates germination. The RDO loci appear to interact with the two different pathways, affecting dormancy imposition. Their exact placement within these two pathways is unknown, as the genes representing these two loci have not yet been isolated. Like *DAG1*, they probably function within the late embryogenesis programs, controlling dormancy induction (modified from MEINKE, FRANZMANN, NICKLE and YEUNG, 1994). 25
- Figure 1.5: Cypsela-type achene of *Lactuca sativa* (lettuce) showing entire fruit with

pappus still attached (A), longitudinal section (B), transverse section (C) and transection of part of the pericarp with subjacent layers (D) (ESAU, 1977). 31

Figure 1.6: Proposed model for temperature control of germination mediated by gibberellins and cytokinins (redrawn from BIDDINGTON and THOMAS, 1978). . . 34

Figure 2.1: Germination curves of *Tagetes minuta* achenes imbibed directly at 25°C (●) and following 72 h pre-treatment at 36°C (○). Vertical bar = maximum standard error of the mean. 48

Figure 2.2: Germination curves of *Tagetes minuta* achenes imbibed at 36°C for varying periods before transfer to an incubation temperature of 25°C. (●) 25°C control; (□) 24 h at 36°C; (Δ) 48 h at 36°C; (○) 72 h at 36°C; (◇) 96 h at 36°C; (☆) 120 h at 36°C. Vertical bar = maximum standard error of the mean. 50

Figure 2.3: Imbibition curves for *Tagetes minuta* achenes imbibed at 25°C (●) and at 36°C (○). Vertical bar = maximum standard error of the mean. 51

Figure 2.4: Germination curves of *Tagetes minuta* achenes stored at 10°C and -20°C when imbibed directly at 25°C and following 72 h imbibition at 36°C. (●), 10°C storage, no pre-treatment; (■) -20°C, no pre-treatment; (○), 10°C, 36°C pre-treatment; (□), -20°C storage, 36°C pre-treatment. Bar = maximum standard error of the mean. 53

Figure 3.1: Mean amounts of protein (mg/g FW) isolated from *Tagetes minuta* achenes imbibed at 25°C for varying periods of time. 66

Figure 3.2: Comparison of separation of proteins isolated from thermoinhibited achenes by Tricine-SDS-PAGE (a) and standard SDS-PAGE (Laemmli system) (b). Lane 1 contains proteins from dry achenes, lanes 2-7 contain proteins from achenes imbibed at 36°C for 24 h, 48 h, 72 h, 96 h and 120 h respectively. 68

Figure 3.3: Comparison of two-dimensional gels stained in the same manner as used for single dimension gels (A) and using the MORRISSEY (1981) method (B). 69

- Figure 3.4: SDS-PAGE separation of proteins isolated from dry achenes (lane 1), achenes incubated in a dry state at 36°C for 10 d (lane 2) and thermoinhibited achenes imbibed at 36°C for periods of 24 h (lane 3), 48 h (lane 4), 72 h (lane 5), 96 h (lane 6), 120 h (lane 7) and 240 h (lane 8). 71
- Figure 3.5: SDS-PAGE separation of proteins from germinating achenes imbibed for varying periods at 25°C. Lane 1 contains proteins isolated from dry achenes, lanes 2-9 contain proteins from achenes imbibed for 6 h, 12 h, 18 h, 24 h, 30 h, 36 h, 42 h and 48 h respectively. Levels of polypeptides labeled d1-d5 increase during germination, whilst levels of polypeptides i1-i4 decline during the germination process. 72
- Figure 3.6: SDS-PAGE separation of proteins from achenes germinated for varying periods following 72 h imbibition at 36°C. Lane 1 contains proteins from dry achenes. Proteins from thermoinhibited achenes (72 h at 36°C) are shown in lane 2. Lanes 3-14 contain proteins from pre-treated achenes which have been allowed to germinate at 25°C for between 2 h (lane 3) and 24 h (lane 14). Samples were taken at two-hourly intervals. 74
- Figure 3.7: Two-dimensional polyacrylamide gels showing changes in polypeptide patterns in germinating achenes imbibed at 25°C. Polypeptides were isolated from dry achenes (A) and achenes germinated for 12 h (B), 24 h (C) and 36 h (D). These three time intervals correspond to germination percentages of approximately 0%, 40% and 95% respectively. 76
- Figure 3.8: Changes in polypeptide patterns in achenes germinating at 25°C (A) and during thermoinhibition at 36°C (B). These diagrams were prepared from data obtained from the gels shown in Figures 3.7 and 3.9. Uncoloured polypeptides show no changes in expression when dry achenes are imbibed. Levels of polypeptides marked in green increase during imbibition, whilst levels of polypeptides marked in blue decline upon imbibition. Polypeptides marked in orange show a slight increase up to 12 h of imbibition at 25°C and subsequently decline. Polypeptides marked in yellow are not present in dry achenes and are synthesized *de novo* upon rehydration, at the times indicated on the diagram (12 h, 24 h or 36 h). The polypeptides marked in purple in (A), were specific to the dry achenes at 25°C, but are still present in thermoinhibited achenes,

| | |
|---|-----|
| albeit at a lower level. Polypeptides marked in red are specific to thermoinhibited achenes and may play a role in the imposition of thermoinhibition of these achenes. | 77 |
| Figure 3.9: Polypeptide patterns obtained by 2D-PAGE of protein isolates from achenes of <i>Tagetes minuta</i> which had been thermoinhibited (A) and subsequently allowed to germinate at 25°C for periods of 2 h (B), 6 h (C) or 20 h (D). Black arrowheads indicate polypeptides which are specific to thermoinhibited achenes. | 79 |
| Figure 3.10: 2D-PAGE separation of proteins from achenes incubated in a dry state at 36°C for 72 h. | 81 |
| Figure 4.1: RNA isolated from control (24 h imbibition at 25°C) achenes (lane 2) and thermoinhibited (72 h at 36°C) achenes (lane 3). Lane 1 contains the molecular weight marker used in the RNA studies (100 bp Ladder, Pharmacia Biotech). The 800 bp fragment of this marker is twice as intense as the other fragments for ease of recognition. | 98 |
| Figure 4.2: Comparison of amplified cDNA from control (24 h at 25°C) achenes (lanes 1-2, 5-6, labeled C) and thermoinhibited (72 h at 36°C) achenes (lanes 3-4, 7-8, labeled T). Note the poor amplification of the control samples. | 100 |
| Figure 4.3: RNA isolated from control (24 h imbibition at 25°C) achenes (lane 2) and thermoinhibited (72 h at 36°C) achenes (lane 3) using the modified guanidinium thiocyanate extraction method. Lane 1 contains molecular weight marker XIV (100 bp ladder, Roche). The 500 bp and 1000 bp fragments of this marker stain more intensely than the other fragments for ease of recognition. | 101 |
| Figure 4.4: Section of a differential display gel showing duplicate control (C) and thermoinhibited (T) samples amplified from the anchored primer T ₁₂ VG and random primers OPJ-03 (left) and OPJ-04 (right). Black arrowheads represent reproducible cDNA fragments which are specific to the thermoinhibited achenes. | 106 |
| Figure 4.5: Reamplification of differentially expressed cDNAs using a dNTP concentration | |

| | |
|--|-----|
| of either 2 μ M (lanes 2-3) or 20 μ M (lanes 4-5). Lane 1 contains molecular weight marker XIV (Roche). | 110 |
| Figure 4.6: Reamplification of four differentially expressed cDNAs showing single (lanes 3-4, 7-8) or multiple (lane 1-2) band synthesis or no reamplification (lanes 5-6). In all cases where multiple bands were generated following reamplification, the brightest band corresponded in size to the original fragment on the differential display gel. As with the differential display reactions, all reamplification reactions were carried out in duplicate and were electrophoresed in adjacent lanes to ensure reproducibility. | 111 |
| Figure 4.7: Effect of changing the amount of <i>Taq</i> DNA polymerase used in the reamplification of differentially expressed cDNAs. All reamplification was eventually conducted using 2.5 U of <i>Taq</i> polymerase per reaction (lanes 3-4). Reducing the amount of <i>Taq</i> polymerase to 1 U per reaction resulted in a significant decrease in the amount of cDNA produced (lanes 1-2), whilst increasing the amount of the polymerase in an effort to increase the amount of product resulted in the formation of non-reproducible bands (lanes 5-6). | 112 |
| Figure 5.1: pGem [®] -T EASY vector map and sequence reference points (Promega catalogue). | 126 |
| Figure 5.2: Colony blotting of recombinant colonies of <i>Escherichia coli</i> JM109 (A) and <i>E. coli</i> XL1 Blue (B). The <i>E. coli</i> JM109 clones in (A) were transformed with the pGEM [®] -T Easy vector with cDNA insert ATJ4-02, whilst the <i>E. coli</i> XL1 Blue colonies (B) contain the vector with cDNA insert GTJ4-43. Whilst <i>E. coli</i> XL1 Blue did prove problematic with the colony blotting protocol used as non-recombinant colonies also generated a positive hybridization signal, some colonies produced a much stronger signal than others and were selected for further analysis. Restriction analysis confirmed the presence of a recombinant vector in these colonies. | 141 |
| Figure 5.3: Negative strand nucleotide sequences of similar 163 bp cDNA inserts GTJ4-43/39 and GTJ4-43/47 as sequenced from the pUC/M13 reverse sequencing primer on the pGEM [®] -T EASY vector. The size of the insert as well as the primers | |

from which it was generated are given in parentheses. Primer sequences are highlighted. Both inserts are present in the same orientation. 142

Figure 5.4: Negative strand nucleotide sequences of similar 249 bp cDNA inserts ATJ4-02/1, CTJ4-27/30, ATJ4-02/2 and ATJ4-02/3 as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. Inserts ATJ4-02/1 and CTJ4-27/30 are cloned in the opposite orientation to ATJ4-02/2 and ATJ4-02/3. 143

Figure 5.5: Negative strand nucleotide sequences of similar 267 bp cDNA inserts TTJ6-53/2, TTJ6-53/3, TTJ6-53/50, ATJ6-01/38 and ATJ6-01/48 as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. Insert TTJ6-53/2 is present in the opposite orientation to the other sequences. 144

Figure 5.6: Negative strand nucleotide sequences of similar 347 bp cDNA inserts CTJ4-24/15 and CTJ4-24/43 as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. Both inserts are in the same orientation. 145

Figure 5.7: Negative strand nucleotide sequences of similar 154 bp cDNA inserts GTJ10-54/1, GTJ10-54/2 and GTJ10-54/3 as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. All inserts are present in the same orientation. . . . 145

Figure 5.8: Negative strand nucleotide sequences of similar 318 bp cDNA inserts GTJ10-52/17, GTJ10-52/25 and GTJ10-52/33 as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. All inserts are present in the same orientation. . . . 146

Figure 5.9: Negative strand nucleotide sequences of cDNA inserts, as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector, which did not align with any other sequences. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. 147

Figure 5.10: Negative strand nucleotide sequences of cDNA insert GTJ4-43/2, as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector, which did not align with any other sequences. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. This insert contains an oligo(dT) section (highlighted in the centre of the sequence) which could have resulted from the ligation of two different sequences, both with T₁₂VG ends and an A/T tail from PCR, in opposite orientations. 148

Figure 5.11: Consensus sequences (positive strand) for the three sequence groups which showed some significant homology to known genes. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. 158

LIST OF TABLES

Table 1.1: Countries in which *Tagetes minuta* is regarded as a weed species (HOLM, PANCHO, HERBERGER and PLUCKNETT, 1979). 4

Table 1.2: Some of the secondary metabolites found in the oil of *Tagetes minuta* L. (RODRÍGUEZ and MABRY, 1977). 7

Table 3.1: Details of proteins from the Combithek Calibration Proteins kit (Roche) and volumes used in the preparation of the molecular weight marker mixture for polyacrylamide gel electrophoresis. This kit is no longer marketed. 60

Table 3.2: Composition of stacking and separating gels for Tricine-SDS-PAGE (SCHÄGGER and VON JAGOW, 1987). 61

Table 4.1: Random decamer primers used for PCR amplification of cDNA fragments generated by reverse transcription of RNA isolated from *Tagetes minuta* achenes. 92

Table 4.2: Random decamer primers tested during optimization of PCR amplification for differential display. 104

Table 4.3: Summary of the numbers of differentially expressed cDNAs synthesized from total RNA isolated from *Tagetes minuta* achenes subjected to the control (24 h at 25°C) treatment using various combinations of primers. 107

Table 4.4: Summary of the numbers of differentially expressed cDNAs synthesized from total RNA isolated from *Tagetes minuta* achenes subjected to thermoinhibitory conditions (72 h at 36°C) using various combinations of primers. 108

Table 5.1: Amount of DNA present in each of the marker bands for 1 µl (250 ng) DNA Molecular Weight Marker III (Roche) as used for estimation of concentration of purified cDNA fragments. For regression analysis, the 3530 bp band was disregarded and the 5148 and 4973 bp bands were combined for a total of

| | |
|---|-----|
| 52.1 ng. | 125 |
| Table 5.2: Ligation reactions for cloning cDNA fragments into the pGEM [®] -T EASY vector. | 127 |
| Table 5.3: List of cDNAs cloned into the pGEM [®] -T EASY vector and their approximate sizes. | 140 |
| Table 5. 4: Alignment groups for sequences grouped by homology following sequencing. | 149 |
| Table 5.5: Summary of results from BLASTN and BLASTX searches for sequences showing homology to the thermoinhibition-associated cDNAs isolated from <i>Tagetes minuta</i> achenes. Only the three highest scoring alignments are listed for each search, in descending order of significance. | 151 |

ABBREVIATIONS

| | |
|---|--|
| 2D-PAGE | Two-dimensional polyacrylamide gel electrophoresis |
| °C | Degrees celsius |
| Ψ | Water potential |
| Ψ _π | Osmotic potential |
| μg | Microgram |
| μl | Microlitre |
| μm | Micrometre |
| μmol photons m ⁻² .sec ⁻¹ | Micromole photons per square metre per second |
| A ₂₆₀ | Absorbance at 260 nm |
| A ₂₈₀ | Absorbance at 280 nm |
| A ₃₂₀ | Absorbance at 320 nm |
| A | Adenine/adenosine |
| ABA | Abscisic acid |
| ACC | 1-aminocyclopropane-1-carboxylic acid |
| APS | Ammonium persulfate |
| AVG | Aminoethoxyvinylglycine |
| BA | Benzyladenine |
| <i>bis</i> acrylamide | N'N' methylene <i>bis</i> acrylamide |
| bp | Base pair |
| BSA | Bovine serum albumin |
| C | Cytosine |
| CA | California |
| cDNA | Complementary deoxyribonucleic acid |
| cfu | Colony forming units |
| CHA | Cyclohexamine |
| Ci | Curie |
| cm | Centimetre |
| CO ₂ | Carbon dioxide |
| cv | Cultivar |

| | |
|--------------------|---|
| d | Day |
| dATP | Deoxyadenosine triphosphate |
| dCTP | Deoxycytidine triphosphate |
| ddH ₂ O | Deionized distilled water |
| DEPC | Diethyl pyrocarbonate |
| dGTP | Deoxyguanosine triphosphate |
| dH ₂ O | Distilled water |
| dITP | Deoxyinosine triphosphate |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| DNase | Deoxyribonuclease |
| dNTP | Deoxynucleoside triphosphate |
| DTT | Dithiothreitol |
| dTTP | Deoxythymidine triphosphate |
| E | East |
| EDTA | Ethylene diamine tetraacetic acid |
| EFE | Ethylene-forming enzyme |
| EGTA | Ethylene glycol <i>bis</i> (β-aminoethyl ether) NNN'N'- tetraacetic acid |
| Fluoridone | 1-methyl-3-phenyl-5-[3-trifluoromethyl-(phenyl)]-4-(1 <i>H</i>)- pyridinone |
| FW | Fresh weight |
| <i>g</i> | Gravitational acceleration (9.806 m.sec ⁻¹) |
| g | Gram |
| G | Guanine |
| G | Single letter code for the amino acid glycine |
| GA | Gibberellic acid |
| Glc-6-P | Glucose-6-phosphate |
| h | Hour |
| hnRNA | heterogenous nuclear ribonucleic acid (pre-mRNA) |
| hnRNP | heterogenous nuclear ribonucleoprotein particle |
| HPLC | High-performance liquid chromatography |

| | |
|--------------------------|---|
| IEF | Isoelectric focussing |
| IPTG | Isopropyl-1-thio- β -D-galactoside |
| kb | Kilobase pairs |
| kDa | Kilodalton |
| K_m | Michaelis constant |
| l | Litre |
| LEA | Late Embryogenic Abundant |
| LFR | Low fluence response |
| M | Molar |
| mA | Milliamperes |
| mM | Millimolar |
| mg | Milligram |
| MGBG | Methyl-glyoxal- <i>bis</i> -guanylylhydrazone |
| min | Minute |
| ml | Millilitre |
| mm | Millimetre |
| mMol | Millimole |
| M-MuLV | Moloney-Murine Leukemia Virus |
| M_r | Relative molecular weight |
| mRNA | Messenger ribonucleic acid |
| ng | Nanograms |
| nm | Nanometre |
| NP40 | Nonidet P-40 |
| OD | Optical density |
| PCR | Polymerase chain reaction |
| PHS | Pre-harvest sprouting |
| pl | Isoelectric point |
| PIPES | Piperazine- <i>N,N'</i> - <i>bis</i> (2-ethane-sulfonic acid) |
| PMSF | Phenylmethylsulfonyl fluoride |
| Poly(A) ⁺ RNA | Polyadenylated ribonucleic acid |
| ppm | Parts per million |
| PVPP | Polyvinylpolypyrrolidone |

| | |
|----------|---|
| R | Single letter code for the amino acid arginine |
| RGG | Arginine-glycine-glycine motif for protein RNA binding |
| RNA | Ribonucleic acid |
| RNase | Ribonuclease |
| rpm | Revolutions per minute |
| RRM | RNA recognition motif for RNA-protein binding |
| rRNA | Ribosomal ribonucleic acid |
| RT | Reverse transcription |
| S | South |
| Sarkosyl | <i>N</i> -laurosarcosine |
| sec | Second |
| SDS | Sodium dodecyl sulfate |
| SDS-PAGE | Sodium dodecyl sulfate-polyacrylamide gel electrophoresis |
| ssDNA | Single-stranded DNA |
| T | Thymine |
| TAE | Tris/acetate electrophoresis (buffer, see Appendix) |
| TBE | Tris/borate electrophoresis (buffer, see Appendix) |
| TE | Tris/EDTA (buffer, see Appendix) |
| T_m | Melting temperature |
| Tris | 2-amino-2(hydroxymethyl)-propane-1,3-diol (Tris[hydroxymethyl] aminomethane) |
| Tris-Cl | Tris hydrochloride |
| tRNA | Transfer ribonucleic acid |
| Temed | NNN'N' tetramethylethylenediamine |
| U | Unit |
| V | Volt |
| v/v | Volume per volume |
| W | Watt |
| w/v | Weight per volume |
| X | Times |
| X-gal | 5-bromo-4-chloro-3-indolyl- β -D-galactoside |

CHAPTER 1

LITERATURE REVIEW

1.1 *Tagetes minuta* L.

Tagetes minuta L. belongs within the tribe Helenieae of the Asteraceae. It is a strongly scented, herbaceous species which varies in height from 5 cm to 3 m (Figure 1.1A) (HILLIARD, 1977). Plant height is related to the conditions under which the plants are growing. Single plants growing in the open usually grow to between 0.5 m and 1 m in height. In dense stands, however, plants generally grow much taller, averaging about 2 m in height (SOULE, 1993). All aerial parts of the plant have elongated, embedded oil glands (Figure 1.1 B, C), which produce an abundance of aromatic essential oils (HILLIARD, 1977).

The leaves may be either alternately or oppositely arranged (HILLIARD, 1977). They are a slightly glossy green, up to 10 cm in length and are pinnately dissected into four to six pairs of pinnae (SOULE, 1993). The lanceolate pinnae have dimensions of about 5 X 0.5 cm with sharply serrate margins. Leaf bases are half-clasping and have several filiform segments (HILLIARD, 1977).

Twenty to eighty capitulae are clustered into terminal corymbose panicles (Figure 1.2 A) (HILLIARD, 1977; SOULE, 1993). The flowering heads (Figure 1.2 B) are small, being approximately 10 mm to 15 mm long and having a diameter of between 10 mm and 20 mm. The involucre is narrowly cylindrical and is composed of three to five fused bracts which are 10 mm to 12 mm long (HILLIARD, 1977). These bracts are dull yellow in colour and are streaked with embedded oil sacs (Figure 1.1 C, Figure 1.2 B) (HILLIARD, 1977). Typically the capitulae contain four to five yellow-orange ray florets and ten to fifteen similarly coloured disc florets (Figure 1.2 C) (SOULE, 1993). In South Africa, flowering occurs between February and June (HILLIARD, 1977).

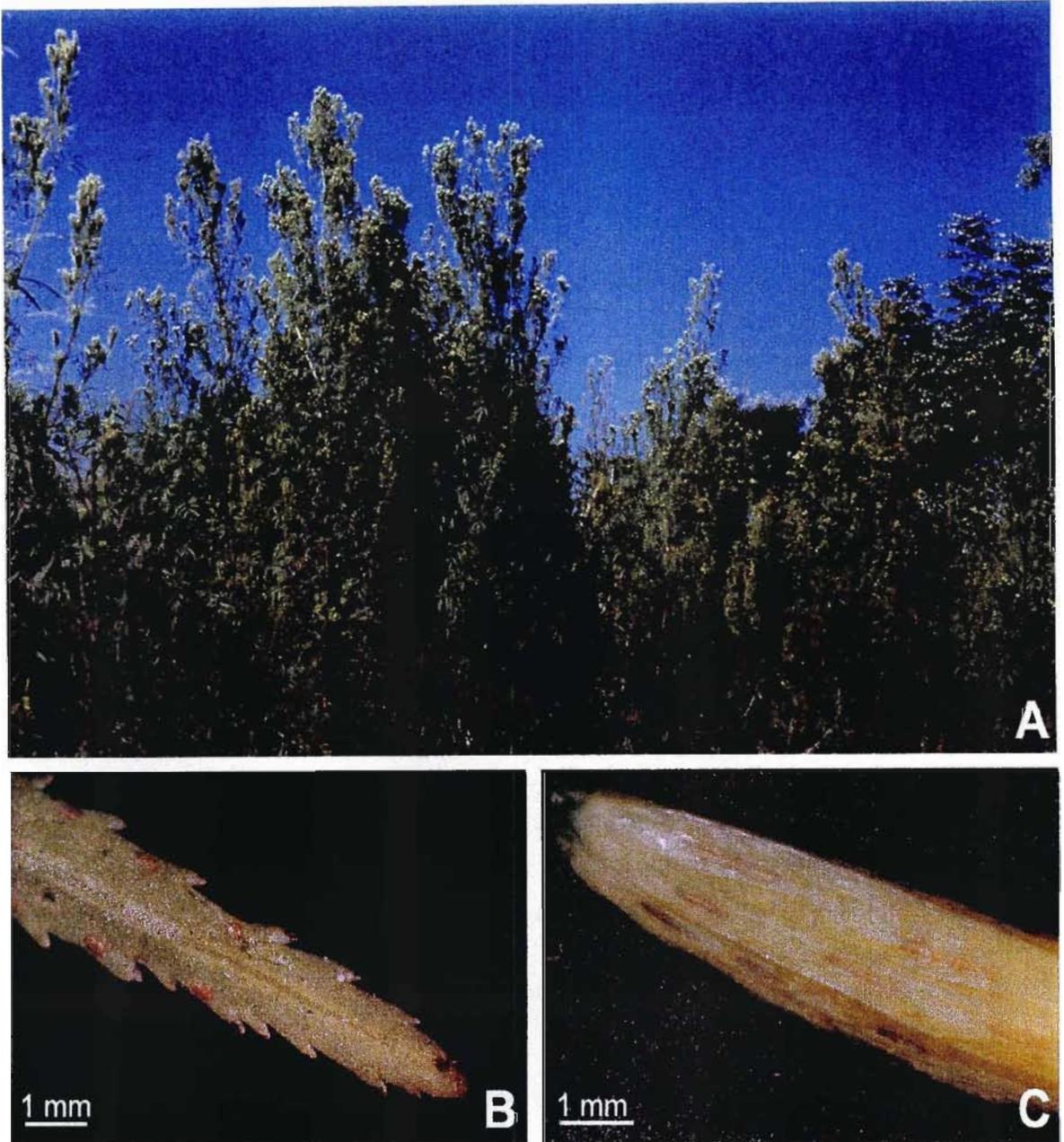


Figure 1.1: External morphology of *Tagetes minuta* L. Mature plants may reach approximately 2-3 m in height in dense stands (A). All aerial parts of the plant, including the leaves (B) and the bracts enclosing the capitulae (C) are covered with orange oil glands which produce an abundance of essential oils.

The achenes are approximately 10 to 12 mm in length and are awl-shaped. They are dark brown to black in colour with appressed golden hairs (Figure 1.2 D, E) (HILLIARD, 1977). The pappus is composed of one to four lanceolate-acuminate scales, one of which is considerably longer than the others (Figure 1.2 D) (HILLIARD,

Figure 1.2: Morphology of inflorescence, capitulae and achenes of *Tagetes minuta* L. Densely packed terminal inflorescence (A). Detail of a pair of capitulae with oil glands visible on the fused bracts (B). Ray (top) and disc (bottom) florets (C). Enlargement of pappus, showing characteristic single long scale and several shorter scales. Note the golden hairs on the achene(D). Awl-shaped mature achenes (E).



A



B



C



D



E

1977; SOULE, 1993).

Tagetes minuta is native to the temperate grasslands and montane regions of South America, including Argentina, Bolivia, Chile, Peru and the Chaco region of Paraguay (SOULE, 1993). The species is often found in disturbed sites during early successional stages (SOULE, 1993). It is now, however, extremely widespread and has been introduced into Europe, Asia, Africa, Madagascar, India, Australia and Hawaii (SOULE, 1993). A list of countries in which *T. minuta* is regarded as a weed species is provided in Table 1.

Table 1.1: Countries in which *Tagetes minuta* is regarded as a weed species (HOLM, PANCHO, HERBERGER and PLUCKNETT, 1979).

| S | P | C | X |
|--------------|-----------|-----------|--------------------------|
| Brazil | Argentina | Australia | Angola |
| Guatemala | Guyana | Ethiopia | Botswana |
| South Africa | Kenya | Hawaii | Chile |
| Tanzania | Swaziland | Peru | Egypt |
| Zambia | Uganda | | New Zealand |
| | Zimbabwe | | Thailand |
| | | | Uruguay |
| | | | United States of America |

S: Serious weed

P: Principal weed

C: Common weed

X: Regarded as a weed species, but status unknown

The earliest record of *T. minuta* in South Africa is at Vlakfontein in 1905 (HILLIARD, 1977). The common names "Khaki weed" and "Khakibos" are derived from its supposed introduction by the Khaki-clad British soldiers in the hay for their horses (HILLIARD, 1977).

In South Africa, *T. minuta* is regarded as a problem weed for maize farming (MALAN,

VISSER and GROBBELAAR, 1981). A problem weed is defined as a weed which is not effectively controlled through normal weed eradication procedures such as spraying with herbicides and mechanical operations (MALAN, VISSER and GROBBELAAR, 1981). The most severe infestations of this species occur during autumn. The weed plants often grow taller than the crop plants and impede harvesting operations (MALAN, VISSER and GROBBELAAR, 1981). The presence of achenes of *T. minuta* in the grain can also cause the grain to have an unpleasant odour, which results in a lower grading (GRABANDT, 1985). Furthermore, the roots of *T. minuta* produce the secondary metabolite α -terthienyl which is released directly into the soil by exudation or through decay of plant matter (MEISSNER, NEL and BEYERS, 1986). This compound has been extracted from the rhizosphere in concentrations of 0.4 ppm (parts per million), which is sufficient to inhibit the growth of other seedlings. When fields which have been overgrown with *T. minuta* are cultivated, the growth of the follow-on crop is often inhibited (MEISSNER, NEL and BEYERS, 1986). These authors showed that aqueous extracts of *Tagetes*-infested soils did not significantly inhibit germination in carrot, cucumber, lettuce, onion, radish, squash and tomato seeds, but that these extracts caused a marked retardation in the subsequent development of the seedlings, especially with regard to radicle elongation. A similar result was observed when the seeds were sown in potted soil which had been heavily infested during the previous season. In all cases except for sunflower seedlings, the dry mass of the top growth was severely reduced and plant height was reduced (MEISSNER, NEL and BEYERS, 1986).

Not all characteristics of *T. minuta* are negative, however. The species has been used by the people in the Americas as a flavourful beverage, a medicinal tea and a condiment since before the discovery of the New World by European explorers (SOULE, 1993). Local names vary according to the region, some of the most common being "chinchilla", "chilquila", "chilca", "zuico" and "suico". The beverage is prepared by steeping half a handful of the dried plant in hot water for between three and five minutes. The resulting concoction is then either drunk hot or cold and is sweetened according to taste (SOULE, 1993). For medicinal purposes, a double handful is used. This tea is used as a cure for the common cold, inflammations of

the upper and lower respiratory tract, upset stomach, diarrhoea and liver ailments (SOULE, 1993). In Chile and Argentina, *T. minuta* is popular as a seasoning, particularly for rice dishes and adding flavour to stews (SOULE, 1993).

In Latin America, many of the smaller-scale farmers leave plants of *T. minuta* in their fields. This “second crop” has several benefits to the farmers. Firstly, the rapid growth of these plants results in the shading out other weed species which are less beneficial to the farmer. Secondly, they may later be harvested for personal use or sale at markets, and thirdly, this practice has been reported to aid in the retention of humidity in the field (SOULE, 1993).

Tagetes minuta is also grown as a commercial crop, chiefly for the essential oils and wide variety of secondary metabolites which the species produces. The oils are widely used in the perfume and flavour industries as “*Tagetes* oil”. This is used as a flavour component in most major food products, including cola and alcoholic beverages, frozen dairy desserts, candy, baked goods, condiments and relishes (SOULE, 1993). Worldwide production in 1984 was estimated at 1.5 million tonnes, with Brazil being the major producer (SOULE, 1993).

Tagetes minuta is extremely rich in a number of secondary compounds, including acyclic, monocyclic and bicyclic monoterpenes, sesquiterpenes, flavonoids, thiophenes and aromatics (Table 2, RODRÍGUEZ and MABRY, 1977). Several of these compounds are effective deterrents of many organisms, including fungi, bacteria, roundworms, trematodes, nematodes and numerous insect pests (SOULE, 1993). This has led to interest in the species as a natural alternative to chemical pesticides and in South Africa, extracts of the plants are relatively widely used as a remedy against fleas.

Table 1.2: Some of the secondary metabolites found in the oil of *Tagetes minuta* L. (RODRÍGUEZ and MABRY, 1977).

| Acyclic monoterpenes | Monocyclic monoterpenes | Bicyclic monoterpenes | Sesquiterpenes |
|----------------------|-------------------------|-----------------------|----------------|
| tagetone | limonene | α -pinene | eudesmol |
| linalool | α -phellandrene | β -pinene | aromadendrene |
| ocimene | ρ -cymene | camphene | |
| citrol | carvone | sabinene | |
| myrcene | α -terpineol | | |

1.2 SEED GERMINATION

The seed is the product of a fertilized ovule and is the prime sexual dispersal unit of the spermatophytes (BRADBEER, 1988). Through the production of this structure the plant is able to achieve four objectives (BRADBEER, 1988). The seed functions as a mechanism for the multiplication of the species, allows for genetic re-assortment through the fusion of the male and female gametes, aids in dispersal of the species and may also function in some cases as a survival mechanism, since desiccated seeds usually display greater tolerance to adverse environmental conditions than do mature plants. The embryo contained within the seed usually does not germinate until it has undergone a period of development on the parent plant and has become air dry. During this process, the metabolic activity of the seed is considerably decreased and it becomes quiescent.

The primary function of any surviving seed is to germinate and establish a new individual of the species. The term germination is often loosely and incorrectly used. Germination commences with the uptake of water by the seed (imbibition) and ends with the start of the elongation of the embryonic axis. This usually begins with the elongation of the radicle (BEWLEY and BLACK, 1986). A number of events occur during this period, including protein hydration, sub-cellular structural changes, respiration, the synthesis of a multitude of macromolecules and cellular elongation

(BEWLEY and BLACK, 1986). In the strictest sense of this definition of germination, seedling growth is not included as a part of the germination process. BRADBEER (1988), however, includes seedling growth as a part of germination and holds the germination process to be complete once all the available energy reserves of the seed have been depleted and the seedling is capable of an independent existence. For the purposes of this study, germination will be considered to be complete once the radicle has emerged.

Three basic requirements have to be met in order for a seed to be able to germinate, these being the availability of sufficient moisture, a suitable temperature and, usually, an aerobic atmosphere (BRADBEER, 1988). Should any of these criteria not be met, germination will not occur. It is obvious then that these factors are critical to the survival of any given plant species.

Dormancy is the inability of a seed to germinate even when exposed to conditions of moisture, temperature and oxygen availability which are normally conducive to the later stages of germination and seedling establishment for that species (BRADBEER, 1988). Seeds which are dormant at the time at which they are shed from the parent plant have primary dormancy (EGLEY, 1995). Should this primary dormancy be broken and conditions then become unsuitable for germination, the seed may enter into a state of secondary dormancy (EGLEY, 1995). Seeds which are released in a non-dormant state may also enter into a secondarily dormant state should conditions become unfavourable for germination. Secondary dormancy should not, however, be confused with enforced dormancy (pseudodormancy), where germination is prevented as a result of passive inhibition or due to the lack of required conditions (EGLEY, 1995).

1.2.1 Physiology of seed germination

Water uptake by seeds follows an imbibition curve which is divided into three distinct phases. Initially there is a rapid uptake of water, which is followed by a lag phase of

slower progress in hydration. The final stage is an acceleration of seed hydration in comparison with the lag phase (BEWLEY and BLACK, 1994). Stages I and II occur during germination, whilst phase III is a post-germinative event and occurs during the elongation of the embryonic axis (BEWLEY, 1997). The initial imbibition phase is a purely physical process in the embryo axis and cotyledons which does not require any action on the part of the seed and which will also occur in seeds which have been killed (OBROUCHEVA and ANTIPOVA, 1997). The predominant forces governing this stage of water uptake are matric forces (OBROUCHEVA and ANTIPOVA, 1997). The second phase is governed by osmotic water uptake and is a physiological process which depends upon the accumulation of osmotically active solutes by the axial tissues (OBROUCHEVA and ANTIPOVA, 1997). Phase III results from cellular expansion and vacuolation in both the growing axis and the cotyledons (OBROUCHEVA and ANTIPOVA, 1997). With rehydration comes a reactivation of the basic metabolic processes in the seed, with water acting as a trigger for the various reactivation processes involved in seed germination. This trigger concept suggests that reaching specific hydration threshold levels switches processes from an "off" state to the "on" state (OBROUCHEVA and ANTIPOVA, 1997).

The respiration rate of the seed is progressively enhanced with an increase in its hydration status. At approximately 20% hydration (as a percentage of fresh weight) there is a slow initial reactivation through glycolysis, the Krebs cycle and the cyanide-sensitive electron transport chain. (OBROUCHEVA and ANTIPOVA, 1997). Once 45-50% hydration is achieved there is a further dramatic increase in the respiration rate, possibly as a result of mitochondriogenesis (OBROUCHEVA and ANTIPOVA, 1997). Following this steep initial increase in oxygen consumption, the rate declines until radicle emergence, when another burst of respiratory activity occurs (BEWLEY, 1997). Respiratory metabolism in germinating seeds has been reviewed in great detail by BOTHA, POTGIETER and BOTHA (1992).

Membranes become stabilized at 20-25% hydration, preventing electrolyte leakage. Amino acid metabolism appears to be activated at 18-20% hydration. Synthesis of new mRNA transcripts appears to commence at between 42-47% hydration

(OBROUCHEVA and ANTIPOVA, 1997). Protein synthesis has been shown to occur at very low levels in quiescent seeds, although polysomes are absent in these dry seeds (OBROUCHEVA and ANTIPOVA, 1997). Polysome formation from monosomes appears to be stimulated at a hydration level of between 47% and 54% (OBROUCHEVA and ANTIPOVA, 1997). The number of single ribosomes then declines as they are incorporated into the polysomes (BEWLEY, 1997). Initial polysome formation is dependant on stored ribosomes, but new ribosomes appear to be synthesized within hours of initial polysomal assembly (BEWLEY, 1997). At threshold hydration levels of between 40% and 50%, starch hydrolysis begins and proteolysis is initiated between 50% and 55% (OBROUCHEVA and ANTIPOVA, 1997). These systems all appear to be invoked during the physical imbibition phase so that they are triggered automatically by water availability (OBROUCHEVA and ANTIPOVA, 1997). If water is removed and the seeds dry out, those processes which have been activated are stopped. When water becomes available again, metabolism proceeds from the point at which it was previously suspended (OBROUCHEVA and ANTIPOVA, 1997).

Once metabolism has been activated, the axial tissues are able to prepare for cell elongation (OBROUCHEVA and ANTIPOVA, 1997). This is, however, dependant on the second phase of imbibition and if hydration is suspended at the level of physical imbibition (60%), elongation does not occur (OBROUCHEVA and ANTIPOVA, 1997). Once hydration has proceeded beyond 60%, osmotically active substances such as sugars, amino acids and K^+ ions are accumulated, which allows for vacuolar enlargement and consequently further water uptake (OBROUCHEVA and ANTIPOVA, 1997). At about 68% hydration, acidification occurs in the cell walls, allowing for expansion (OBROUCHEVA and ANTIPOVA, 1997). The extension of the radicle through the structures surrounding the embryo generally marks the end of germination *per se*. Radicle extension is a turgor-driven process and is not necessarily accompanied by cell division (BEWLEY, 1997). Three possible reasons for radicle extension have been proposed (BEWLEY, 1997). Firstly, during germination the osmotic potential ($\Psi\pi$) of the radicle cells becomes more negative as a result of solute accumulation from hydrolysis of stored reserve polymers. This

decrease would lead to an increased water uptake and a consequent increase in turgor which would drive expansion of the cells. Secondly, increased extensibility of the cell walls themselves as a result of enzyme activity may allow for their elongation. The third possibility is that the tissues surrounding the tip of the radicle weaken, again possibly as a result of enzyme action, allowing the tip to penetrate through them (BEWLEY, 1997).

Imbibition takes place more slowly in the cotyledons and only proceeds to a level of about 55-60% (OBROUCHEVA and ANTIPOVA, 1997). The cotyledons only undergo the first stage of germination, the activation of basal metabolism, prior to radicle emergence (OBROUCHEVA and ANTIPOVA, 1997). At 40-50% hydration, respiration is fully activated, starch mobilization at 45% hydration and protein reserves are mobilised at about 52%, providing a pool of sugars and amino acids to the cotyledons (OBROUCHEVA and ANTIPOVA, 1997). This pool is then able to support respiration and is used for the synthesis of hydrolytic enzymes which then hydrolyze the bulk of the reserves. This process is induced by a hormonal cue following radicle emergence when phytohormones synthesized in the axis during imbibition are exported to the cotyledons (OBROUCHEVA and ANTIPOVA, 1997).

Whilst the importance of various plant growth regulators in dormancy breaking and relief of thermoinhibition has been established, germination of non-dormant seeds under favourable conditions seems to occur without the immediate participation of phytohormones. Plant growth regulators appear to play a role in the control of seed development and maturation, whereafter they drop to very low levels in dry seeds (OBROUCHEVA and ANTIPOVA, 1997). In a number of studies, inhibitors of biosynthesis of a variety of phytohormones do not appear to prevent radicle emergence (OBROUCHEVA and ANTIPOVA, 1997). Gibberellins have long been considered to play a major role in germination. In tomato (*Lycopersicon esculentum*), seeds of a GA-deficient mutant were unable to germinate without the application of exogenous GA (GROOT and KARSSSEN, 1987). However, removal of the micropylar region of the tissues surrounding the embryo resulted in germination without GA. Inhibitors of GA biosynthesis had no effect on pea (*Pisum sativum*) germination until

elongation of the epicotyl (SPONSEL, 1983). Thus the role of GA in tomato seed germination (and most likely in other "GA-requiring" seeds) may be limited to decreasing the mechanical resistance of the structures surrounding the embryo to allow for radicle emergence. Whilst some phytohormones or their precursors are synthesized in germinating seeds, the phytohormonal system only appears to develop after radicle emergence (OBROUCHEVA and ANTIPOVA, 1997).

The expression of the genes required for germination begins during early embryogenesis, approximately 15-20 days after pollination (OBROUCHEVA and ANTIPOVA, 1997). During seed maturation and desiccation, most of the mRNAs related to embryo development are degraded, whilst the transcripts for germination related processes are stored as complexes with proteins in either the cytoplasm or in the nucleus (OBROUCHEVA and ANTIPOVA, 1997). Although transcription does occur within germinating seeds, these stored mRNAs may be all that is required for early seed germination up to radicle emergence. While protein synthesis is critical to seed germination and inhibition of this process by inhibitors such as cyclohexamide and chloramphenicol prevents germination (WALTON and SOOFI, 1969; KLEIN, BARENHOLZ and BUDNIK, 1971; JENDRISAK, 1980), seeds are able to germinate to a certain extent in the presence of transcription inhibitors such as actinomycin D and α -amanitin (WALTON and SOOFI, 1969; KLEIN, BARENHOLZ and BUDNIK, 1971; JENDRISAK, 1980; DATTA, MARSH and MARCUS, 1983). Low concentrations of α -amanitin inhibit RNA polymerase II, which is the polymerase responsible for the synthesis of heterogenous nuclear RNA (hnRNA), the precursor to mRNA, while higher concentrations of the compound also inhibits RNA polymerase III, responsible for the synthesis of tRNA precursors and 5 S rRNA (JENDRISAK, 1980). Actinomycin D is a chromopeptide which forms nonionic complexes with the DNA to prevent the action of the RNA polymerases (JENDRISAK, 1980).

The various authors do, however, interpret these transcription inhibitor studies differently. In isolated wheat embryos, low concentrations (0.1 to 1.0 $\mu\text{g/ml}$) of α -amanitin inhibited polyadenylated RNA synthesis and embryo growth, suggesting that *de novo* mRNA transcription is required for germination (JENDRISAK, 1980). A

similar result was obtained by DATTA, MARSH and MARCUS (1983), who suggested that rather than implying that transcription is necessary for germination, these results showed that the early stages of germination are transcription-independent. Indeed, in the earlier experiments of JENDRISAK (1980), when embryos exposed to α -amanitin showed an inhibition of mRNA synthesis of 80%, embryo growth was reduced by only 20%. This discrepancy in the analysis of these results may be due to a difference in the time-scale of the experiments. JENDRISAK (1980) ran the experiments over a period of 40 h, whilst DATTA, MARSH and MARCUS (1983) conducted their studies over 24 h. These latter researchers also noticed an effect on growth after approximately 17 h of imbibition. Actinomycin D caused a reduction in fresh weight gain in imbibing excised *Phaseolus vulgaris* embryos (WALTON and SOOFI, 1969), as well as a 60-70% reduction in growth in excised lima bean (*Phaseolus lunatus*) axes (KLEIN, BARENHOLZ and BUDNIK, 1971). JENDRISAK (1980) found that actinomycin D was considerably less effective at inhibiting embryo growth than α -amanitin, being effective only at high concentrations (greater than 25 $\mu\text{g/ml}$). The effectiveness of this compound on transcription in seeds may be reduced due to poor penetration of the drug into the embryo tissues (JENDRISAK, 1980).

In all of the inhibitor studies mentioned above, whether seeds were treated with actinomycin D or α -amanitin, some reduction in germination or embryo growth was observed after longer periods of imbibition. Unfortunately, since all of these studies used excised embryos and determined embryo growth only as an increase in fresh weight over time, the timing of these processes in relation to radicle emergence could not be determined. Thus, germination in its strictest sense may indeed be able to take place using only stored mRNA transcripts, with newly transcribed mRNAs becoming necessary at a slightly later stage for post-germinative events related to seedling establishment. This postulate appears to be supported by the fact that proteins which have been identified as being produced from newly transcribed mRNAs prior to radicle emergence, such as phytochrome, oxalate oxidase (germin) and polyamine oxidase, all function in the axial tissues after radicle emergence has taken place (OBROUCHEVA and ANTIPOVA, 1997). As time progresses, stored

transcripts are replaced with newly synthesized identical mRNAs and protein synthesis becomes increasingly dependant on the new transcripts (BEWLEY, 1997).

1.2.2 Molecular biology of seed germination

Besides a basic knowledge of gross physiological changes that occur in seeds following the application of water and the onset of imbibition, as described above, very little is known about what happens at a molecular level to initiate and control seed germination. Often, the knowledge that has been gained is confused by varying perceptions of what constitutes germination. From an agricultural point of view, germination is often considered to be the period from the sowing of the seed until the emergence of a vigorous seedling (BOTHA, POTGIETER and BOTHA, 1992). When this standard is applied, the time of radicle emergence is usually not recorded. Thus, when the strictest definition of germination is followed, where germination is considered to be complete once the radicle has ruptured the testa, it cannot be determined which processes are a part of germination and which represent seedling establishment. Literature dealing with reserve mobilization, which by strict definition is a post-germinative event, is often a good example of this problem (BOTHA, POTGIETER and BOTHA, 1992).

Seeds, by their very nature, are difficult to work with. Often they are small and large numbers must be used for physiological experiments. They are composed of a number of different tissues, each tissue being involved with different processes. Within the testa, the seed may be divided into storage tissues (the endosperm or cotyledons) and the embryonic axis (radicle and plumule). These two tissue types often have different metabolic patterns. When whole seed studies are conducted, all metabolic activity is averaged over the whole seed and the metabolic reactions in the various tissues cannot be differentiated (BOTHA, POTGIETER and BOTHA, 1992). Dissection of the seed into the various tissues immediately changes the environment surrounding those tissues and prevents any interactions which may have taken place in the whole seed, resulting in alterations to the metabolism and

functioning of that tissue (BOTH A, POTGIETER and BOTH A, 1992). A second major contributing factor to our poor understanding of germination is that germinating seeds have a very slow metabolic rate, making it difficult to label intermediates from exogenously applied compounds. Added to this is the slow penetration of compounds into the embryonic axis through the enclosing structures. These two factors often result in a gradient of label from the outside of the seed inward towards the axis, which means that some labelled intermediates may be cycled and recycled in some tissues before they have even reached other sites within the seed (BOTH A, POTGIETER and BOTH A, 1992). Again, separation of the various structures by dissection immediately has undefined and important changes on the metabolism of the seed parts. This problem has long been a limiting factor for germination studies using conventional approaches. The recent advances in molecular biology do, however, open up new approaches for understanding the metabolic processes occurring within germinating seeds through transgenic methods of gene manipulation, allowing changes to be made at the level of individual reactions in individual tissues and observing their effect on the germination process.

The vast majority of molecular studies undertaken focus either on seed maturation and the deposition of storage reserves, or on seedling establishment. This is probably an economically driven trend, as seeds form the major food sources for both humans and livestock. These aspects of seed molecular biology have been reviewed by BEWLEY and MARCUS (1990). Whilst embryogenesis and germination are distinct and separate processes, which, under normal circumstances are mutually exclusive, the latter is not entirely autonomous of the former process. Many of the factors which regulate germination are set up during late embryogenesis. Indeed, it is during this stage that all of the genes identified to date as having an effect on germination are expressed. Most of these relate to the inhibition of germination during the later stages of embryogenesis, when the seed has already matured to the point of being capable of germinating but before it has attained desiccation tolerance or dormancy, or to the imposition and regulation of primary embryo dormancy. Many of these genes appear to encode regulators of gene expression and may control the embryogenic programs, including inhibition of precocious germination, by ensuring

the timeous expression of various program-related genes. Studies pertaining to the genes involved in germination from the time at which rehydration begins up to the point of radicle emergence are much more difficult to find and the fundamental metabolic reactions necessary for germination, as well as the genes involved, have yet to be identified.

Embryogenesis is typically seen as consisting of three main stages: early embryogenesis, during which period the morphological patterns of the seed are laid out; the green cotyledonary stage, during which the embryos enlarge and storage reserves are accumulated; and late embryogenesis, when the seeds become quiescent, desiccation tolerant and during which period primary dormancy is instituted (KEITH, KRAML, DENGLER and McCOURT, 1994). Germination is then viewed as a separate process occurring after rehydration. Based on temporal mRNA abundance, GALAU, JAKOBSEN and HUGHES (1991) suggested that late embryogenesis and early germination in dicotyledonous seeds should rather be viewed as consisting of four to five temporally discrete abundance components, each attributable to a developmental program. These five programs would be the cotyledon stage (COT), the maturation (MAT) stage which includes all seed development from the cotyledon stage up to desiccation, the ABA stage of transiently high ABA levels in the embryo during the water potential decline in the maturation stage, the post-abscission stage (PA) and early germination (GRM) at one to two days post imbibition. According to their model, five global temporal regulatory factors exist in which genes and consequently the developmental programs respond to different *cis* elements which are recognized by these factors or their derivatives. The factors controlling each program accumulate in the developing seed until the program is complete, then decline allowing the next factor to accumulate. The PA program appears to be autoregulated, switching itself off once sufficient PA proteins have accumulated within the embryo. These seem to be largely LEA (late embryogenic abundant) proteins involved in desiccation tolerance in the seed. With the removal of the factor controlling the PA program, the GRM program is free to proceed and appears only to require imbibition to be expressed (GALAU, JAKOBSEN and HUGHES, 1991). A similar result was found in *Brassica napus*, where analysis of

transcription in isolated nuclei suggested that the shift from an embryonic to a postgerminative pattern of gene expression occurs during rehydration (COMAI and HARADA, 1990).

A number of loci which appear to act in the regulation of the transition from embryogenesis to germination have been isolated through analysis of mutant phenotypes, mainly relating to a loss in dormancy. The majority of these loci have been identified from *Arabidopsis thaliana* and *Zea mays*. In *Arabidopsis*, these loci include *ABSCISIC ACID INSENSITIVE3 (ABI3)* (KOORNNEEF, REULING and KARSSSEN, 1984), *FUSCA3 (FUS3)* (BÄUMLEIN, MISERA, LUERSSEN, KÖLLE, HORSTMANN, WOBUS and MULLER, 1994; KEITH, KRAML, DEGLER and McCOURT, 1994) and *LEAFY COTYLEDON1 (LEC1)* (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994; WEST, YEE, DANAQ, ZIMMERMAN, FISHER, GOLDBERG and HARADA, 1994). In maize, the *Viviparous1 (VP1)* (McCARTY, HATTORI, CARSON, VASIL, LAZAR and VASIL, 1991) locus has been identified as being important in vivipary.

The *vp1* mutation in maize results in precocious germination of mutant embryos on heterozygous ears and was one of the first loci characterized as playing a role in seed germination. Embryos carrying this mutation contain wild type levels of ABA but exhibit reduced sensitivity to this hormone (ROBICHAUD and SUSSEX, 1986). In *Triticum aestivum* (wheat), pre-harvest sprouting (PHS) or vivipary may be associated with incorrect splicing of introns of *vp1*-related transcripts, such that sections of intron are either left in the transcript or that sections of exon are deleted (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001). The caryopses may therefore contain sub-optimal levels of the VP1 protein, which may explain the susceptibility to PHS and the low levels of embryo dormancy of modern wheat varieties (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001). The *ABI3* locus in *Arabidopsis* was also identified from seeds which were ABA-insensitive (KOORNNEEF, REULING and KARSSSEN, 1984). *ABI3* appears to play a central role in modulating the ABA response and regulating gene expression during embryo development in *Arabidopsis*

(MEINKE, FRANZMANN, NICKLE and YEUNG, 1994). The *ABI3* mutants are phenotypically similar to the *VP1* mutant, suggesting that these proteins have similar functions (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001).

The *leafy cotyledon* class of genes defines at least three genes with distinct but related functions during development (*LEC1*, *LEC2* and *FUS3*) (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994). *FUS3* and *LEC1* seeds have similarities in phenotype, although *LEC1* ultimately results in a much more dramatically altered embryo. The mutation was named for the fact that mutant cotyledons show a vascular pattern intermediate between that of wild-type leaves and cotyledons and also produced trichomes and stomata when embryo rescue was performed (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994). None of these mutants differ from the wild type in their perception of ABA (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994).

The *fus3* mutation alters late embryo functions such as the establishment of dormancy and desiccation tolerance as well as reducing levels of storage proteins in the embryo. Mutant seeds also accumulate anthocyanins at maturity (KEITH, KRAML, DEGLER and McCOURT, 1994). This is a heterochromatic mutation which affects the timing of the developmental processes of late embryo development and leaf development, such that *fus3* shoot apices resemble germinating wild-type seedlings more than wild-type embryos (KEITH, KRAML, DEGLER and McCOURT, 1994). Mutant seeds are nondormant and respond normally to ABA and GA (KEITH, KRAML, DEGLER and McCOURT, 1994).

Of the *leafy cotyledon* mutations, *lec1* exhibits the most dramatic alteration in late embryogenesis (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994). Mutant embryos have a reduced hypocotyl and rounded cotyledons which remain green until a late stage of development, as well as lacking storage reserves characteristic of wild type embryos. The seeds are deformed, occasionally viviparous and are not desiccation tolerant. When immature embryos are rescued in culture, viable plants

are produced which show a normal phenotype except that the cotyledons are leafy and have trichomes on their adaxial surfaces (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994). The severity of the *LEC1* mutant phenotype suggests that this gene is responsible for the regulation of a much wider range of embryo-specific programs and has a more global function than *VP1* or *ABI3* (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994). Mutant embryos therefore would show precocious germination as a result of a failure to activate programs which usually inhibit germination during seed development (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994).

The *lec2* mutant produces desiccation tolerant seeds which germinate at maturity (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994). Similarities to *lec1* and *fus3* are limited to trichomes and stomata on the cotyledons as well as differentiated vascular elements (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994).

These loci all show similarities to transcription factors and may function by activation or repression of gene expression programs associated with seed development and germination (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001). *ABI3* and *VP1* are homologues with a high degree of homology. The *FUS3* gene is related to the *ABI3/VP1* gene family, but is reduced in length (WOBUS and WEBER, 1999). Homology to the other two genes is restricted to a 100 amino acid section referred to as the B3 domain in *VP1*, which is able to mediate sequence specific binding *in vitro* and which is critical for gene activation at low ABA concentrations (WOBUS and WEBER, 1999). The *FUS3* protein lacks an amino-terminal A-domain, responsible in *VP1* and *ABI3* for mediating the ABA response. *FUS3* is thus not responsive to ABA levels (WOBUS and WEBER, 1999). *LEC1* is not related to *ABI3/VP1* or *FUS3* and may encode a protein related to a transcription factor subunit of the HAP3 type (WOBUS and WEBER, 1999).

ABI3, *FUS3* and *LEC1* affect a wide, broadly overlapping but not identical set of seed-specific characters (WOBUS and WEBER, 1999). Mutant embryos enter the germination program immediately by effectively skipping the maturation phase

(WOBUS and WEBER, 1999). These three genes may act synergistically, acting as transcription factors and forming hetero-oligomeric combinations which regulate different developmental processes (WOBUS and WEBER, 1999). FUS3 and ABI3 have been shown to induce maturation-specific promoters in a co-operative and synergistic fashion (WOBUS and WEBER, 1999). This promotion was strongly dependant on an intact RY promoter (CATGCAT) *cis*-element and suggests that the proteins may act by direct binding to the DNA (WOBUS and WEBER, 1999). LEC1 may also bind directly to the RY repeat (WOBUS and WEBER, 1999). In grasses, the same element is known as the Sph *cis*-element and is functionally dependant on the B3 domain of VP1 (WOBUS and WEBER, 1999). These proteins may act as central regulators, up- and down-regulating other regulators such as MYB transcription factors as well as directly regulating effector genes through the RY promoter *cis*-element (WOBUS and WEBER, 1999).

Late embryogenesis in *Arabidopsis* appears to be regulated through two different pathways which ensure the embryo attains desiccation tolerance and dormancy (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994). According to the model (Figure 1.3), *ABI3* regulates embryogenic programs and suppresses germination in response to ABA, whilst the LEC genes (including *FUS3*) regulate another pathway which is insensitive to ABA (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994). *LEC2* acts at a regulatory position early in the pathway, as mutations do not affect as many features of normal embryogenesis. *FUS3* would act further down the pathway where the embryogenic programs (desiccation tolerance and dormancy imposition) are regulated, with *LEC1* probably acting upstream from it since it affects a wider range of functions, as evidenced by the more severe defects in its phenotype (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994). This model is supported by the fact that the *lec1*, *fus3* and *abi3* mutations all inactivate different maturation phase-specific genes (WEST, YEE, DANA, ZIMMERMAN, FISHER, GOLDBERG and HARADA, 1994). Further evidence to support this fact comes from the additive effects of double mutants between *ABI3* and *LEC* mutations (MEINKE, FRANZMANN, NICKLE and YEUNG, 1994; LÉON-KLOOSTERZIEL, VAN DE BUNT, ZEEVAART and KOORNNEEF, 1996), as well as the fact that genes expressed

during maturation are regulated by different *cis*-elements and *trans*-acting factors (THOMAS, 1993).

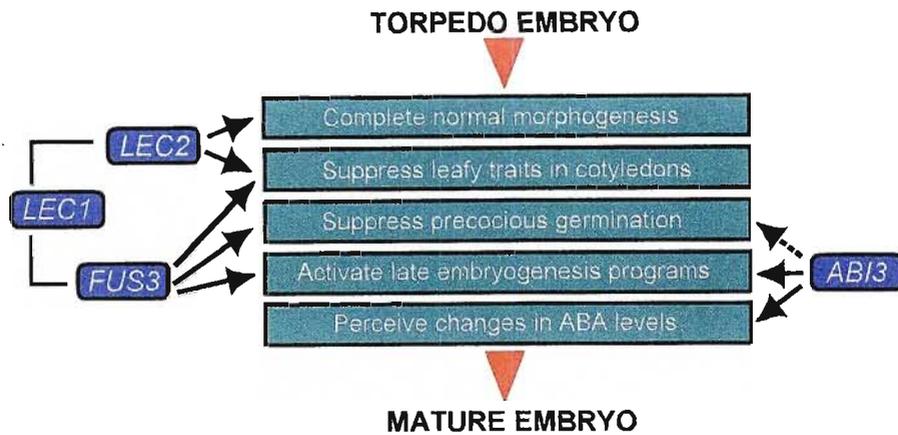


Figure 1.3: Model of genetic regulation of late embryogenesis in *Arabidopsis* based on mutant phenotypes *leafy cotyledon* and *ABA-insensitive*. Major events occurring during normal development are shown in the rectangles. These are not presented in sequential order. Arrows indicate which wild type genes are proposed to be required to complete each event (redrawn from MEINKE, FRANZMANN, NICKLE and YEUNG, 1994).

Insight into possible mechanisms of action for the ABA-responsive proteins ABI3 and VP1 has been provided by identification of some of the proteins with which they interact. These include TRAB (transcription factor responsible for ABA regulation) from rice, and ABI5 from *Arabidopsis*, which shows homology to TRAB (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001). The TRAB protein also interacts with ABREs (ABA-responsive *cis*-elements), providing a link between VP1 (and ABI3) and the ABA-response pathway (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001). Several other proteins with features similar to transcription factors and response regulators have also been identified which interact with these two loci, suggesting that VP1 and ABI3 also have an important interaction with transcription and other signalling pathways (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001). Several other species have extremely conserved homologues of ABI3/VP1. This suggests that this protein is a central component in

the regulation of embryo maturation and germination in flowering plants (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001).

In *Mesembryanthemum crystallinum*, the expression of a *Vp1*-like gene appears not to be related to embryo maturation as in other *Vp1/ABI3*-like genes, but rather to be directly linked to seed dormancy (FUKUHARA and BOHNERT, 2000). The deduced amino acid sequence encoded by *Mvp1* (*Mesembryanthemum Vp1*) is collinear with other VP1 homologues and is 39% identical and 50% similar to the *Arabidopsis* ABI3 protein (FUKUHARA and BOHNERT, 2000). *Mvp1* transcripts were not detectable in immature or mature desiccated seeds, but appeared during imbibition. Following imbibition, dormant (late-germinating) seeds are characterized by increased expression of *Mvp1*, which declines gradually over time. In early-germinating seeds, the *Mvp1* transcripts disappear within a few days (FUKUHARA and BOHNERT, 2000). When treated with cycloheximide, late germinating seeds germinated at the same rate as the early-germinating seeds and contained much-reduced levels of *Mvp1* (FUKUHARA and BOHNERT, 2000). These results indicate that as well as playing a role in seed development, *Vp1/ABI3*-like genes may also play a direct role in the control of seed germination.

Two further mutant loci which result in reduced dormancy in *Arabidopsis*, *RDO1* and *RDO2*, have been identified and their phenotypic effects characterized, although the genes corresponding to these loci have yet to be identified (LÉON-KLOOSTERZIEL, VAN DE BUNT, ZEEVAART and KOORNNEEF, 1996). Seeds of the mutant *rdo1* showed a slow reduction in dormancy which greatly preceded dormancy release in the wild type seeds, whilst *rdo2* seeds were non-dormant. Both mutants contained wild type levels of ABA in their seeds and had similar ABA sensitivities as the wild type, indicating that the reduced dormancy is not due to reduced ABA or ABA sensitivity in the seeds. The *rdo2* mutant had a slightly reduced sensitivity to GA biosynthesis inhibitors. It was hypothesized that the *RDO2* locus controls a step between the initial occurrence of ABA and ABA-induced dormancy (LÉON-KLOOSTERZIEL, VAN DE BUNT, ZEEVAART and KOORNNEEF, 1996). The *fus3*

mutation does not abolish the GA requirement for germination and *RDO1* may control a step in the same dormancy mechanism as the one affected by *fus3* (LÉON-KLOOSTERZIEL, VAN DE BUNT, ZEEVAART and KOORNNEEF, 1996). Additive effects of these genes with *ABI3/LEC1/FUS3* in double mutants adds further evidence to the model of a double pathway controlling embryogenesis (LÉON-KLOOSTERZIEL, VAN DE BUNT, ZEEVAART and KOORNNEEF, 1996). From these double mutant experiments, it seems that the two *RDO* may therefore be involved in different pathways regulating dormancy and embryogenesis. *RDO1* appears to be involved in the pathway regulated by *the LEC*-genes which leads to developmental arrest. *RDO2* is then involved in the *ABI3* pathway which induces a requirement for *de novo* GA biosynthesis in addition to the developmental arrest (LÉON-KLOOSTERZIEL, VAN DE BUNT, ZEEVAART and KOORNNEEF, 1996).

The *Arabidopsis DAG1* (Dof affecting germination) gene encodes a zinc finger transcription factor of the Dof transcription factor family and appears to be involved in the control of seed dormancy (PAPI, SABATINI, BOUCHEZ, CAMILLERI, CONSTANTINO and VITTORIOSO, 2000). A *dag1* knockout mutant produced seeds which were nondormant and capable of germination in the absence of light, as compared with wild-type plants which produce dormant seeds which have an absolute requirement for light for their germination following dormancy release (PAPI, SABATINI, BOUCHEZ, CAMILLERI, CONSTANTINO and VITTORIOSO, 2000). The mutation did not appear to bypass the requirement for phytochrome in the germination of the seeds, as a pulse of far-red light prevented the mutant seeds from germinating (PAPI, SABATINI, BOUCHEZ, CAMILLERI, CONSTANTINO and VITTORIOSO, 2000). The mutation did not affect the sensitivity of the seeds to ABA and treatment with this growth regulator induced dormancy in the mutant seeds (PAPI, SABATINI, BOUCHEZ, CAMILLERI, CONSTANTINO and VITTORIOSO, 2000). As with the wild type seeds, the mutant seeds required GA for germination and treatment with paclobutrazol prevented their germination (PAPI, SABATINI, BOUCHEZ, CAMILLERI, CONSTANTINO and VITTORIOSO, 2000). Segregation and expression analysis revealed that the *DAG1* gene is expressed maternally, in the vascular tissues of the mother plant (PAPI, SABATINI, BOUCHEZ, CAMILLERI,

CONSTANTINO and VITTORIOSO, 2000). These authors suggested a model to explain the mode of action of the gene. Expression of *DAG1* results in inactivity of a (protein) component required for dormancy initiation which is transported from the mother plant into the seed which is activated to act as a trigger for germination by red light. Disruption of the gene leaves this component in the active form, independent of red light, such that dormancy is not established in the seeds which also no longer have a light requirement for their germination. However, this component would still be sensitive to reversion to the inactive form on exposure to far-red light (PAPI, SABATINI, BOUCHEZ, CAMILLERI, CONSTANTINO and VITTORIOSO, 2000).

The *comatose* (*cts*) mutation in *Arabidopsis* appears to prevent or severely retard dormancy breaking by after-ripening in mutant seeds (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001). Germination in dormant *Arabidopsis* seeds can be enhanced by the application of GA, however, this does not induce germination of *cts* seeds (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001). This suggests that the *CTS* locus encodes protein required for the enhancement of germination potential and the activation of germination (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001). RUSSELL, LARNER, KURUP, BOUGOURD and HOLDSWORTH (2000) have shown that *CTS* represses *ABI3*, *FUS3* and *LEC1* function. This suggests that embryo maturation and germination are controlled by interactions which repress precocious germination (*ABI3*, *FUS3* and *LEC1*) and activate germination (*CTS*) at the appropriate time (Figure 1.4) (RUSSELL, LARNER, KURUP, BOUGOURD and HOLDSWORTH, 2000). Several other loci, such as *ERA1*, *GA1,2,3*, *GAI*, *SLY* and *DAG1*, which enhance germination potential have been identified (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001). The *CTS* locus, however, appears to be the only locus identified to date which has a seed-specific phenotype (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, McKIBBIN, BAILEY, LARNER and RUSSELL, 2001).

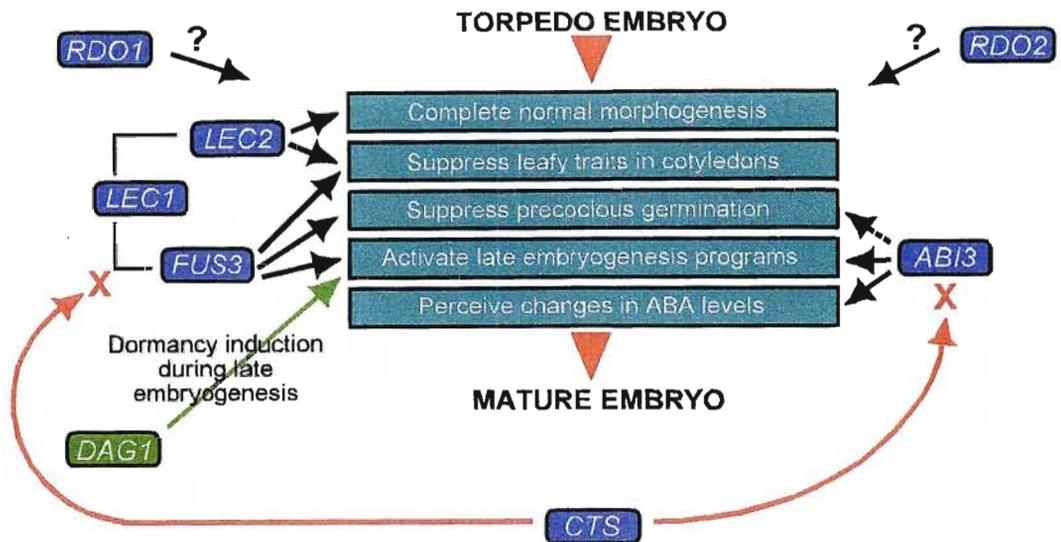


Figure 1.4: Modified model of genetic regulation of late embryogenesis in *Arabidopsis*, encompassing data from all germination-related mutants. Major events occurring during normal development are shown in the rectangles. Black arrows indicate which wild type genes are proposed to be required to complete each event. *DAG1* (green arrow) is expressed in the mother plant and causes the induction of dormancy during late embryogenesis. *CTS* reduces the expression of the *LEC*-genes and *ABI3* and activates germination. The *RDO* loci appear to interact with the two different pathways, affecting dormancy imposition. Their exact placement within these two pathways is unknown, as the genes representing these two loci have not yet been isolated. Like *DAG1*, they probably function within the late embryogenesis programs, controlling dormancy induction (modified from MEINKE, FRANZMANN, NICKLE and YEUNG, 1994).

The genes described above, with the exception of *VP1*, have all been isolated from *Arabidopsis* mutants and appear to interact with each other. From double mutant analysis, it seems as though there are two regulatory pathways by which these act. A revised model attempting to show the effects of all these genes is presented in Figure 1.4. It is not clear whether pathway is indeed an accurate description of these interactions, as the genes involved may not interact in a linear fashion. Rather, they may be associated with each other in a complex network which includes other factors which have not been identified to date. Many of these genes encode transcription-factor like proteins. It appears that these genes may play a central role in the

regulation of various programs during embryogenesis. Since mutations of these genes often result in precocious germination of the mutant embryos, or in the case of CTS completely prevent germination, some of the genes controlled by these regulatory proteins obviously encode critical steps in germination. The elucidation of the nature of those genes is the next challenge facing plant biologists in the attempt to more completely understand the process of seed germination.

1.2.3 Effect of temperature on seed germination

Temperature strongly influences both physiological and biochemical processes, including those occurring in germinating seeds. In the field, temperature acts to regulate germination in three main ways (BEWLEY and BLACK, 1986). Firstly, temperature may be involved in the removal of either primary or secondary dormancy. Secondly, temperatures outside of the normal limits for germination may cause the establishment of secondary dormancy in a seed when conditions are not favourable for seedling establishment. Finally, the temperature at which seeds are incubated determines their capacity for germination and the rate at which this occurs.

Each species has a range of temperatures over which germination may proceed and at which seedling establishment is possible (BRADBEER, 1988). There are three cardinal temperatures for germination, namely the maximum, minimum and optimum germination temperatures (BEWLEY and BLACK, 1982). The maximum and minimum temperatures represent the extremes of the range of temperatures over which the seed is able to germinate. The optimum germination temperature may be defined as that temperature at which the highest percentage germination may be obtained in the shortest possible period of time (MAYER and POLJAKOFF-MAYBER, 1975). These temperatures vary considerably between species, but may also vary between cultivars of the same species (BEWLEY and BLACK, 1982) and are thus determined by the source of the seeds, genetic differences within a species and the age of the seed (MAYER and POLJAKOFF-MAYBER, 1975). These temperatures may be considerably modified by several other factors, such as exposure to light of

different wavelengths or the application of different compounds, including phytohormones such as ethylene, abscisic acid (ABA) and gibberellic acid (GA).

Within the temperature range of a species, the rate of germination usually increases as the temperature rises, although the final percentage germination may decline (HEYDECKER, 1977). As the temperature approaches the maximum germination temperature, however, the germination rate begins to slow (HEYDECKER, 1977). Variations in the rates of germination of different seeds ensures that the seeds in a population germinate at different times, thereby leading to the temporal dispersal of germination (BEWLEY and BLACK, 1986). In wild populations this is a distinct advantage. It is, however, disadvantageous in a field crop where uniformity is an important consideration.

In the environment, temperatures greater than the maximum favourable temperature for germination will result in the suspension of germination. At these supra-optimal temperatures, seeds may enter into a state of either thermodormancy or thermoinhibition. It is important that a distinction be made between these two phenomena. In thermodormant seeds, a state of secondary dormancy is induced by exposure to the elevated temperature, which must then be released by some form of dormancy-breaking treatment before the seeds are able to germinate again at their optimal temperature (VIDAVER and HSIAO, 1975). Seeds are said to be thermoinhibited where they fail to germinate at a high temperature, but where germination proceeds immediately upon transfer to a temperature suitable for germination of seeds of that species (HOROWITZ and TAYLORSON, 1983). In both thermodormant and thermoinhibited seeds, if the temperature is too high or is maintained for an extended period, thermal death will result (HOROWITZ and TAYLORSON, 1983).

In *T. minuta*, germination is completely inhibited at temperatures in excess of 35°C (FORSYTH and VAN STADEN, 1983). This cessation of germination is due to thermoinhibition rather than thermodormancy, since 100% germination was recorded when the achenes were transferred to their optimum germination temperature of

25°C (FORSYTH and VAN STADEN, 1983).

Thermoinhibition may be used as a form of seed pre-treatment. In certain species, after a brief exposure to temperatures marginally in excess of their maximum germination temperature, processes occur within the seed which enable rapid completion of germination upon return to temperatures conducive to germination (HEYDECKER, 1977). This has been shown for *T. minuta* achenes (FORSYTH and VAN STADEN, 1983). When achenes were imbibed at 25°C, germination was spread over a period of several days. Where the achenes were first imbibed at 35°C for 10 days, 100% germination was recorded within 24 h of the temperature shift. The maximum germination response (100% germination within 24 h of transfer to 25°C) was only obtained when achenes were imbibed for at least 120 h at 35°C. DRENNAN and VAN STADEN (1989), however, reported that achenes became thermoinhibited between 18 and 24 h of imbibition at 35°C and that incubation at 35°C beyond 24 h showed no further advantage with regard to accelerated germination at 25°C. These differences may be due to differences in the seed lots used by these authors.

Imbibition is usually a triphasic process, comprising an initial period of water uptake, followed by a lag phase during which all of the major metabolic events leading up to germination occur, and ending with radicle protrusion (BEWLEY and BLACK, 1982). Presumably, whilst exposed to thermoinhibitory temperatures, *T. minuta* achenes as well as the seeds of other species showing a similar response, are able to undergo a number of these metabolic processes up to a certain stage, although final germination is prevented by high-temperature inhibition of one or more of the final processes necessary for radicle emergence to occur. The embryos are therefore all brought up to a similar state of maturation and so germinate rapidly and concomitantly once the temperature inhibition is relieved (DRENNAN and VAN STADEN, 1989). It is the nature of this thermal block which is the subject of this study.

1.2.4 Thermoinhibition

Observation of temperature effects on the germination of seeds of a number of different species has led to the formulation of a number of hypotheses concerning the basis of thermoinhibition in those species. An analysis of the available literature reveals that there appear to be a variety of possible mechanisms resulting in the arrest of germination. The situation is made more confusing by the fact that there appears to be a complex interplay between several different factors in a number of species, particularly in lettuce. In this species, not only do different factors appear to be involved in thermoinhibition in different cultivars, but in certain cultivars, a number of factors seem to act together or synergistically to prevent germination at supra-optimal temperatures. In the cultivar Grand Rapids, for instance, thermoinhibition has been proposed to be due to restrictions on radicle expansion by the endosperm or pericarp (DREW and BROCKLEHURST, 1984), by changes in endogenous cytokinin and gibberellin levels (SAINI, BASSI, CONSOLACION and SPENCER, 1986), by changes in phytochrome in the seeds (KRISTIE and FIELDING, 1994), by abscisic acid (ABA) (YOSHIOKA, ENDO and SATOH, 1998) and reduced anaerobic respiration and ATP production (SMALL, SCHULTZ and CRONJE, 1993).

Another problem in the study of thermoinhibition lies in terminology. Many authors, particularly in earlier papers, refer to thermodormant seeds, where the seeds in fact appear to be thermoinhibited. VIDAVER and HSIAO (1975) stated: "Secondary dormancy is a distinctly different condition from so-called thermodormancy in these seeds. The term thermoinhibition should probably be substituted for thermodormancy since seeds which have not yet become dormant will germinate merely by lowering the temperature." This distinction is a particular problem in lettuce seed studies. The threshold temperature at which seeds fail to germinate varies between different cultivars. When this threshold temperature is exceeded during imbibition, seeds become thermoinhibited, but if this supra-optimal temperature is maintained, the seeds eventually enter into a state of thermodormancy, which is indicated by the fact that these seeds will then not germinate even when the

temperature is reduced. A similar situation appears to occur in celery seeds, where seeds which do not germinate at high temperatures will germinate when transferred to a lower temperature. The longer the seeds are incubated at the supra-optimal temperature, the longer they appear to take to germinate when the temperature is reduced (BIDDINGTON and THOMAS, 1978). It is therefore important to examine the actual germination curves of the various seed lots used in the literature to determine whether the seeds used were in fact thermodormant or thermoinhibited.

1.2.4.1 Involvement of the embryo coverings in thermoinhibition

In many species, seed dormancy is imposed by the structures surrounding the embryo, including the endosperm, perisperm, pericarp and testa (BEWLEY and BLACK, 1982). These structures may impose dormancy through a number of mechanisms, including mechanical restraint on embryo extension, the establishment of permeability barriers interfering with water uptake or gaseous exchange or with the outward diffusion of endogenous germination inhibitors. Similar mechanisms may also be associated with thermoinhibition in some species. In the legume *Lathyrus sativus* L., supra-optimal temperatures increase hardseededness, preventing seed hydration and germination (AGRAWAL, SETHI and MEHRA, 1980).

The integuments of apple seeds are rich in phenolic compounds which fix oxygen through oxidation, thereby lowering the levels of oxygen available to the embryo (CÔME and TISSAOUI, 1973). When the seeds are imbibed at high temperatures the embryo's requirement for oxygen is increased as a result of an increase in the metabolic rate. The amount of oxygen available to the embryo is, however, decreased as oxidation of phenolics rises and oxygen becomes less soluble in the water in the seed coat and germination is consequently inhibited (CÔME and TISSAOUI, 1973). Hypoxic conditions imposed by the seed coat at supra-optimal temperatures may have a different effect in seeds which require ethylene for germination. Since the ethylene-forming enzyme (ACC oxidase) is oxygen-dependant, the conversion of ACC (1-aminocyclopropane-1-carboxylic acid) to

ethylene may be inhibited, preventing germination (YANG, 1985).

Embryo coverings also play a role in the thermoinhibition of lettuce seeds at supra-optimal temperatures (DREW and BROCKLEHURST, 1984). The involvement of different tissues surrounding the embryo (Figure 1.5) in thermoinhibition appears to vary, however, between seeds of different cultivars. In seeds of *Lactuca sativa* cv. Cobham Green, control over germination appears to be exerted by the pericarp, since relief of thermoinhibition at 35°C resulted when the pericarp was weakened by hypochlorite treatment (DREW and BROCKLEHURST, 1984). Hypochlorite treatment of the pericarp also appeared to overcome thermoinhibition in cultivars Empire and Benita, whilst treatment of seeds of cultivars Sabine and Bellona resulted in faster recovery from thermoinhibition than non-treated seeds (DREW and BROCKLEHURST, 1987). In seeds of the cultivar Grand Rapids however, thermoinhibition was relieved by weakening of the endosperm (DREW and BROCKLEHURST, 1984).

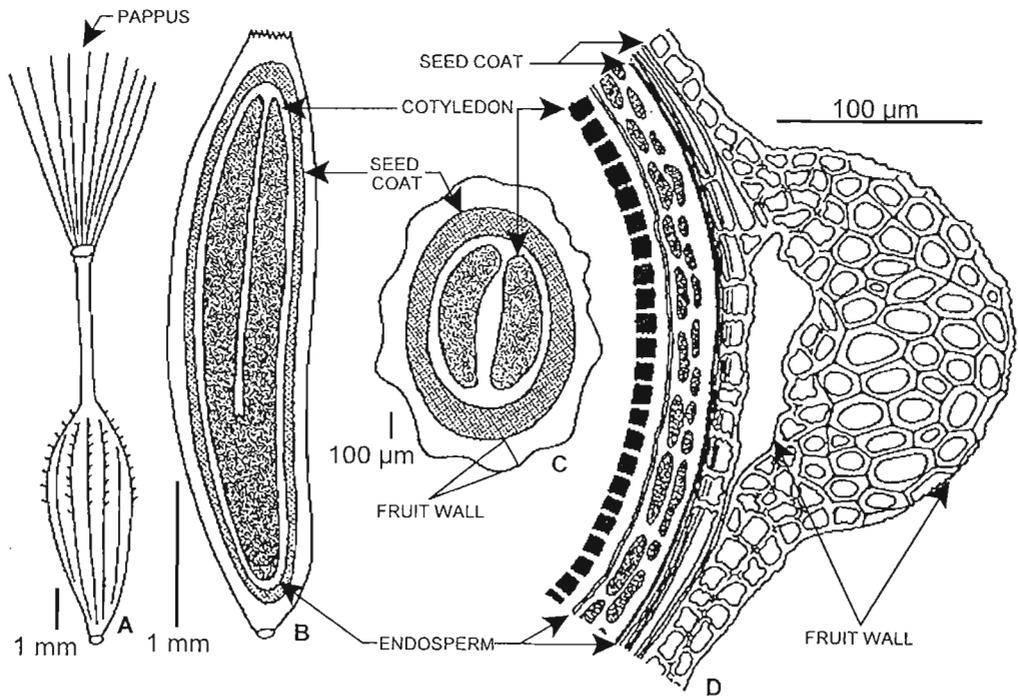


Figure 1.5: Cypsela-type achene of *Lactuca sativa* (lettuce) showing entire fruit with pappus still attached (A), longitudinal section (B), transverse section (C) and transection of part of the pericarp with subjacent layers (D) (ESAU, 1977).

BRADFORD and SOMASCO (1994) explained the role of embryo coverings in thermoinhibition in lettuce seeds by analysing water relations in lettuce cv. Empire seeds. This work did not attempt to distinguish between the effect of the endosperm and the pericarp, but rather refers to the effects of the embryo coverings as a whole (Figure 1.5). The ability of the embryo to absorb water from its surroundings and initiate growth is determined by the water potential (Ψ) of its cells. Thermoinhibition has been shown to be associated with an increase in the yield threshold that must be exceeded in order for radicle emergence to take place (KARSSSEN, HAIGH, VAN DER TOORN and WEGES, 1989). This total yield threshold is made up of a component from the radicle cell walls (Y_r) and a component from the tissues surrounding the embryo (Y_e). As the incubation temperature increases, the osmotic potential of the embryo (Ψ_{π}), the embryo turgor threshold for growth (Y_r) and endosperm resistance (Y_e) all increase relative to seeds incubated at optimal temperatures, causing thermoinhibition (BRADFORD and SOMASCO, 1994). The removal of the structures enclosing the embryo thus results in the alleviation of thermoinhibition by eliminating the contribution of Y_e to the total yield threshold (BRADFORD and SOMASCO, 1994). This was also shown in the lettuce cv. Pacific, where slitting the seeds to disrupt the integrity of the endosperm extended the temperature range for germination and maintained lower base water potential in the slit seeds as compared with intact seeds (DUTTA and BRADFORD, 1994). The fact that removal or scarification of the endosperm does not affect differences in temperature sensitivity between seeds of different lettuce seed cultivars (DUNLAP, SCULLY and REYES, 1990, PRUSINSKI and KHAN, 1990) however, indicates that the embryo itself also plays a key role in thermoinhibition (BRADFORD and SOMASCO, 1994).

Weakening of the endosperm through cell wall autohydrolysis may be important for germination in lettuce seeds (DUTTA, BRADFORD and NEVINS, 1994). Autolytic activity appeared to be an enzymatic process as it was dependant on temperature and pH (DUTTA, BRADFORD and NEVINS, 1994). When isolated cell walls from thermoinhibited or ABA-treated seeds were examined, autohydrolysis was reduced by up to 25% compared with control endosperms isolated from germinating seeds.

Treating seeds with kinetin and GA increased the rate of hydrolysis by 20 to 30% compared with thermoinhibited controls (DUTTA, BRADFORD and NEVINS, 1994). Although a direct relationship between autolysis and actual weakening of endosperm walls in intact seeds was not established, temperature inhibition of enzymatic cell wall autohydrolysis may play a slight role in thermoinhibition.

1.2.4.2 The role of gibberellins and cytokinins in thermoinhibition

From their work on celery (*Apium graveolens*) cv Lathom Blanching, BIDDINGTON and THOMAS (1978) reported that temperature may control germination in these seeds by affecting endogenous levels of cytokinins and gibberellins. Exogenous application of these hormones increased the germination of celery seeds subjected to supra-optimal temperatures in both the light and the dark. These two hormones also appeared to act in a synergistic manner in the dark (BIDDINGTON and THOMAS, 1978). A similar result has also been shown for the lettuce cultivar Grand Rapids (SAINI, BASSI, CONSOLACION and SPENCER, 1986). BIDDINGTON and THOMAS (1978) proposed that temperature may control germination in celery seeds by affecting the levels of endogenous gibberellins or a gibberellic acid/protein complex. They suggested that germination is primarily induced by gibberellins, endogenous levels of which decline as temperatures increase and rise again when the temperature is reduced (Figure 1.6).

Relief from thermoinhibition was enhanced by white or red light (BIDDINGTON and THOMAS, 1978). Quantitative and qualitative changes in the cytokinin contents of seeds have been reported in response to dormancy-breaking red light treatment of celery, lettuce (VAN STADEN, 1973) and *Rumex obtusifolius* (VAN STADEN and WAREING, 1972). These results suggest that phytochrome-mediated control over the germination process is expedited via endogenous cytokinins. BIDDINGTON and THOMAS (1978) therefore suggested that red-light may enhance the effect of gibberellins on the relief of thermoinhibition through increased cytokinin levels.

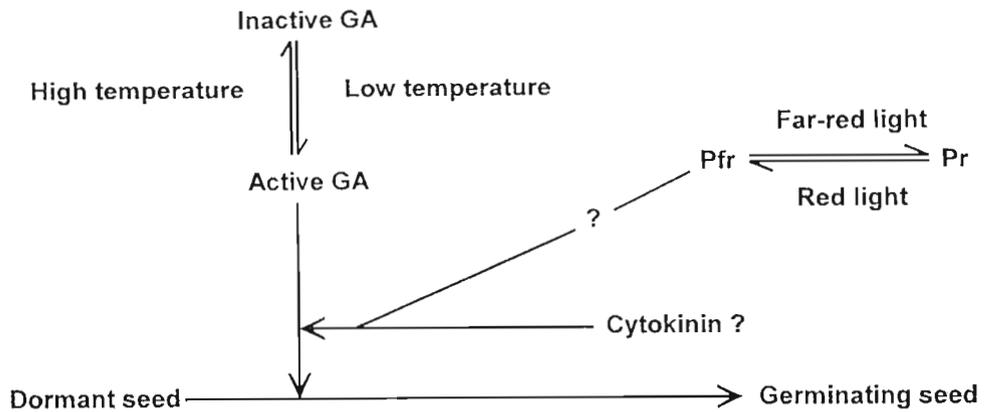


Figure 1.6: Proposed model for temperature control of germination mediated by gibberellins and cytokinins (redrawn from BIDDINGTON and THOMAS, 1978).

Cytokinins may play a subsidiary role in alleviating thermoinhibition in lettuce seeds by increasing ethylene biosynthesis or the sensitivity of the seeds to ethylene. Exogenously applied kinetin and ACC act synergistically to increase pre-germination levels of ethylene and enhance germination at temperatures of 32°C and 35°C (KHAN and PRUSINSKI, 1989). Removing the restriction imposed by the tissues enclosing the embryo by cutting the testa allowed ACC to be readily taken up and converted to ethylene, removing the synergistic effect of the added kinetin. This result suggests that cytokinins have an important regulatory role for ethylene biosynthesis and hence for germination in intact lettuce seeds at high temperatures, and that the seed coat may be an important factor in this regulation (KHAN and PRUSINSKI, 1989). This may involve enhanced utilization of ACC as a result of activation of the ethylene-forming enzyme (EFE) or ACC oxidase (HUANG and KHAN, 1992) and an interaction of the ACC-derived ethylene with the cytokinin (KHAN and PRUSINSKI, 1989). In Grand Rapids seeds, treatment with kinetin and oxygen also stimulated germination at thermoinhibitory temperatures, possibly either by causing seeds to bypass their ethylene requirement or increasing their sensitivity to ethylene (SMALL, SCHULTZ and CRONJE, 1993).

1.2.4.3 The role of ethylene in thermoinhibition

In a number of species, seed germination is strongly related to the biosynthesis and action of ethylene (GALLARDO, SANCHEZ-CALLE, MUÑOZ DE RUEDA and MATILLA, 1996). HUANG and KHAN (1992) found that preconditioning seeds of the lettuce cultivar Mesa 659 with the moist solid carrier Micro-Cel E conferred greater thermotolerance on these seeds, alleviating thermoinhibition at 35°C. The ability of the pre-conditioned seeds to rapidly synthesize ACC and to respond to ACC and ethylene on exposure to temperatures which were normally thermoinhibiting appeared to be the critical factor for the alleviation of thermoinhibition. Exposing seeds to aminoethoxyvinylglycine (AVG), an inhibitor of ACC synthesis, prevented pre-conditioned seeds from germinating at 35°C. This inhibitory effect of AVG was reversed by supplying the seeds with exogenous ACC, ethephon or ethylene (HUANG and KHAN, 1992). The use of ACC is enhanced by cytokinins under conditions of temperature stress (KHAN and HUANG, 1988), and the production of cytokinin or cytokinin-like factors may also play a regulatory role in ethylene biosynthesis during thermoinhibition (Section 1.2.4.2) (HUANG and KHAN, 1992).

Alleviation of thermoinhibition in chickpea (*Cicer arietinum*) may also be induced by the addition of exogenous ethylene or by the stimulation of ethylene production within the seeds (GALLARDO, SANCHEZ-CALLE, MUÑOZ DE RUEDA and MATILLA, 1996). Thus the addition of ethrel (chloroethylphosphonic acid), ACC or spermine to the germination medium or the inhibition of polyamine biosynthesis, which competes with the ethylene biosynthetic pathway for the same precursors, by CHA (cyclohexamine) or MGBG (methyl-glyoxal-*bis*-guanylhydrazone), could induce germination of chickpea seeds at thermoinhibitory temperatures (MUÑOZ DE RUEDA, GALLARDO, SANCHEZ-CALLE, and MATILLA, 1994).

Unlike lettuce seeds, where high temperatures inhibit ACC production, supra-optimal temperatures appear to stimulate ACC synthase activity in chickpea seeds. Despite this stimulation, ACC levels decrease in thermoinhibited seeds possibly due to an increase in ACC conjugation. (GALLARDO, DEL MAR DELGARDO, SANCHEZ-

CALLE and MATILLA, 1991). With an increase in temperature, a decrease in the levels of ACC in the embryonic axis of chickpea seeds and a concomitant increase in the levels of conjugated ACC in the embryonic tissues was observed. Furthermore, malonyl ACC-transferase, the enzyme responsible for the conjugation of ACC, has a very low K_m value. It therefore appears that in chickpea seeds, ACC malonisation and the consequent inhibition of ethylene synthesis is one of the principle causes of thermoinhibition (GALLARDO, SANCHEZ-CALLE, MUÑOZ DE RUEDA and MATILLA, 1996). The conversion of ACC to ethylene may also be affected at high temperatures, possibly through some temperature effect on ACC oxidase (GALLARDO, SANCHEZ-CALLE, MUÑOZ DE RUEDA and MATILLA, 1996). In sunflower (*Helianthus annuus*) too, thermoinhibition is associated with the loss of the seed's ability to synthesize ethylene (CORBINEAU, RUDNICKI and CÔME, 1989).

Besides changes in ethylene biosynthesis, ethylene and ethylene action have also been implicated in the alleviation of stress on lettuce seed germination (PRUSINSKI and KHAN, 1990). SAINI, CONSOLACION, BASSI and SPENCER (1989) have reported that in lettuce seeds, thermoinhibition does not appear to be related only to a reduction in the ability of the seeds to synthesize ethylene, but may rather be due to high temperatures causing an increase in the threshold concentrations of this hormone required to elicit germination. PRUSINSKI and KHAN (1990) have also reported that seeds and seedlings have a greater sensitivity to ethylene when stressed than under non-stressful conditions, which may suggest a greater need for this phytohormone for growth processes under stress conditions. These authors found a positive correlation between the ethylene-producing capacity of seeds of a number of lettuce cultivars and the ability of the seeds to germinate under conditions of salinity, osmotic and temperature stress.

Furthermore, thermoinhibition may be related to other metabolic changes within the seeds at supra-optimal temperatures, including changes in phytochrome (SAINI, CONSOLACION, BASSI and SPENCER, 1989). Ethylene, GA_3 , kinetin, carbon dioxide (CO_2) and light were all shown to have no effect on overcoming

thermoinhibition at 32°C, but that the action of any one of the hormones required the presence of at least one other growth regulator, light or CO₂ or a combination of these. To achieve 100% germination, at least three stimuli were required (SAINI, BASSI, CONSOLACION, and SPENCER, 1986). This suggests that thermoinhibition in relation to ethylene may be regulated at a number of points.

The mechanism whereby ethylene is involved in the regulation of germination appears therefore to differ in different species of plants. It should, however, be noted that the alleviation of thermoinhibition by ethylene is limited to species in which seed germination is dependant on that phytohormone.

1.2.4.4 The role of abscisic acid (ABA) in thermoinhibition

Fluoridone (1-methyl-3-phenyl-5-[3-trifluorlmethyl-(phenyl)]-4-(1*H*)-pyridinone) is an inhibitor of the carotenoid biosynthetic pathway enzyme phytoene desaturase. Since carotenoids are the main precursors of ABA in plants, inhibition of carotenogenesis also inhibits ABA biosynthesis (YOSHIOKA, ENDO and SATOH, 1998). Application of this compound to *Lactuca sativa* cv. Grand Rapids seeds was able to overcome thermoinhibition up to 33°C, although at this temperature germination was markedly delayed as compared to 23°C (YOSHIOKA, ENDO and SATOH, 1998). Fluoridone application also overcame inhibition of germination at supra-optimal temperatures in seeds of *Chrysanthemum parthenim*, *Freesia hybrida*, *Cerastium glomeratum*, *Stellaria neglecta*, *Agrostis alba*, *Conyza canadensis*, *Cryptotaenia japonica*, *Dactylis glomerata*, *Festuca rubra*, *Medicago sativa*, *Plantago lanceolata*, *Trifolium repens* and *Vicia angustifolia*, although the range of temperatures at which fluoridone promoted germination differed between species (YOSHIOKA, ENDO and SATOH, 1998). Fluoridone had no effect on two gramineous weeds, *Bromus catharticus* and *Lolium perenne*, which were primarily dormant as opposed to thermoinhibited. These authors therefore suggested that ABA plays a decisive role in the regulation of seed germination at supra-optimal temperatures.

1.2.4.5 Repression of metabolic reactivation as a factor in thermoinhibition

In *Phacelia tanacetifolia* seeds, germination is inhibited at temperatures in excess of 26°C (PIROVANO, MORGUTTI, ESPEN and COCUCCI, 1997). This inhibition is correlated with the lack of activation of the seed's metabolic processes. In the first 24 h of imbibition, an increase in the levels of glucose-6-phosphate (Glc-6-P) was inhibited first, followed by inhibition of increases in enzymic activity and then by inhibition of increases in levels of ATP, reducing sugars, RNA and DNA. The inhibition of radicle emergence correlated most strongly with the levels of Glc-6-P. After 24 h, when germination was inhibited by 60%, only the levels of Glc-6-P was still inhibited to the same extent (PIROVANO, MORGUTTI, ESPEN and COCUCCI, 1997). Appreciable qualitative differences in transcriptional and translational activities were detected at 16°C and 30°C. Both activities were higher in seeds incubated at 30°C than in seeds incubated at the lower temperature after 9 h, but with time the translational activities became much lower at 30°C than at 16°C (PIROVANO, MORGUTTI, ESPEN and COCUCCI, 1997). Examination of protein patterns also revealed that a number of polypeptides which disappeared during imbibition at 16°C did not disappear in temperature-inhibited seeds (PIROVANO, MORGUTTI, ESPEN and COCUCCI, 1997).

The high temperature may inhibit germination at two levels (PIROVANO, MORGUTTI, ESPEN and COCUCCI, 1997). Firstly, the high temperature may have a general effect on the rate of the reactivation processes occurring within the imbibing seeds. This may be linked to the fact that cells of seeds incubated at 30°C also had lowered levels of mobile potassium (K^+) ions and the plasma membranes showed greater permeability to K^+ , compared with cells of seeds incubated at 16°C. The control of K^+ concentrations in cells is crucial for the regulation of metabolic activity and these authors suggested that reduced cellular concentrations of K^+ may inhibit the activation of K^+ -dependant mechanisms and hence germination in seeds at 30°C. Secondly, elevated temperature may also work directly on single reactions involved in the inhibitory mechanism of germination in *P. tanacetifolia* seeds, either inducing a lack of function through denaturation of key components or by increasing processes

related to the synthesis of inhibitory substances.

In lettuce cv. Grand Rapids, thermoinhibition at 38°C markedly reduced ATP and total adenylate content of the seeds, as well as adenylate energy charge (SMALL, SCHULTZ and CRONJE, 1993). Seeds had a high ethanol content at this temperature and appeared to undergo ethanolic fermentation, probably as a result of a reduction in oxygen solubility at this temperature. Thermoinhibition was alleviated and seeds contained normal levels of ATP when treated with 100% oxygen and kinetin at 38°C (SMALL, SCHULTZ and CRONJE, 1993). A reduction in aerobic respiration and consequently in ATP production may therefore be a contributing factor in thermoinhibition (SMALL, SCHULTZ and CRONJE, 1993).

1.2.4.6 The involvement of a thermo-labile factor in thermoinhibition

Before PIROVANO, MORGUTTI, ESPEN and COCUCCI (1997) proposed that denaturation may be a factor in thermoinhibition, TAKEBA and MATSUBARA (1976) reported that a thermo-labile process is involved in the germination of lettuce cv. New York seeds. Germination was completely inhibited at a temperature of 30°C in the dark, but was re-activated after incubation at 20°C. Although TAKEBA and MATSUBARA did not establish the nature of the thermo-labile factor, they proposed that it was probably protein-based. The longer the seeds were pre-incubated at 30°C, the longer the period required at the lower temperature for recovery of germination. A similar trend has been observed in *T. minuta*, where seeds incubated for extended periods of more than 20 days at 35°C begin to take longer to achieve germination (FORSYTH and VAN STADEN, 1983). TAKEBA and MATSUBARA suggested that this thermo-labile factor was thermoreversible and that the degree of its re-activation upon return to favourable temperatures is dependant on the extent of its inactivation. Three models were proposed to attempt to interpret these results in terms of the responses of the lettuce seeds to light and temperature (TAKEBA and MATSUBARA, 1976):

1. Active phytochrome may promote germination by increasing the amount of the

- precursor of the reaction mediated by the thermo-labile factor,
2. The product(s) of the phytochrome system is(are) changed to the essential metabolite for germination by combining with the product(s) of the reaction mediated by the thermo-labile factor, and
 3. The phytochrome system increases the stability of the thermo-labile factor by unknown mechanisms.

1.2.4.7 Role of phytochrome in thermoinhibition

In lettuce, the promotion of seed germination through the red/far-red reversible or Low Fluence Response (LFR) is closely linked to temperature (KRISTIE and FIELDING, 1994). For the majority of lettuce cultivars, a pulse of red light is able to promote germination within a certain range of temperatures. When this range is exceeded, the response to the red light pulse is diminished and seeds become thermoinhibited (VIDAVER and HSIAO, 1975). In lettuce cv. Grand Rapids, dark germination at 20°C results in almost 100% germination, but thermoinhibition prevents germination at 27°C and above (SAINI, CONSOLACION, BASSI and SPENCER, 1989). This has been attributed to the reversion of Pfr to Pr before it has time to produce an effect, a hypothesis which is supported by the finding that repeated red light pulses raise the upper threshold temperature for germination by several degrees as compared to a single pulse of red light (SAINI, CONSOLACION, BASSI and SPENCER, 1989). These authors also suggested that ethylene was essential for this light-induced alleviation of thermoinhibition. KRISTIE and FIELDING (1994) suggested that dark germination was governed by levels of Pfr and that the mean level of Pfr required to induce dark germination was increased as temperatures rose. It was proposed that this may be attributable to an increase in dark-reversion of Pfr to Pr at high temperatures (KRISTIE and FIELDING, 1994). Preliminary Pfr requirement curves generated from a number of different lettuce cultivars suggest that the shape of the Pfr requirement curve may partially account for differences in the temperature response curves for dark germination that occur in the different cultivars (KRISTIE and FIELDING, 1994). The response to Pfr:Pr

levels may differ in different species, however. It appears that in *Plantago major* (PONS, 1986, 1992, cited in KRISTIE and FIELDING, 1994), the opposite situation may be true in that lowered levels of Pfr are responsible for promoting germination at higher temperatures.

1.2.4.8 Thermoinhibition in *Tagetes minuta*

From the above, it would seem that in all cases described thus far in the literature, thermoinhibition is a passive state, imposed on seeds held under conditions of temperature stress through heat-induced changes in their normal physiological functioning. Whilst most studies have focussed on specific aspects of seed physiology, such as plant growth regulators, enzymic reactions or phytochrome, it would appear that control of germination is levied at multiple sites and that thermoinhibition is actually the result of a complex interaction of a number of factors in these species.

In *T. minuta*, thermoinhibition of the achenes is a very tightly regulated phenomenon. If the temperature is reduced even slightly below 35°C, germination proceeds rapidly (DREWES, 1989). None of the factors described in the literature appear to be involved in thermoinhibition in this species. The achenes easily take up water when imbibed at 35°C, indeed, with careful manipulation the embryo can be squeezed from its coverings once imbibed. Whilst the seeds have a light requirement for germination, seeds incubated at 35°C in white light become thermoinhibited. Ethylene does not appear to be involved in thermoinhibition in *T. minuta* achenes. The application of ethrel to achenes imbibed at 35°C, either as short pulses or through continuous exposure, had no statistically significant alleviatory effect on thermoinhibition of achenes of this species (DREWES, 1989). When ethrel was applied at 25°C, the germination rate was not affected either one way or the other, indicating that *T. minuta* achenes are not dependant on ethylene for their germination (DREWES, 1989). Indeed, ethrel treatment of the achenes resulted in abnormal post-germinative development and both hypocotyl and radicle extension were

severely retarded. The cytokinins kinetin and benzyladenine (BA) also have no effect on thermoinhibited achenes (DREWES, 1989). A commercial mixture of gibberellin GA_{4+7} and BA, Promalin, also had no alleviatory effect on thermoinhibition, indicating that even a combination of light and multiple growth regulators was unable to prevent thermoinhibition in this species. The application of gibberellic acid (GA) by itself, however, did partially alleviate thermoinhibition in this species, with GA_{4+7} being more effective than GA_3 (DREWES, 1989). Yet, even GA application was unable to induce more than approximately 75% germination at 35°C and no major changes in endogenous gibberellin-like substances were detected in thermoinhibited achenes (DREWES, 1989). This suggests that thermoinhibition in this species may be imposed in a manner not previously described, possibly by an actively-imposed block on germination through gene expression, in a manner similar to the inhibition of precocious germination or to genetically-imposed seed dormancy.

1.3 AIMS AND OBJECTIVES

When achenes of *T. minuta* are incubated at 25°C, germination is spread over between two and seven days. When seeds are imbibed at 35°C they are prevented from germinating and become thermoinhibited (FORSYTH and VAN STADEN, 1983). However, when the achenes are imbibed for a short period at 35°C and are then transferred to a temperature of 25°C, 100% germination is achieved within 24 hours. This suggests that during imbibition at 35°C the achenes undergo a number of metabolic processes leading towards radicle emergence but that this final step of germination is prevented through the high-temperature inhibition of one or more of the final metabolic steps of germination. This appears to be an actively-imposed block on germination, unlike most other mechanisms previously proposed to account for thermoinhibition in a variety of species. The aim of this project is thus to attempt to elucidate the factors underlying thermoinhibition in *Tagetes minuta* at the molecular level.

CHAPTER 2

GENERAL CHARACTERISTICS OF *Tagetes minuta* GERMINATION

2.1 INTRODUCTION

Despite its weed status, *Tagetes minuta* is an exceptionally useful plant as a result of the secondary metabolites it produces in such abundance (Section 1.1). In South Africa, attempts have been made to grow this species in the Eastern Cape in the regions formerly known as the Ciskei and Transkei (FORSYTH and VAN STADEN, 1983). Difficulties were experienced, however, in the domestication of the species, since the erratic germination which ensure the species survival as a weed makes *T. minuta* intractable in agronomic situations, where synchronicity is a virtue. This has led to several studies on the germination characteristics of the species, which are now fairly well characterized.

FORSYTH and VAN STADEN (1983) have shown that *T. minuta* achenes have both a temperature and a light requirement for germination. At 25°C in white light, 100% germination was achieved, compared with only 27% germination in the dark at the same temperature. These authors established that 25°C is the optimum germination temperature for *T. minuta*, with 100% germination being achieved over a seven day period. This period appears to differ between different seed lots, since DRENNAN and VAN STADEN (1989) and DREWES and VAN STADEN (1990) observed that 100% germination could be achieved within two days. No germination was observed at 10°C and 35°C (FORSYTH and VAN STADEN, 1983), these obviously representing the limiting temperatures for germination of the achenes.

Whilst achenes did not germinate at 35°C, they were not in any way harmed by incubation at this temperature. When the temperature was dropped to 25°C, 100% germination was observed within 24 h. DREWES (1989) showed that similar results

may be achieved for temperatures up to 38°C. After 24 h imbibition at 40°C, a slight lag was observed before the achenes began to germinate, presumably indicating that some damage was sustained by the embryos during incubation at this temperature which had first to be repaired before germination could be achieved. Since no further dormancy-breaking treatment is needed to induce germination besides a drop in the incubation temperature, achenes are only thermoinhibited at high temperatures and do not enter a state of thermodormancy. It was observed that if the thermoinhibited achenes were dried and later re-imbibed at 25°C, 100% germination was again achieved within 24 h (FORSYTH and VAN STADEN, 1983; DRENNAN and VAN STADEN, 1989). The possibility thus arises that thermoinhibition, followed by rapid drying of the achenes, could be used as a form of pre-treatment to increase the synchronicity of germination to make farming of the species a viable option. Although a 10 d pre-treatment was used by FORSYTH and VAN STADEN (1983), later studies showed that this 'accelerated germination' could be achieved after only 24 h of incubation at 35°C (DRENNAN and VAN STADEN, 1989; DREWES and VAN STADEN, 1990).

DREWES (1989) subsequently conducted further studies on the physiology of thermoinhibited *T. minuta* achenes. Whilst ethylene and cytokinins appeared not to play any role in thermoinhibition in this species, the exogenous application of gibberellic acids was found to result in a partial alleviation of thermoinhibition. A mixture of GA₄₊₇ was more effective than GA₃, and approximately 75% germination could be achieved at 35°C following the application of a 1 h pulse treatment of the achenes with GA₄₊₇. Endogenous levels of GA did not appear to differ in thermoinhibited and germinating achenes, however. These results, and the fact that exogenously-applied GA did not completely restore germination, suggest that thermoinhibition is not purely due to a lack of active gibberellins at high temperatures, but do indicate that GA is probably somehow involved in thermoinhibition.

The aim of the experiments outlined in the following Chapter was to confirm that the achenes used in this study conform to the trends described in the literature. It was also necessary to characterize exactly the germination events of the achenes

collected for use in this study, so that any events observed at the molecular level could be correlated with the various stages of the germination process, particularly as the literature appears to indicate that the timing of germination differs slightly between different populations of achenes.

2.2 MATERIALS AND METHODS

2.2.1 Collection of plant material

Achenes of *Tagetes minuta* L. were collected during May and June 1997. All achenes were harvested from a single large population growing in a field at the intersection of Durban and Blackburrow Roads in Pietermaritzburg (29° 36' S; 30° 23' E). Inflorescences were collected as soon as they began to dry. These were then allowed to dry fully on open trays in the laboratory for six weeks. Once dry, the achenes were easily shaken free from the inflorescences. Most of the other plant debris was removed by sieving the collected achenes through a 1.7 mm and a 1.0 mm brass mesh sieve in an Endecotts test sieve shaker. Any remaining small debris and most of the unfilled achenes were separated by the use of a Hearson's apparatus. By regulating the airflow through the apparatus, the lighter debris and non-viable achenes were blown away from the viable achenes and discarded. The remaining achenes were used for experimentation. These were placed into brown paper bags and divided into two groups. Each of the groups was stored in a desiccator over silica gel. Half of the achenes were stored in a cold room at 10°C, whilst the other group were stored in a freezer at -20°C.

2.2.2 Determination of moisture content

Four groups of 25 achenes each were massed to the nearest 0.1 mg, placed in small manilla envelopes and dried at 70-75°C for 72 h. The achenes were then reweighed. Moisture content was then calculated as a percentage of the fresh mass of the

achenes, using the following equation:

$$\frac{FW - DW}{FW} \times 100 = \% \text{ moisture content}$$

Results were given as the mean of the four replicates.

2.2.3 Germination trials

All germination trials were conducted using four replicates of 25 achenes each. Achenes were incubated in plastic petri dishes on two layers of Whatman's No. 1 filter paper moistened with dH₂O. Where the effects of gibberellic acid on germination were to be examined, various concentrations of GA solutions were used instead of dH₂O. Petri dishes were sealed in plastic bags to reduce moisture loss. Additional dH₂O (or GA solution) was placed in the incubators to allow the filter paper to be remoistened if necessary. Achenes were incubated in Labex Labcon LTGC controlled-environment chambers at either 25±1°C or 36±1°C, with a light intensity of 4.8 μmol photons/m²/sec. Petri dishes were examined for signs of radicle protrusion, which was taken as the criterion for germination, at two hourly intervals.

2.2.4 Imbibition curves

The rate of water uptake by achenes was measured at 25°C and at 36°C. Four replicates of 25 achenes each were weighed to the nearest 0.1 mg. Achenes were imbibed in plastic petri dishes on two layers of Whatman's No. 1 filter paper moistened with dH₂O. Following the addition of water, the samples were weighed at 4 hourly intervals. Before weighing, samples were removed from the petri dishes and gently blotted to remove excess surface water. Time at which radicle protrusion was first observed was also recorded.

2.2.5 Freezing Trial

Achenes were sealed in a desiccator over silica gel and held at -20°C for 7 d. At the end of this period, the achenes were removed and allowed to return to room temperature. They were then used in a germination trial conducted as above at 25°C and 36°C. A control group composed of achenes which had not been frozen was included at each temperature.

2.3 RESULTS AND DISCUSSION

2.3.1 Germination trials

Before any molecular analysis of the events surrounding germination and thermoinhibition of achenes of *T. minuta* was undertaken, the general germination characteristics of the achenes which were collected for use in this study was made. This study, it was hoped, would fulfill two purposes. Firstly, it was necessary to characterize the exact timing of the various germination events in the sample of achenes under study so that these could later be related to any changes observed at the molecular level. Secondly, a detailed physiological study of thermoinhibition in *T. minuta* achenes has previously been conducted (DREWES, 1989) and, it was hoped that this would show that the germination characteristics of the achenes used for this molecular study were sufficiently similar to those of this previous study as to allow this study to be viewed as a logical and scientifically valid extension of that analysis. Discussion will therefore be limited to a comparison between the results obtained in this study and those obtained in previous reports on the germination characteristics of *T. minuta* achenes, such as those by FORSYTH and VAN STADEN (1983), DRENNAN and VAN STADEN (1989), DREWES (1989) and DREWES and VAN STADEN (1990).

After collection of the achenes, they were laid out to dry for six weeks before any further studies were carried out, since DREWES (1989) showed that *T. minuta*

achenes have a short afterripening requirement. After this period, achenes had a moisture content of $10.5 \pm 3.4\%$. Although much higher than the $5.2 \pm 0.3\%$ reported by DREWES (1989), this figure is in keeping with the achene moisture contents of $11.37 \pm 3.42\%$ and $11.7 \pm 1.3\%$ reported by FORSYTH and VAN STADEN (1983) and DRENNAN and VAN STADEN (1989). Achenes were then stored at 10°C over silica gel whilst initial germination tests were conducted. All germination tests were conducted in illuminated growth chambers, as FORSYTH and VAN STADEN (1983) have shown that *T. minuta* achenes have a light requirement for germination. When achenes were imbibed at 25°C , radicle germination was first observed in the sample after 14-18 h. Thereafter germination proceeded rapidly and germination was generally completed by between 42 and 48 h of imbibition (Figure 2.1). Germination was considered to be complete once no changes in the level of germination had been observed for a period of more than 6 h. Although no viability tests were performed on the achenes, the embryos proving to be impossible to isolate and hence making

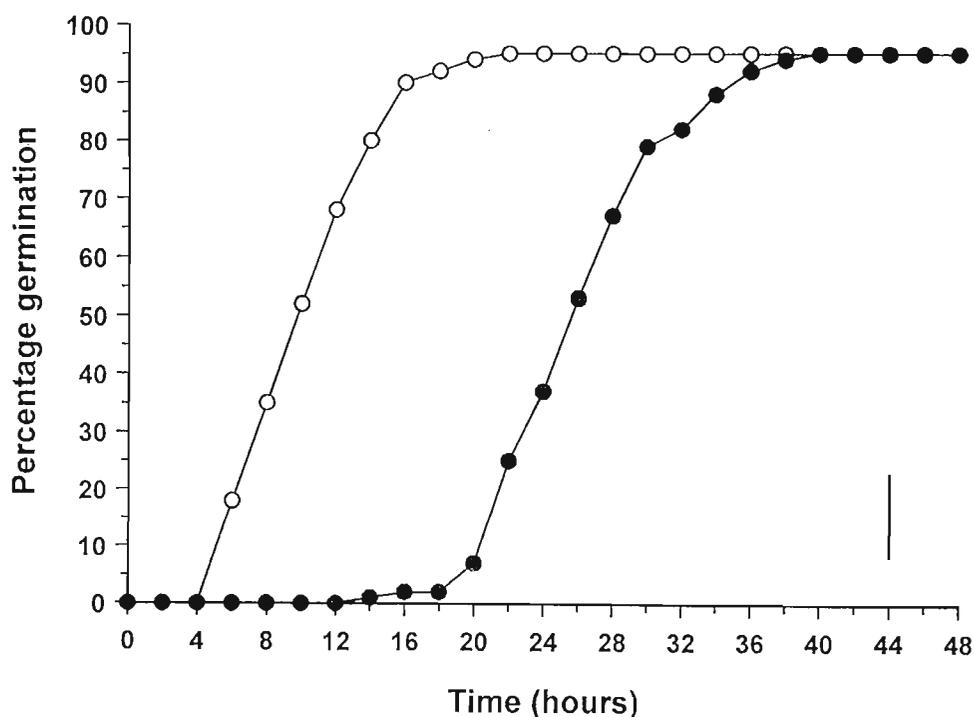


Figure 2.1: Germination curves of *Tagetes minuta* achenes imbibed directly at 25°C (●) and following 72 h pre-treatment at 36°C (○). Vertical bar = maximum standard error of the mean.

tetrazolium testing non-feasible, in all germination tests conducted more than 96% germination was achieved. At 36°C, as expected, no germination was observed. FORSYTH and VAN STADEN (1983) reported that after 4 d of incubation at 35°C, germination of *T. minuta* achenes was accelerated when the temperature was reduced to 25°C, so that 100% germination was achieved within 24 h. Similar results were observed by DRENNAN and VAN STADEN (1989), DREWES (1989) and DREWES and VAN STADEN (1990). The achenes tested in this study proved to be no different, and a similar result was observed after 3 d of high-temperature pre-treatment (Figure 2.1). Radicle emergence was first observed 4 h after the temperature shift and germination was completed within 20-24 h. The germination curve produced by such achenes is, however, identical to that produced by achenes germinated at 25°C. The observed difference between the two curves translates into a mere temporal shift of an otherwise identical germination rate. Pre-treating achenes at 36°C for periods of 24, 48, 72, 96 and 120 h all produced similar results (Figure 2.2), being consistent with the results obtained by DREWES (1989).

It appears that, during pre-treatment of seeds at thermoinhibitory temperatures, these seeds are able to undergo the majority of the germination-related processes that take place in seeds germinated at more favourable temperatures (HEYDECKER, 1977). At some point, however, the high temperatures form a restraint against further development, leading to the arrest of the germination process, or thermoinhibition. When thermoinhibition is relieved through a reduction in the incubation temperature, these seeds are already primed for germination, resulting in the "acceleration" in the rate of germination which has been observed. Although only 24 h of imbibition at 36°C was required to induce the full 'accelerated germination' effect, 72 h of imbibition at 36°C was selected as the standard thermoinhibitory treatment for use in this study. Since germination was spread over 48 h when achenes were imbibed at 25°C, it is likely that not all achenes imbibed at 36°C for periods of less than 48 h have reached the same stage of development. Since the thermoinhibitory block presumably affects all achenes at the same point in the germination process, allowing achenes to imbibe at 36°C for longer than 48 h is likely to allow all achenes to reach this point, hence producing a more uniform population for molecular studies. A

further advantage is that this longer incubation period allows for visual confirmation that achenes are indeed thermoinhibited and not merely germinating more slowly than usual.

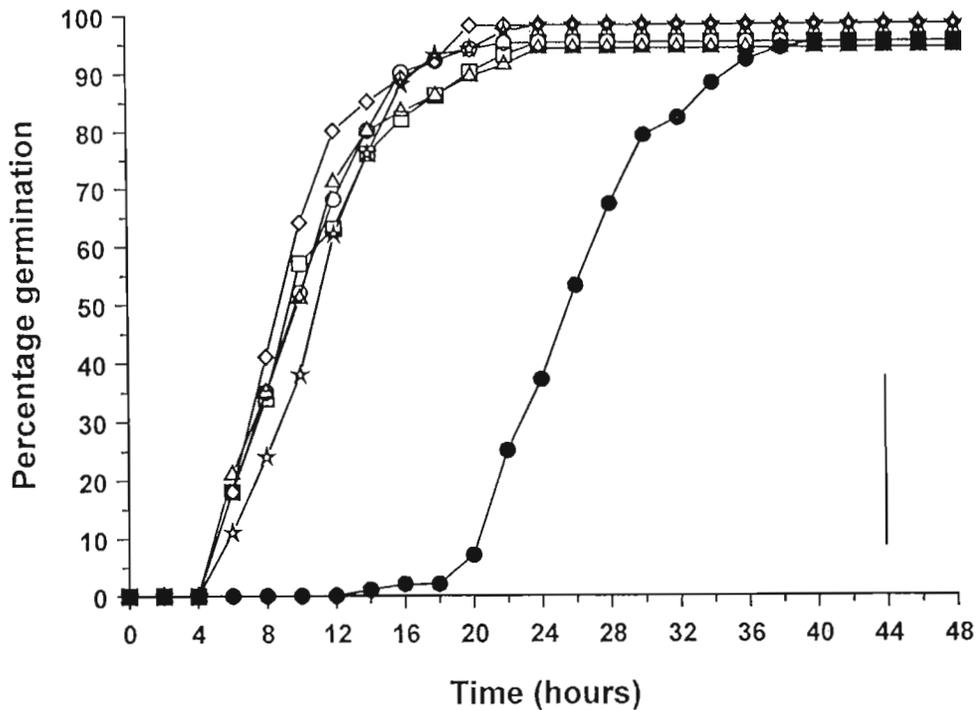


Figure 2.2: Germination curves of *Tagetes minuta* achenes imbibed at 36°C for varying periods before transfer to an incubation temperature of 25°C. (●) 25°C control; (□) 24 h at 36°C; (△) 48 h at 36°C; (○) 72 h at 36°C; (◇) 96 h at 36°C; (☆) 120 h at 36°C. Vertical bar = maximum standard error of the mean.

2.3.2 Imbibition curves

Imbibition occurred more quickly at 35°C than at 25°C (Figure 2.3). This is in keeping with the results observed by DREWES (1989). The initial rapid uptake of water characteristic of Stage I of the typical water uptake curve (BEWLEY and BLACK, 1994) lasted approximately 8-10 h, slightly longer than the 6 h reported by DREWES (1989). After 10 h, achenes imbibing at both temperatures entered a lag phase

(Stage II) where the increase in mass was very slight over the next 10 h. The first signs of radicle emergence in the achenes incubated at 25°C was observed after 18 h. After 20 h, the imbibition rate of these achenes accelerated sharply and the water uptake curve entered into Stage III, where water uptake is driven by cell expansion and vacuolation of the growing axis and cotyledons (OBROUCHEVA and ANTIPOVA, 1997). The achenes held at 35°C did not germinate and the imbibition curve for these achenes did not progress beyond Stage II. As was observed by DREWES (1989), the fact that imbibition did occur in these thermoinhibited achenes is an indication that thermoinhibition in this species is not due to the isolation of the embryo from the environment at elevated temperatures. Further confirmation that the embryo coverings play no part in thermoinhibition was obtained when achenes from which the tip had been removed at the radicle end (the end opposite the pappus) also became thermoinhibited when germinated at 35°C. The removal of the tip had no influence on germination, which proceeded normally when the temperature was reduced to 25°C (results not shown).

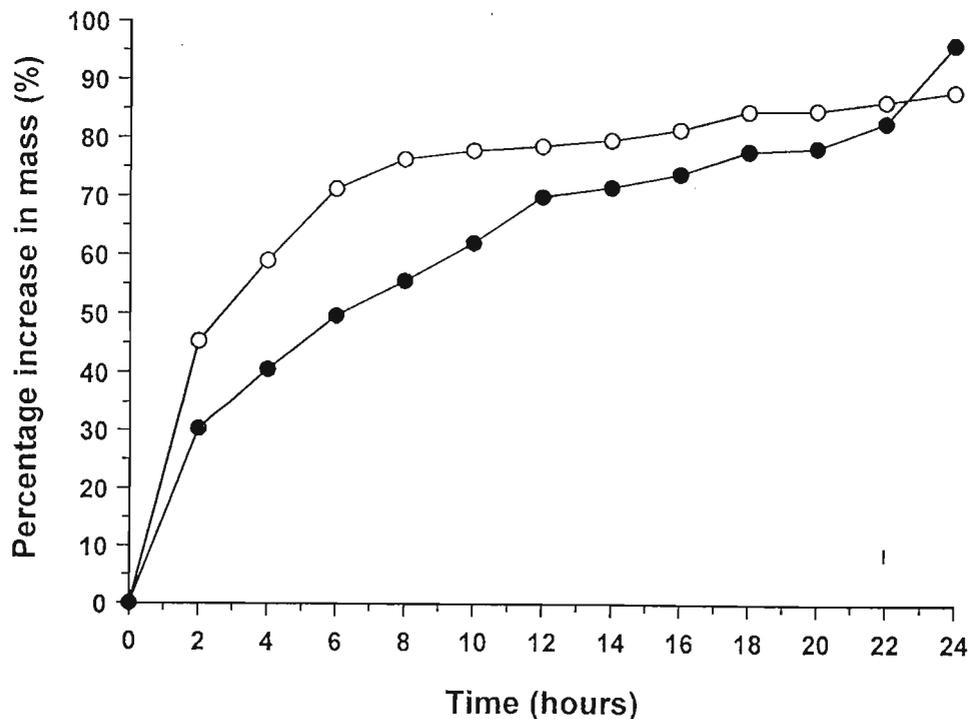


Figure 2.3: Imbibition curves for *Tagetes minuta* achenes imbibed at 25°C (●) and at 36°C (○). Vertical bar = maximum standard error of the mean.

2.3.3 Freezing trials

DREWES (1989) reported that achenes appeared to be able to germinate to full capacity for a period of only nine months following this afterripening period, and that the level of germination in the seed lot dropped by approximately 20% after one year of storage. A sample of achenes were therefore frozen over silica gel at -20°C for 7 d and their germination characteristics compared with those of a control group of achenes which had not been frozen, to determine whether storage at sub-zero temperatures was a viable method of delaying ageing effects. DREWES (1989) showed that storage of achenes over silica gel at -10°C resulted in no significant changes in their germination characteristics. Although achenes stored at -20°C did appear to germinate fractionally slower when achenes were germinated immediately at 25°C and following 72 h of high-temperature pre-treatment, there was no major change in their germination characteristics (Figure 2.4). After one year of storage, one half of the achenes collected were therefore stored at -20°C to ensure a continuous supply of achenes with identical germination characteristics for the duration of this study. At this stage, the moisture content of the seeds had dropped to $4.48 \pm 0.65\%$. Germination trials were repeated every two months to ensure that no change in germination efficiency had occurred in the stored achenes. Over a period of 18 months, no changes in germination rates or levels were observed in either the cold- (10°C) or freezer- (-20°C) stored achenes. Ageing effects were, however, noticed in a sample of achenes stored at room temperature over the same period. These changes were in keeping with those reported by DREWES (1989) (results not shown). It therefore appears that storage at 10°C over silica gel is an effective means of reducing ageing effects in achenes of *T. minuta* in the short term. With extended storage (longer than two years), however, achenes stored at 10°C began to show slower germination characteristics and a slight loss in viability (results not shown).

Only freezer-stored achenes were used in experiments after 18 months of storage. After 6 years the germination characteristics of achenes stored at -20°C had not changed from the original germination experiments (results not shown). For long

term storage of achenes, therefore, it appears that storage at -20°C is recommended.

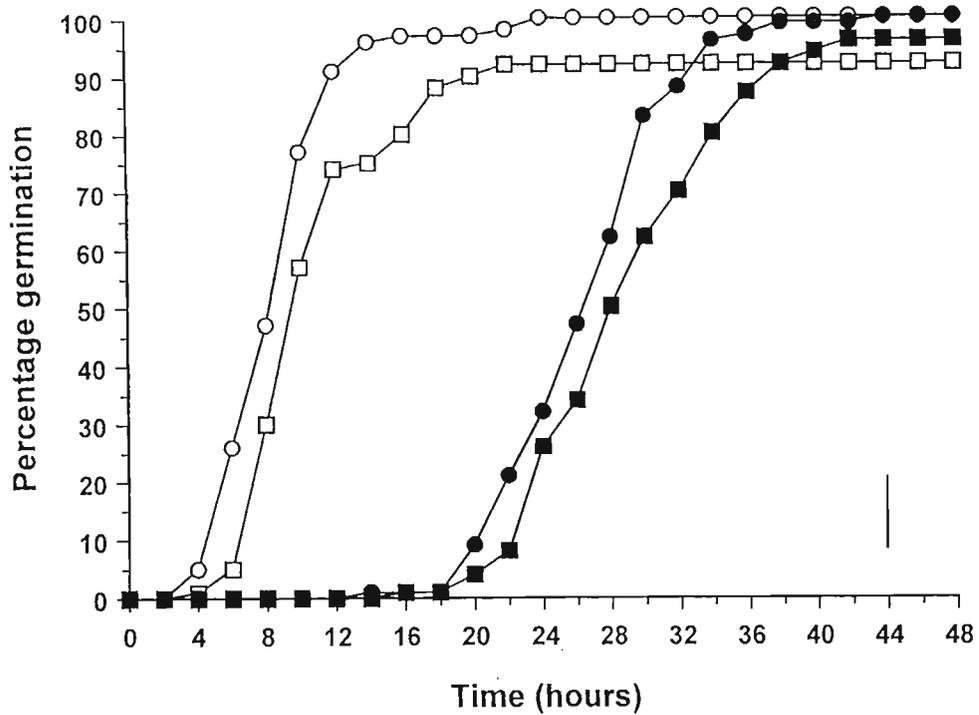


Figure 2.4: Germination curves of *Tagetes minuta* achenes stored at 10°C and -20°C when imbibed directly at 25°C and following 72 h imbibition at 36°C . (●), 10°C storage, no pre-treatment; (■) -20°C , no pre-treatment; (○), 10°C , 36°C pre-treatment; (□), -20°C storage, 36°C pre-treatment. Bar = maximum standard error of the mean.

2.3.4 Effect of gibberellic acid on germination

DREWES and VAN STADEN (1990) reported that the application of exogenous GA_3 or GA_{4+7} to thermoinhibited achenes of *T. minuta* resulted in a partial alleviation of thermoinhibition. The application of GA_{4+7} was more effective than application of GA_3 , resulting in a maximum level of germination of approximately 75% being attained at 35°C following a 1 h pulse of this GA mix at a concentration of 3×10^{-4} M. Since it was thought that this treatment could provide a means of confirming the involvement in thermoinhibition of any proteins or mRNAs identified as being differentially expressed in thermoinhibited achenes in this study by determining

whether or not levels of these factors were affected when GA was applied, this experiment was repeated. However, despite several attempts, the maximum level of germination observed at 36°C was 24±19.2%, also obtained after a 1 h pulse treatment with GA₄₊₇ (ProVide[®], Abbot Laboratories) at a concentration of 3x10⁻⁴ M (results not shown). Germination was spread over a period of 14 d. The fact that consistently low levels of germination were achieved and that germination was not synchronized meant that this experiment was abandoned. Although, as shown by DREWES and VAN STADEN (1990), thermoinhibition in achenes of *T. minuta* achenes can to some extent be overcome by gibberellins, the extent of this relief appears to differ widely between different populations of achenes.

2.4 CONCLUSION

In general, the germination patterns observed in germination trials using the achenes collected for this study showed almost identical trends to those reported in previous studies of *T. minuta*. When achenes are imbibed at 25°C, radicle emergence may first be observed within the sample group of achenes after between 14 h and 18 h. Thereafter, germination is rapid and all viable achenes in the sample germinate over the next 22 h so that germination takes approximately 48 h. At 36°C, however, the achenes become thermoinhibited and do not germinate until the incubation temperature is reduced. After as little as 24 h of pre-treatment at 36°C, the achenes appear to be primed for germination so that an 'accelerated germination' effect is observed following the temperature shift. Radicle emergence can be seen after 4 h at 25°C and germination is completed within 24 h of the temperature being reduced. There is no benefit in pre-treating the achenes for longer periods at elevated temperatures. Nevertheless, 72 h imbibition at 36°C was selected as the standard thermoinhibitory treatment to be used in this study as this allows all achenes to reach a similar stage of development, ensuring greater uniformity of the study samples. This is of particular value for molecular studies since it ensures that any subtle changes in the protein or mRNA populations between thermoinhibited and non-thermoinhibited achenes will be most pronounced.

Although gibberellins have been shown to result in a partial alleviation of thermoinhibition in *T. minuta* achenes, this resulted in less than 25% germination being achieved over a two week period. This differed from the previous report (DREWES and VAN STADEN, 1990), where up to 75% germination was observed at 36°C. The difference in the extent of this relief from thermoinhibition is probably due to differences between the two populations of achenes used in these different studies. It does, however, mean that the effect of GA on the levels of any proteins or mRNAs differentially expressed in thermoinhibited and non-thermoinhibited achenes cannot be analyzed for the achenes used in this study.

CHAPTER 3

PROTEIN ANALYSIS

3.1 INTRODUCTION

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) is probably the most widely used technique in the protein sciences. The basis of this technique is that a charged molecule will migrate in an electric field at a rate which is determined by both its size and its charge. In SDS-PAGE, proteins are denatured and reacted with SDS, an anionic detergent, to form negatively-charged complexes (SMITH, 1994). In the denatured state, most proteins bind SDS in a constant weight ratio of 1.4 g SDS per gram protein (SMITH, 1994), so that the proteins end up with similar charge densities (COPELAND, 1994). When an electric field is placed across an polyacrylamide gel, which acts as a barrier to molecular motion, the SDS-coated proteins will migrate through the gel under the influence of the electrical field. The rate of migration will, however, no longer depend on the inherent charge of the protein, but solely on the basis of size since larger proteins will be more severely impeded by the polymeric gel than smaller proteins (COPELAND, 1994).

The resolution of SDS-PAGE is enhanced by casting a stacking gel of a different acrylamide concentration and pH on top of the separating gel. When an electric field is applied to the gel in the presence of a glycerol-containing electrophoresis buffer, ions will move towards the electrodes. At the pH used in the stacking gels (typically pH 6.8), the mobility of the negatively-charged SDS-protein complexes lies between that of Cl⁻ ions present in the system and of the glycinate ions in the buffer (SMITH, 1994). The Cl⁻ ions have the greatest mobility and are known as leading ions, whilst the glycinate ions have the lowest mobility at this pH and are hence referred to as the trailing ions. The various ions thus concentrate into narrow zones in the stacking gel. When the interface between the stacking and separating gels is reached, the change in pH alters the relative mobilities of the various ions. The glycinate ions overtake the SDS-protein complex zone, compressing the proteins into a narrow band and leaving

these molecules in an evenly-buffered electrical field that allows for increased separation and resolution in the separating gel (SMITH, 1994). This is the basic principle behind SDS-PAGE as described by LAEMMLI (1970).

More recently, a system of SDS-PAGE using tricine buffers has been developed (SCHÄGGER and VON JAGOW, 1987). This technique was mainly developed to allow for the separation of proteins of less than 10 kDa, since proteins of this size are not effectively separated in the LAEMMLI (1970) system. The system is, however, able to separate proteins in the range 1-100 kDa (SCHÄGGER and VON JAGOW, 1987) and is the preferred method for SDS-PAGE in some laboratories.

Single-dimension SDS-PAGE is, however, fairly limiting in terms of the number of protein bands which can be resolved. Furthermore, several proteins of similar molecular weight can co-migrate as a single band and hence not be distinguished as separate proteins. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) provides one of the most powerful tools for the separation of proteins from complex mixtures (POLLARD, 1994). By far the most commonly followed procedure is that developed by O'FARRELL (1975) and involves separation of proteins based on both their isoelectric point and their molecular weight. In the first dimension, isoelectric focusing in tube gels is used to separate proteins based on their isoelectric points. These gels are then laid across the top of a slab gel and separated on the basis of molecular weight by the SDS-PAGE second dimension gel. Since the technique allows up to 1800 individual proteins to be separated on a single gel (POLLARD, 1994), it has been extensively used to study proteins whose expression is modified by various external factors or which are developmentally-regulated (BAUW, VAN MONTAGU and INZE, 1992).

The aim of these protein studies was to examine the proteins produced by both germinating and thermoinhibited achenes to determine whether or not any significant changes in gene expression could be identified which might be related to thermoinhibition.

3.2 MATERIALS AND METHODS

3.2.1 Protein extraction

All protein extractions were conducted using 0.25 g samples of achenes. Samples were ground to a fine powder in a pre-chilled mortar and pestle with the aid of liquid nitrogen. The powder was then added to 1 ml of protein extraction buffer (0.1 M potassium phosphate buffer, pH 7.0; 30% [v/v] glycerol; 0.3 M NaCl; 0.1 M DTT; 1 mM EDTA; 1 mM PMSF; 1.5% [w/v] PVPP) in a centrifuge tube and stirred magnetically on ice for 30 min. Debris was pelleted by centrifugation at 12 000 g at 4°C for 10 min. The supernatant liquid was transferred to a sterile 1.9 ml microfuge tube using a sterile 1 ml micropipette. Any remaining debris was pelleted at maximum speed for 10 min at ambient temperature in a Sigma 113 desktop microfuge. The supernatant liquid was then transferred to a fresh microfuge tube and stored at -20°C.

3.2.2 Quantification of protein samples

The concentration of proteins in extracts was measured using the BioRad Protein Assay Kit, which is based on the Coomassie blue dye-binding method developed by BRADFORD (1976). A dilution series of a bovine serum albumin (BSA) protein standard was prepared, ranging in concentration from 200 µg/ml to 1400 µg/ml, with increments of 200 µg/ml. A new standard curve was prepared each time the assay was performed to eliminate errors introduced by slight changes in the concentration of the dye reagent, which was freshly prepared for each assay. This was necessary as once the concentrated dye reagent provided in the kit has been diluted to the working concentration its storage life is very short. Thus, only enough dilute reagent was prepared as was needed for each assay. One part of the five-fold concentrated dye reagent was diluted with four parts ddH₂O and filtered through Whatman's No. 1 filter paper. For construction of the standard curve, 20 µl of the standard solutions were placed into clear plastic cuvettes with 1 µl extraction buffer. Dye reagent was

then added to a total volume of 1 ml and the contents of the cuvette mixed thoroughly by inversion. After 5 min, optical density (OD) was measured at 595 nm in a Beckman DU-65 spectrophotometer. For quantification of samples, 1 μ l of the protein extract was placed in a cuvette and dye reagent added to a total of 1 ml. After 5 min, OD₅₉₅ was measured against a reagent blank (1 μ l extraction buffer in dye reagent). All readings were done in triplicate and the mean values used in further calculations. The mean absorption values for the concentration standards were subjected to linear regression. The mean OD readings of the protein extracts were then substituted into the regression equation to estimate their concentration.

3.2.3 Polyacrylamide gel electrophoresis of proteins

3.2.3.1 SDS-polyacrylamide gel electrophoresis (Laemmli system)

Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) was performed according to SMITH (1994), a method which is based on the original LAEMMLI (1970) protocol. Proteins were separated on a 1.5 mm thick gel consisting of a stacking gel (pH 6.8) and a separating gel (pH 8.8). Thirty millilitres of separating gel solution (Appendix A) were prepared from 15 ml acrylamide stock (Appendix A), 7.5 ml ddH₂O and 7.5 ml separating gel buffer (Appendix A). This solution was then degassed under vacuum and polymerization initiated by the addition of 45 μ l 10% [w/v] ammonium persulfate (APS) and 15 μ l NNN'N'tetramethylethylenediamine (Temed). The gel was then poured between two ultra clean glass plates, prepared by washing each one twice with warm, soapy water, followed by two washes with absolute alcohol and one wash with acetone. Plates were then buffed using a clean, lint-free towel. The gel was allowed to polymerize for 1 h after the addition of a ddH₂O overlay. This was found to be as effective as a water-saturated butanol overlay as suggested in the original protocol, which has the disadvantage that the top of the separating gel must be rinsed several times with ddH₂O to remove all traces of butanol which can interfere with polymerization of the stacking gel. Once the gel had completely polymerized, the ddH₂O overlay was poured off and the stacking gel

poured.

The stacking gel solution was prepared by mixing 750 μ l acrylamide stock and 3 ml ddH₂O. After degassing, 1.25 ml stacking gel buffer (Appendix A) was carefully added and polymerization initiated by adding 15 μ l 10% [w/v] APS and 5 μ l Temed. The stacking gel was then poured onto the separating gel and the comb inserted between the plates to form the loading wells. The stacking gel extended below the comb to a depth of 1 cm.

Protein extracts stored at -20°C were thawed and aliquots containing 10 μ g protein each were made up to 10 μ l with ddH₂O. These samples were then mixed with an equal volume of double-strength sample solvent and denatured by boiling for 5 min. A mixture of proteins of known molecular weight from the Combithek Calibration Proteins kit (Roche, no longer available) was also prepared and electrophoresed in the outer lanes of each gel to allow for size determination of the unknown proteins. Details of this mixture are given in Table 3.1. Samples were electrophoresed at a constant current of 10 mA through the stacking gel. Once the dye front reached the

Table 3.1: Details of proteins from the Combithek Calibration Proteins kit (Roche) and volumes used in the preparation of the molecular weight marker mixture for polyacrylamide gel electrophoresis. This kit is no longer marketed.

| Protein | Concentration of stock | M _r | Volume |
|-----------------------------|------------------------|----------------|--------------|
| β -galactosidase | 0.5 mg/ml | 116 400 Da | 0.5 μ l |
| Fructose-6-phosphate kinase | 0.65 mg/ml | 85 200 Da | 0.39 μ l |
| Glutamate dehydrogenase | 0.5 mg/ml | 55 600 Da | 0.5 μ l |
| Triose-phosphate isomerase | 0.5 mg/ml | 26 600 Da | 0.25 μ l |
| Aldolase | 0.65 mg/ml | 39 200 Da | 0.39 μ l |
| Trypsin inhibitor | 0.8 mg/ml | 20 100 Da | 0.16 μ l |
| Lysozyme | 0.5 mg/ml | 14 300 Da | 0.25 μ l |

interphase between the stacking and the separating gels, the current was increased to 25 mA until the bromophenol blue had reached the end of the gel. As the dye front migrated approximately 2-3 cm ahead of the first detectable protein bands, the gels were then allowed to run for a further 40 min at 20 mA to allow for the greatest possible separation of the protein bands.

3.2.3.2 Tricine-SDS-polyacrylamide gel electrophoresis

In order to determine the best system for single-dimensional separation of the proteins in the extracts, the Tricine-SDS-PAGE method of SCHÄGGER and VON JAGOW (1987) was also tested. Stock solutions (Appendix A) were stored at 4°C. The acrylamide-*bis*acrylamide stock solution was filtered through Whatman's No. 1 filter paper and was stored in the refrigerator in a brown bottle.

As with the Laemmli-system gels, proteins were separated on 1.5 mm thick gels consisting of a stacking and a separating gel. The composition of these gels is given in Table 3.2. After degassing the separating gel mixture, polymerization was initiated by adding 45 µl 10% (w/v) APS and 15 µl Temed. The gel was overlaid with ddH₂O until polymerization was complete. Thirty microlitres of 10% (w/v) APS and 10 µl Temed were then added to the degassed stacking gel solution. This was then poured onto the separating gel and the comb inserted. As with the Laemmli gels, the stacking gel extended to a depth of 1 cm below the base of the loading wells.

Table 3.2: Composition of stacking and separating gels for Tricine-SDS-PAGE (SCHÄGGER and VON JAGOW, 1987).

| Stock solution | Stacking gel (4% T, 3% C) | Separating gel (10% T, 3% C) |
|-----------------------------------|------------------------------|---------------------------------|
| Acrylamide- <i>bis</i> acrylamide | 1.0 ml | 6.1 ml |
| Gel buffer | 3.1 ml | 10.0 ml |
| ddH ₂ O | 8.4 ml | 13.9 ml |
| Total | 12.5 ml | 30 ml |

Protein samples were incubated at 40°C for 30 min in the loading buffer before electrophoresis. Unlike the Laemmli system, Tricine-SDS-PAGE gels were run at constant voltage. Voltage was set at 70 V through the stacking gel and was increased to 100 V as the dye front entered the separating gel.

3.2.3.3 Two-dimensional-polyacrylamide gel electrophoresis

Two-dimensional electrophoresis was conducted according to O'FARRELL (1975), with modifications as set out by MAYER, HAHNE, PALMER and SCHELL (1987). Glass tubes with an internal diameter of 1.5 mm were cleaned by soaking in chromic acid for 30 min. The acid was then washed out with ddH₂O and the tubes neutralized with 2% (w/v) potassium hydroxide dissolved in 95% ethanol. This was then followed by several washes with ddH₂O. Using a micropipette, 1 ml of ethanol was run through each tube, followed by 1 ml of acetone. Finally, 250 µl of Gel Slick™ (AT Biochem) was run through the tubes to prevent the gels from adhering to the glass, and the tubes dried thoroughly in a drying oven.

Isoelectric focusing (IEF) gels (4.2%) were prepared from acrylamide stock (29.16:1.33 acrylamide:*bis*acrylamide), 5% (v/v) ampholytes (Isolyte 3-10, Pharmacia Biotech), 4% Nonidet P-40 (NP-40) and 9.0 M urea. All solutions were prepared with ddH₂O. The gel solution was filtered through a sterile 0.22 µm filter and then briefly degassed. Polymerization was initiated by adding 2 µl of 20% (w/v) APS and 0.7 µl Temed per ml gel solution. Tubes were marked to a distance of 9 cm from the base of the tube. The gel solution was then poured into the glass tubes to this level using a hypodermic syringe with a long needle, taking care not to trap air bubbles in the tubes. Finally, the gels were overlaid with ddH₂O and allowed to polymerize fully.

Protein extracts were prepared for IEF by adding NP-40 to a final concentration of 4% (v/v) and urea crystals to a concentration of 9 M. Gels were pre-run to establish the pH gradient before loading. During the pre-run, voltage was increased in a step-wise fashion from 200 V (20 min) to 300V (20 min) and finally to 400V (20 min). Proteins

were then applied to the basic end of the gels. Routinely, 20 µg protein were loaded per gel. Isoelectric focusing was conducted for 20 h at 400 V. Cathode and anode electrolytes were 0.02 M NaOH and 0.01 M phosphoric acid respectively. After focusing gels were extruded from the glass tubes and the acidic end marked by briefly dipping it into a 0.05% bromophenol blue solution. Focused gels were equilibrated for 3 min in equilibration buffer (Appendix A) before storage at -70°C or being subjected to the second dimension of electrophoresis. The second dimension was conducted using a standard SDS-PAGE slab gel, as described in Section 3.2.3.1. Where IEF gels had been stored at -70°C, these were first thawed to room temperature in equilibration buffer before electrophoresis. The second dimension was then run under the same electrophoresis conditions as previously described for SDS-PAGE.

3.2.4 Detection of electrophoresed proteins

Proteins separated on both single- and two-dimensional gels were detected by silver-staining. All single-dimension gels were subjected to a two-stage pre-fixing step consisting of a 20 minute wash in a solution containing 30% (v/v) ethanol and 10% (v/v) glacial acetic acid followed by a second 20 min wash in a 30% ethanol / 7% acetic acid solution. Gels were then fixed for 20 min in a 10% (v/v) glutaraldehyde solution. After fixing, the gels were soaked overnight in ddH₂O. Before continuing with the staining procedure, the gels were rinsed in fresh ddH₂O for 30 min. Gels were stained for 20 min in a 0.5% (w/v) solution of silver nitrate and then rinsed well with water to remove unbound silver. To develop the stain the gels were soaked in developing solution (3% [w/v] sodium carbonate, 0.0005% [v/v] formaldehyde) until the desired level of staining was achieved. The reaction was then terminated by adding 10 ml acetic acid and agitating for 10 min. Finally, gels were rinsed well with several volumes ddH₂O. After being photographed, gels were stored in a small amount of ddH₂O in sealed plastic bags at 4°C.

Two-dimensional gels were stained according to MORRISSEY (1981). Gels were

pre-fixed for 30 min in pre-fixing solution I (50% [v/v] methanol, 10% [v/v] acetic acid) and a further 30 min in pre-fixing solution II (5% [v/v] methanol, 7% [v/v] acetic acid). Fixing was achieved by soaking gels for 30 min in 10% (v/v) glutaraldehyde. Gels were washed overnight and then rinsed again briefly the following morning in ddH₂O. Before adding the silver stain, the gels were soaked for 30 min in a 5 µg/ml solution of DTT. This was then poured off and, without rinsing the gel, the 0.1% (w/v) silver nitrate solution added. After staining for 30 min, the gels were rinsed briefly with two rapid washes of ddH₂O and a third quick rinse in a small volume of developer. Gels were then soaked in developer until staining reached the required levels, when the reaction was again stopped by adding 10 ml acetic acid. Gels were then washed, photographed and stored as before.

3.3 RESULTS AND DISCUSSION

3.3.1 Optimization of protein extraction buffer

In order to determine the optimal extraction buffer for the isolation and stable storage of proteins from *Tagetes minuta* achenes, a stability trial was carried out in a manner similar to that outlined by GEGENHEIMER (1990). Samples of 0.25 g of achenes were incubated in dH₂O for 72 h at 36°C. These were then ground to a fine powder in a pre-chilled mortar and pestle using liquid nitrogen. Before it could thaw, the powder was added to 1 ml of the appropriate buffer under test in a sterile 1.5 ml microfuge tube. Seven different extraction buffers were tested, as follows:

- i) 0.1 M potassium phosphate buffer, pH 7.0; 30% (v/v) glycerol;
- ii) 0.1 M potassium phosphate buffer, pH 7.0; 30% (v/v) glycerol; 1.5% (w/v) polyvinylpyrrolidone (PVPP); 0.3 M NaCl; 0.1 M dithiothreitol (DTT); 1 mM ethylene diamine tetraacetate (EDTA); 1 mM phenyl methyl sulfonyl fluoride (PMSF), 2% (v/v) Triton X-100;
- iii) as for (ii), but with DTT replaced by 14 mM 2-mercaptoethanol;
- iv) as for (ii), but without Triton X-100;
- v) as for (iii), but without Triton X-100;

- vi) as for (ii), but with EDTA replaced by 1 mM ethylene glycol *bis*(β -aminoethyl ether) NNN'N'-tetraacetic acid (EGTA);
- vii) as for (iii), but with EDTA replaced by 1 mM EGTA; and
- viii) as for (ii), but with Triton X-100 replaced with 1% Nonidet P40 (NP40).

Extracts were incubated on ice for 30 min and then centrifuged for 10 min at ambient temperature in a Sigma 113 desktop microfuge. The supernatant liquid was then transferred to fresh microfuge tubes and re-centrifuged. Samples were analyzed by SDS-PAGE electrophoresis. Aliquots of 1 μ l of each sample were run immediately after extraction, whilst the remainder of each isolate was stored at -20°C. To determine stability of the samples during storage, further 1 μ l aliquots of each sample were electrophoresed after periods of 24 h and 4 weeks of storage .

Extraction buffers containing DTT as an anti-oxidant appeared to produce better results than those containing 2-mercaptoethanol. Samples extracted in buffers containing 2-mercaptoethanol tended to show a greater degree of background staining than those extracted in DTT-containing buffers . This effect has been noted in two-dimensional gels (BEIS and LAZOU, 1990). Furthermore, 2-mercaptoethanol has been reported to cause staining artefacts in the form of intense horizontal bands in the range 50 to 65 kDa in two-dimensional gels which have been stained with silver (TASHEVA and DESSEV, 1983; BEIS and LAZOU, 1990).

Although samples extracted in buffers containing surfactants (Triton X-100, NP40) tended to produce slightly crisper bands than those without, these detergents were eventually left out of the buffer. This was done because these samples tended to precipitate out of solution after denaturing whilst they were being loaded onto the gels, and it was feared that some less abundant proteins may be lost. This effect is rather difficult to understand as surfactants are generally added to protein extraction buffers to increase the solubility of normally insoluble proteins.

The metalloprotease inhibitors ethylene diamine tetraacetate (EDTA) and ethylene glycol *bis*(β -aminoethyl ether) N,N,N',N'-tetraacetic acid (EGTA) were compared.

There appeared to be no difference in the efficiency of these two compounds and EDTA was selected as the cheaper option. On the basis of these results, extraction buffer (iv) was selected as the standard extraction buffer to be used in all further analyses. Typically, the use of this buffer, along with the modified extraction procedure described in Section 3.2.1 resulted in the isolation of between 60 and 80 μg protein per gram fresh weight of achenes. There was much variation in the amount of protein isolated from achenes subjected to different incubation temperatures and times, although the amount of protein isolated from different samples of achenes treated in the same fashion did not differ very much. Typically, the amount of protein isolated from achene samples where radicle emergence had occurred in a large proportion of the achenes (for example, between 36 and 48 h of incubation at 25°C) was approximately half of that obtained from achene samples where little or no radicle emergence was observed (Figure 3.1).

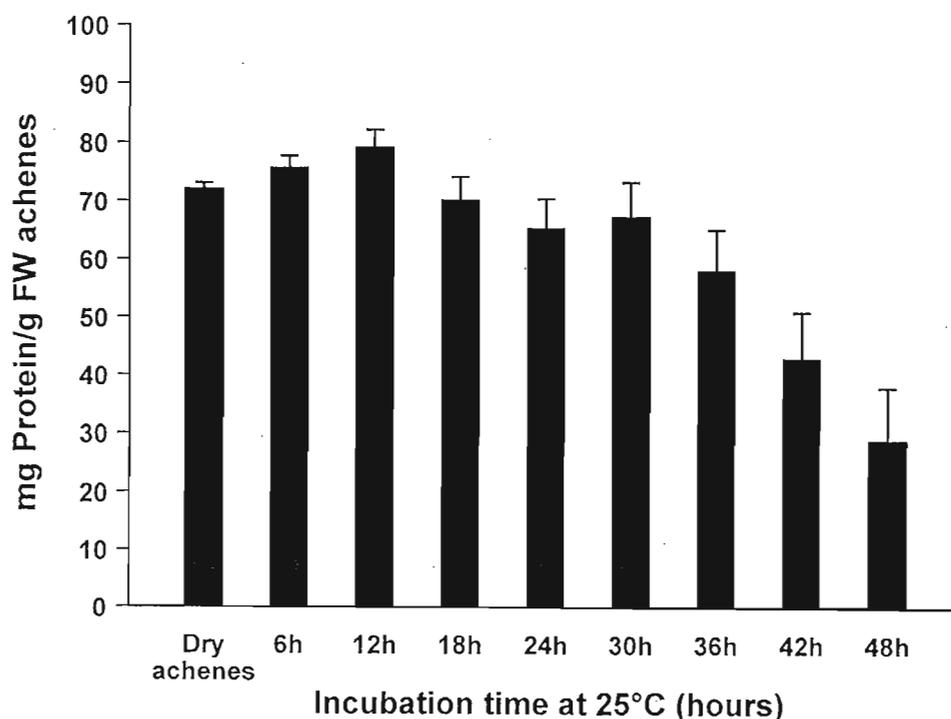


Figure 3.1: Mean amounts of protein (mg/g FW) isolated from *Tagetes minuta* achenes imbibed at 25°C for varying periods of time.

A similar result was observed when thermoinhibited achenes were allowed to germinate (result not shown). This phenomenon correlated with a noticeable greening of the protein isolate, presumably due to the presence of photosynthetic pigments from the young seedlings. Although the polypeptide patterns of these samples were largely obscured, it was not considered necessary to modify the extraction buffer to eliminate this problem as these samples were representative of achenes in the post-germinative phase. Protein isolates from thermoinhibited achenes and achenes still undergoing germination and thus of greatest interest for the purposes of this study were not affected.

3.3.2 Comparison of standard SDS-PAGE (Laemmli system) and Tricine-SDS-PAGE

In order to determine the best method for separation of the proteins in the samples, the standard SDS-PAGE technique based on the original method of Laemmli was compared to the Tricine-SDS-PAGE technique developed by SCHÄGGER and VON JAGOW (1987) (Figure 3.2). Tricine-SDS-PAGE resulted in slightly better separation of protein bands above about 55 kDa and below 14 kDa. Resolution of bands in the size range between these two extremes was, however, far superior in the standard SDS-PAGE gels. The vast majority of proteins resolved in both systems ranged in size from about 14 kDa to 85 kDa. Standard SDS-PAGE was thus selected over Tricine-SDS-PAGE as the method of choice for this study.

3.3.3 Staining of two-dimensional gels

Two-dimensional gels were initially stained using the same staining protocol as was used for staining the single-dimensional gels. This method was not successful in resolving the polypeptides separated by 2D-PAGE. Polypeptides were stained as

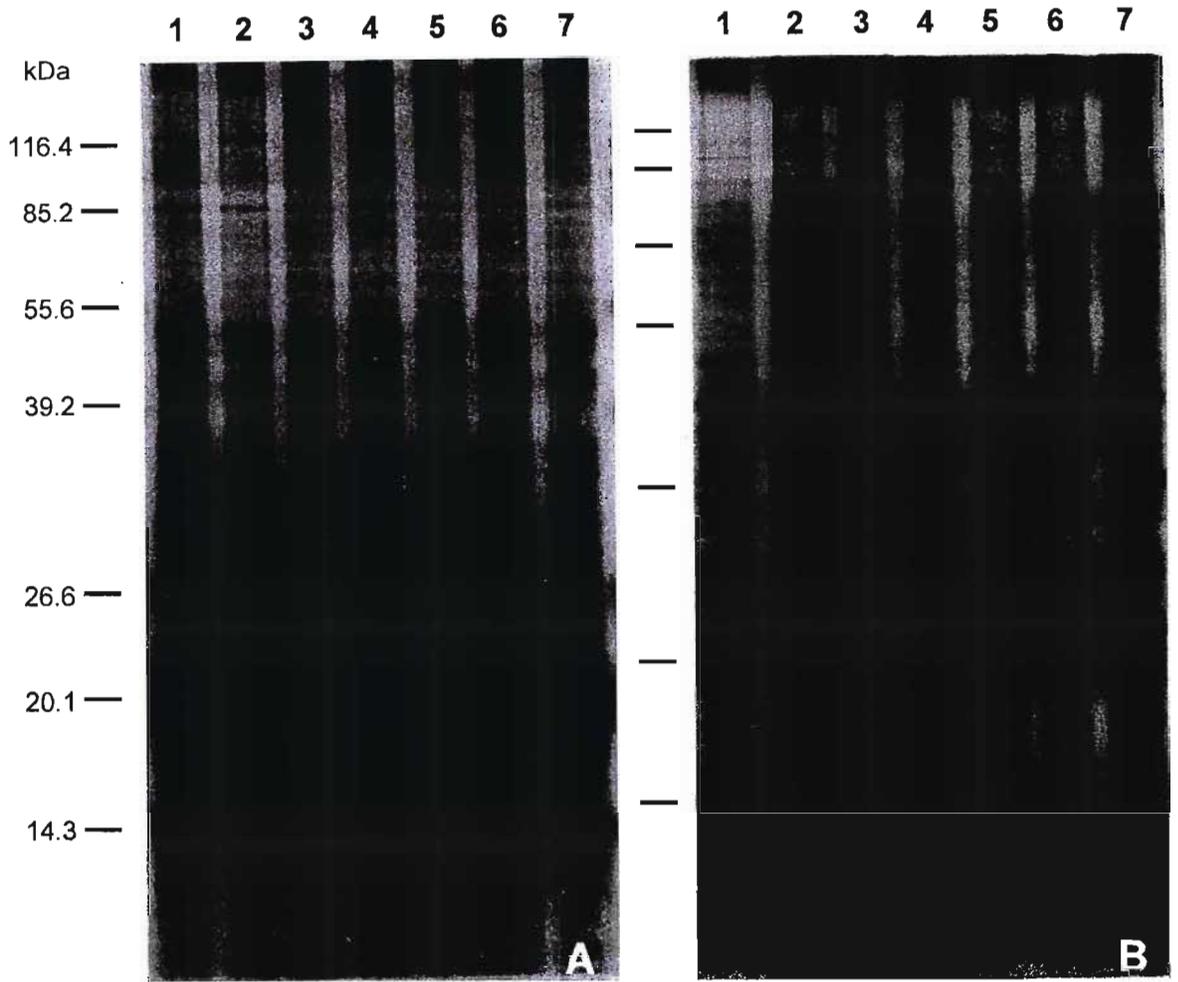
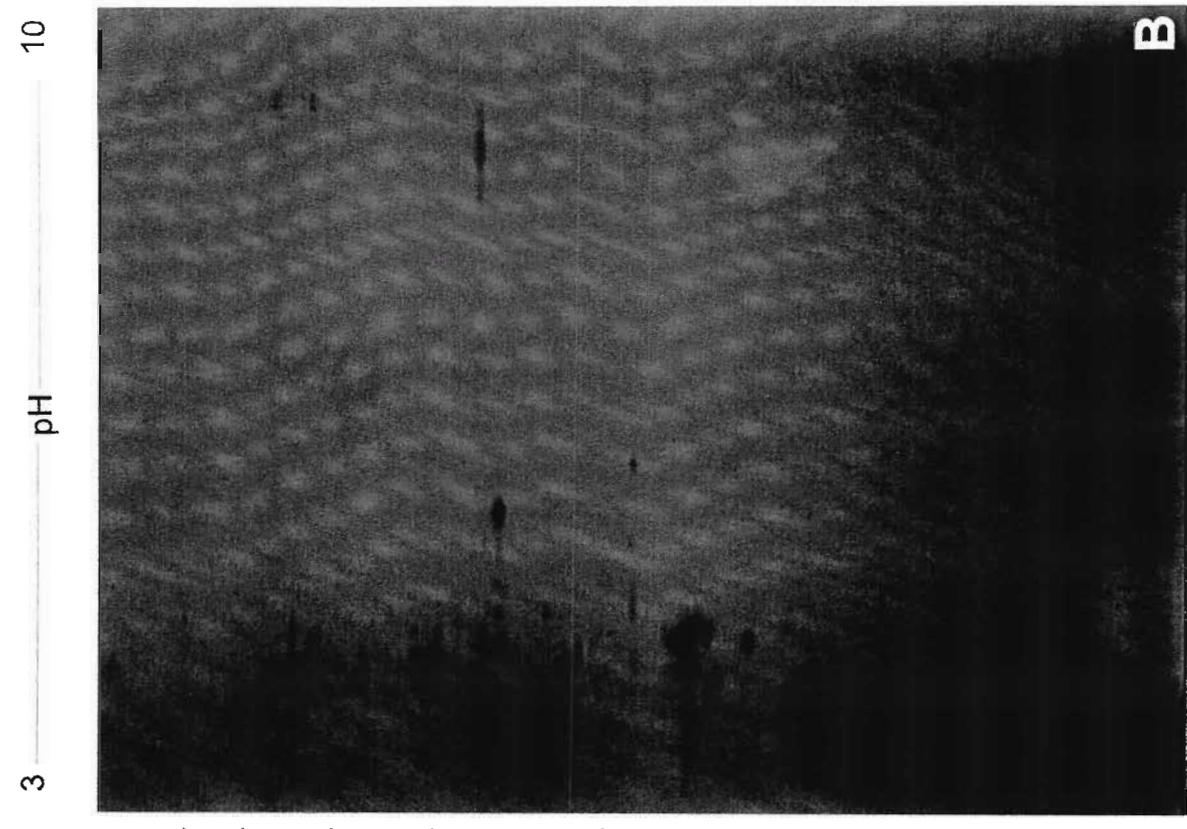


Figure 3.2: Comparison of separation of proteins isolated from thermoinhibited achenes by Tricine-SDS-PAGE (a) and standard SDS-PAGE (Laemmli system) (b). Lane 1 contains proteins from dry achenes, lanes 2-7 contain proteins from achenes imbibed at 36°C for 24 h, 48 h, 72 h, 96 h and 120 h respectively.

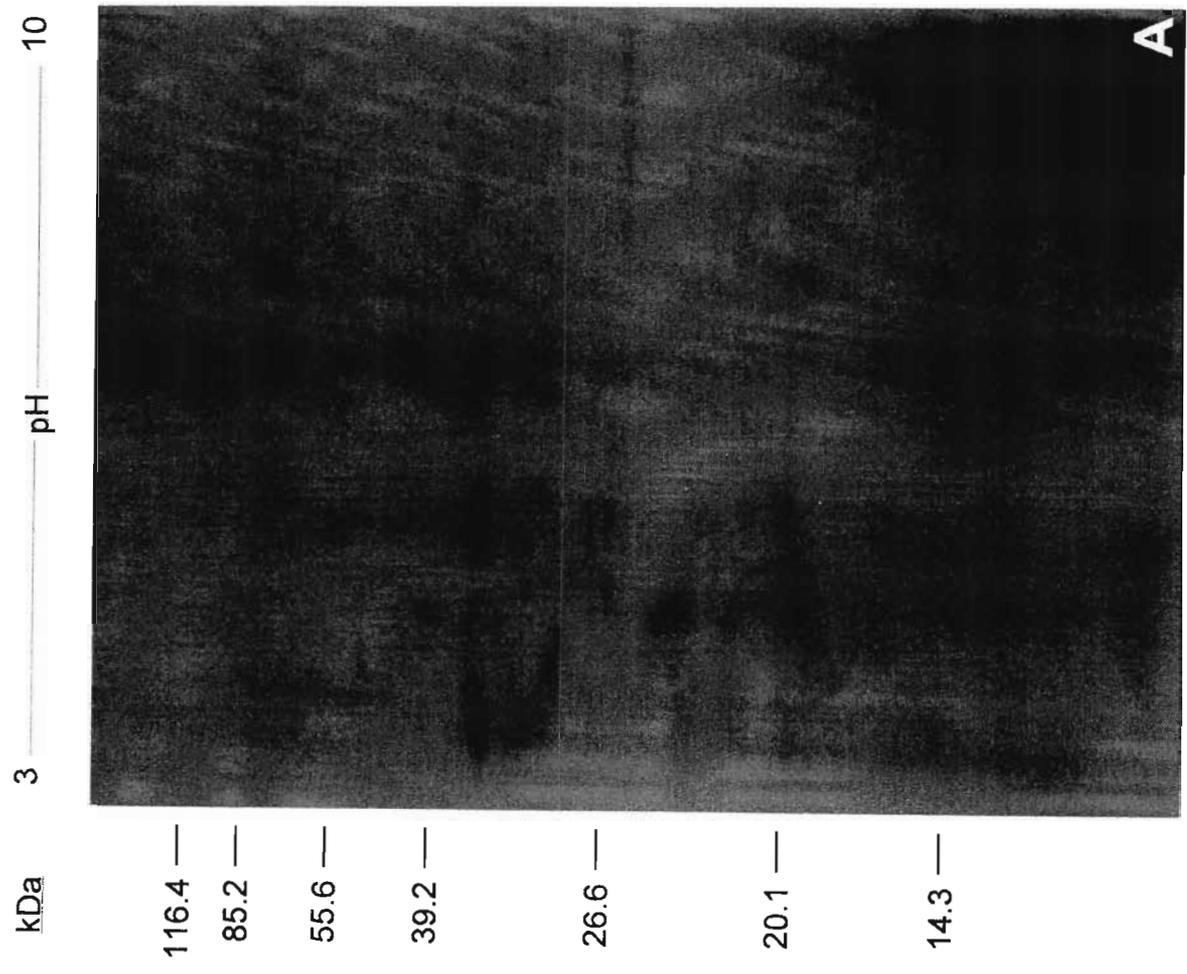
large, indistinct areas. Polypeptides of similar size and pI could not be distinguished as separate entities and less abundant peptides were not consistently resolved in replicate gels. Furthermore, severe background staining problems were encountered which obscured the polypeptide patterns even further (Figure 3.3A).

To try to improve this situation, gels were stained using the staining protocol developed by MORRISSEY (1981). This protocol involves washing the fixed gel with the reducing agent dithiothreitol (DTT) before silver staining. Since silver staining probably involves the reduction of silver ions to silver metal by proteins, the highest

Figure 3.3: Comparison of two-dimensional gels stained in the same manner as used for single dimension gels (A) and using the MORRISSEY (1981) method (B).



B



A

degree of sensitivity will be obtained when the proteins are fully reduced (MORRISSEY, 1981). During the isolation of proteins, different proteins may be subjected to varying degrees of oxidation in the different tissues (MORRISSEY, 1981) and cellular compartments. Since all proteins are reduced to the same extent by this wash, the method also results in reduced variation in staining intensity between replicates. Using this method, over 180 different polypeptides could be resolved (Figure 3.3B).

3.3.4 Single-dimension SDS-PAGE

The use of single-dimension SDS-PAGE did not identify any protein bands associated with thermoinhibition. Banding patterns were completely reproducible in all replicates of the various gels. No changes were observed in the banding patterns produced by protein extracts from dry achenes and achenes imbibed at 36°C for periods of 24 h, 48 h, 72 h, 96 h, 120 h and 240 h (Figure 3.4). A number of differences were, however, observed during the course of germination of achenes imbibed at 25°C (Figure 3.5). Five polypeptides (d1 to d5) present in the dry achenes decreased in concentration as germination progressed. A further four polypeptides (i1 to i4) appeared to increase in concentration (i1, i2 and i4) or were synthesized *de novo* (i3) in the germinating achenes. Polypeptide i3 appeared at slightly different times in the three replicates, appearing as a very faint band after 18 h in the first replicate. In the two subsequent replicates, i3 only appeared after 24 h, but at approximately the same staining intensity as this band at the same time in the first replicate. Very little could be observed in lanes containing samples of protein isolated from achenes imbibed for more than 36 h (Figure 3.5). At this time, background staining in these lanes became severe and by 42 h of imbibition virtually nothing could be detected in these lanes. This increase in background staining also correlated with a decrease in the amount of soluble protein which could be isolated from these achenes using the standard extraction technique. As discussed previously (Section 3.3.1), these isolates were distinctly green, probably due to the presence of chlorophylls from the growing seedlings, the majority of which had produced the first green leaflets by this stage. The increase in background staining

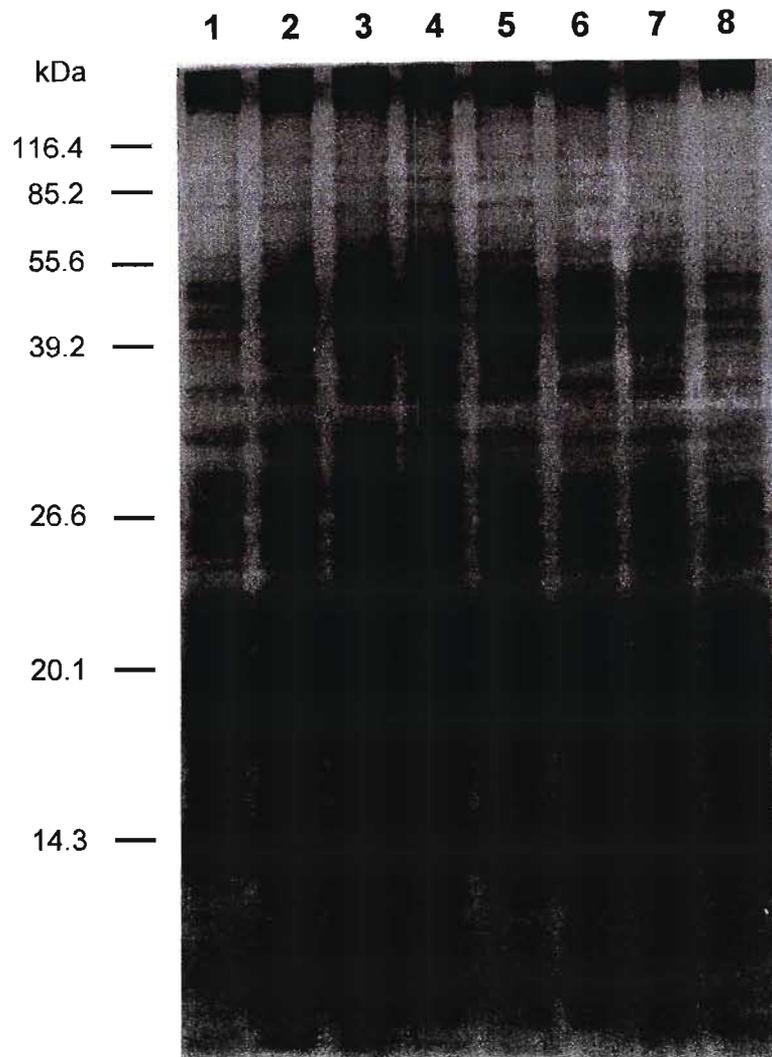


Figure 3.4: SDS-PAGE separation of proteins isolated from dry achenes (lane 1), achenes incubated in a dry state at 36°C for 10 d (lane 2) and thermoinhibited achenes imbibed at 36°C for periods of 24 h (lane 3), 48 h (lane 4), 72 h (lane 5), 96 h (lane 6), 120 h (lane 7) and 240 h (lane 8).

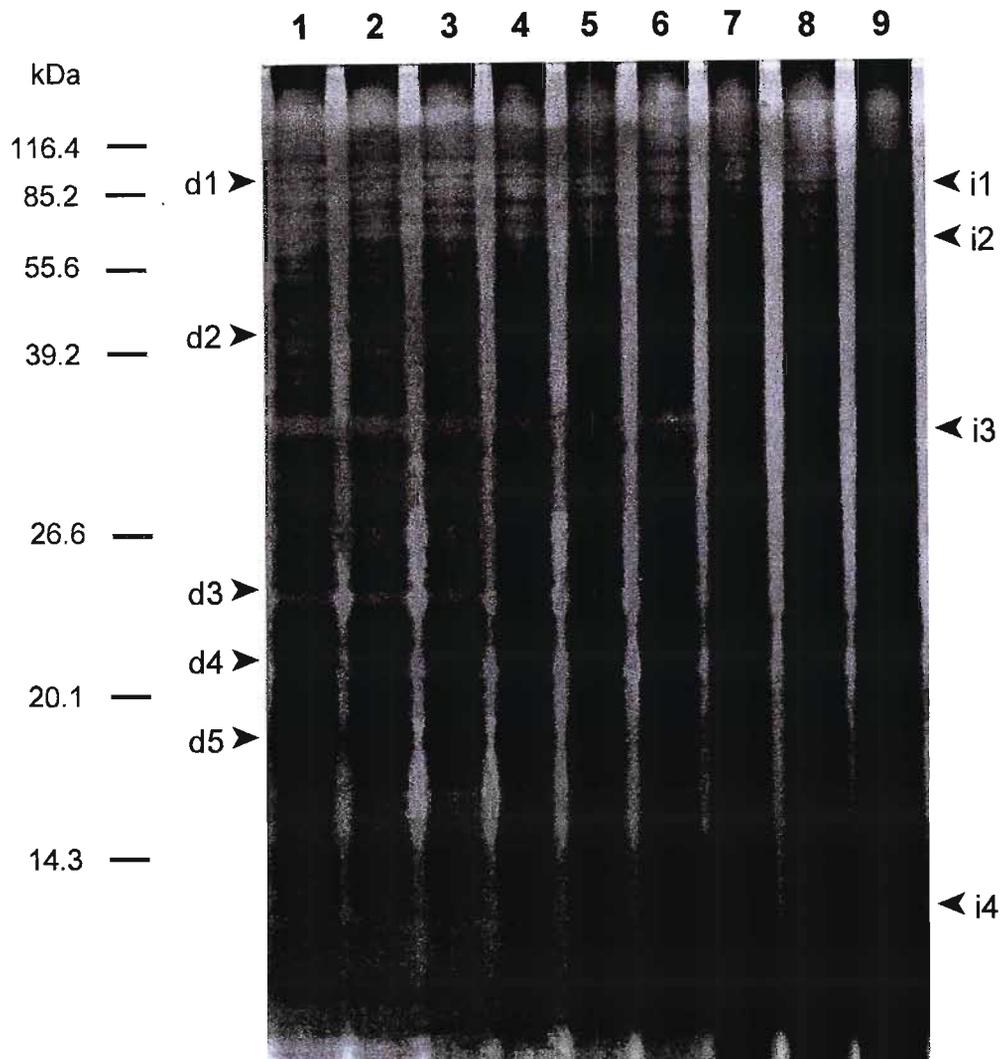


Figure 3.5: SDS-PAGE separation of proteins from germinating achenes imbibed for varying periods at 25°C. Lane 1 contains proteins isolated from dry achenes, lanes 2-9 contain proteins from achenes imbibed for 6 h, 12 h, 18 h, 24 h, 30 h, 36 h, 42 h and 48 h respectively. Levels of polypeptides labeled d1-d5 increase during germination, whilst levels of polypeptides i1-i4 decline during the germination process.

is thus possibly due to the presence of various photosynthetic pigments in the samples. Similar changes in polypeptides were also apparent when thermoinhibited achenes were removed to the 25°C incubator and allowed to continue germinating (Figure 3.6). Polypeptide i3 first appeared after 8-10 h at 25°C. This corresponds well with the appearance of this polypeptide in achenes which were not exposed to the thermoinhibitory treatment. In both cases, the appearance of this polypeptide corresponds to approximately 35% germination in the germination trials (Figure 2.1, Section 2.3.1). Background staining was also observed in lanes containing samples from achenes which had been allowed to complete germination and had begun to produce photosynthetic shoots.

The fact that no differences were observed between germinating and thermoinhibited achenes was not surprising, since SDS-PAGE was only able to separate about 40 distinct polypeptide bands, whereas seeds produce thousands of proteins. A similar result was obtained by MAHHOU and DENNIS (1994), who concluded that SDS-PAGE was not sufficiently powerful a technique to resolve proteins involved in breaking dormancy. If thermoinhibition is caused by specific gene expression which in some way arrests the course of germination, it is likely that this is due to the expression of a small number of genes and hence, proteins. Although no previous molecular studies on thermoinhibition have been conducted, several studies involving seed dormancy in different species where dormancy is suspected to be actively imposed in imbibing seeds by differential gene expression (e.g. DYER, 1993; HANCE and BEVINGTON, 1992) have only identified a few differentially-expressed proteins in the dormant seeds of these various species which may be associated with dormancy. It would consequently be easily possible to overlook those few proteins specific to thermoinhibited achenes amongst the myriad of other constitutively-expressed and germination-related proteins, particularly in SDS-PAGE gels where several polypeptides of similar molecular weight can co-migrate as a single band. Two-dimensional SDS-PAGE offers a much greater likelihood of resolving any differentially-expressed polypeptides as polypeptides are separated in terms of both their molecular weights and their isoelectric points.

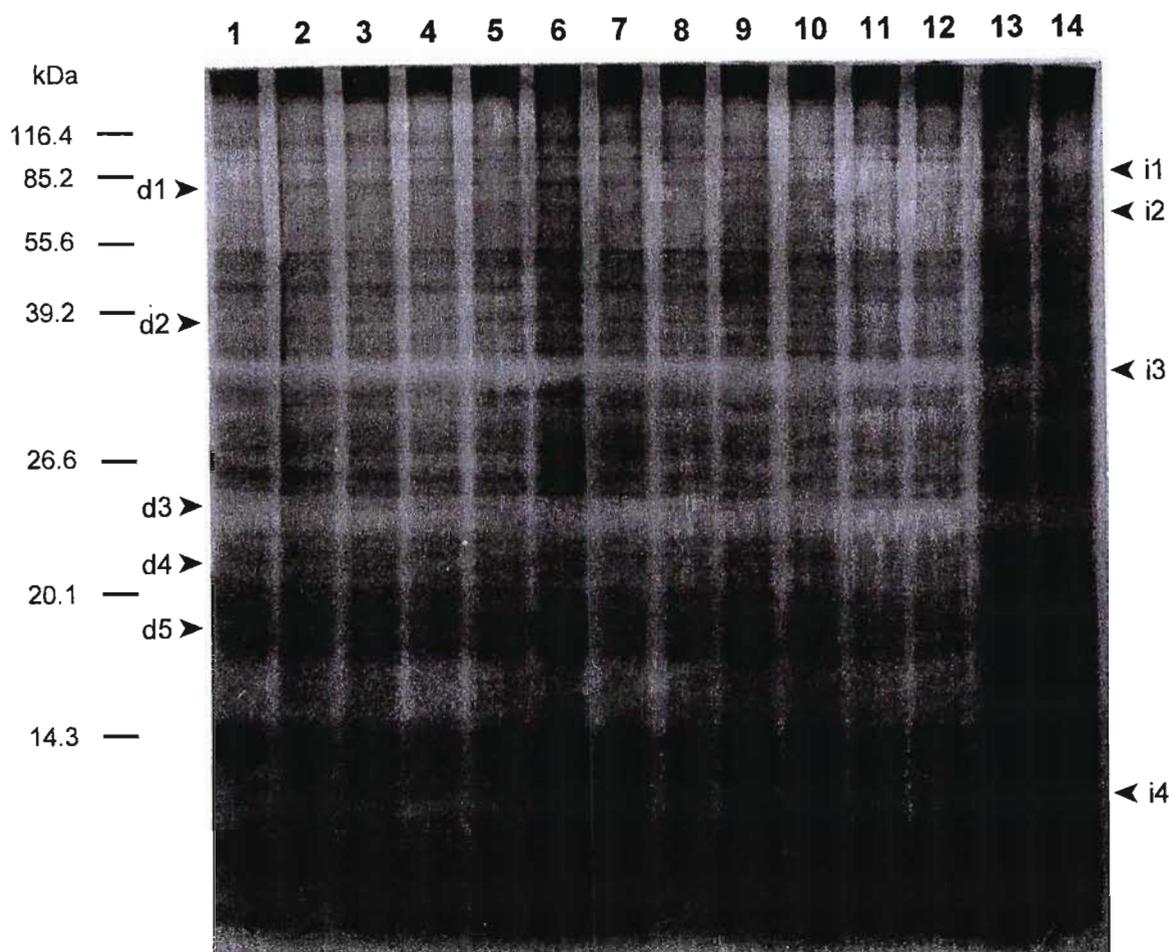


Figure 3.6: SDS-PAGE separation of proteins from achenes germinated for varying periods following 72 h imbibition at 36°C. Lane 1 contains proteins from dry achenes. Proteins from thermoinhibited achenes (72 h at 36°C) are shown in lane 2. Lanes 3-14 contain proteins from pre-treated achenes which have been allowed to germinate at 25°C for between 2 h (lane 3) and 24 h (lane 14). Samples were taken at two-hourly intervals.

3.3.5 Two-dimensional polyacrylamide gel electrophoresis

In contrast to the relatively small number of polypeptides resolved by single-dimension SDS-PAGE, 2D-PAGE resulted in the separation of approximately 200 polypeptides. As was the case with single-dimension SDS-PAGE, gels were almost perfectly reproducible. The vast majority of polypeptides isolated were found in the acidic regions of the gels for all samples. Relatively few neutral to basic polypeptides were observed in dry achenes, but those polypeptides which were synthesized *de*

novo during the course of germination appeared to be more evenly distributed over a broader range of pI values.

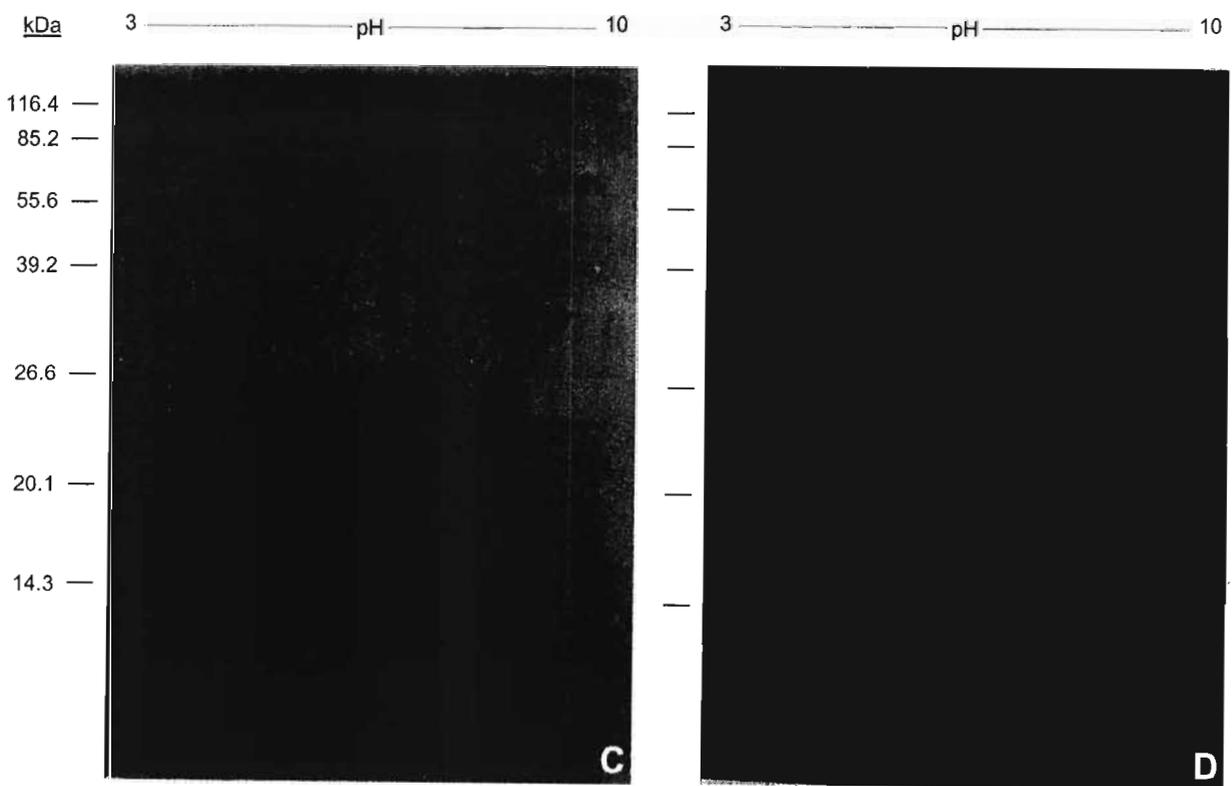
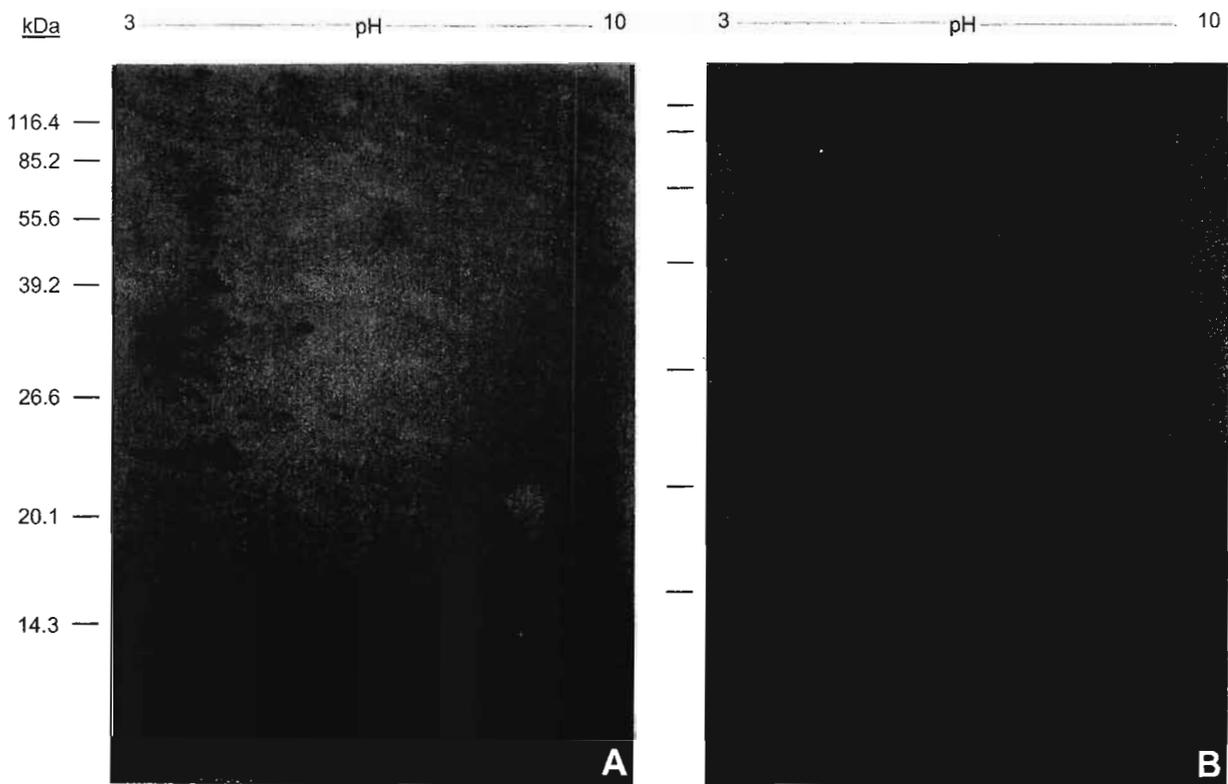
Four main classes of polypeptides could be distinguished during the course of germination at the optimum temperature for germination of *T. minuta* achenes of 25°C (Figure 3.7, Figure 3.8A). Approximately 70 of the polypeptides isolated appeared to be constitutively expressed (uncoloured polypeptides, Figure 3.8A). Levels of these polypeptides did not change during germination from the levels established in the dry achenes. A second subset of about 20 polypeptides decreased in abundance during germination (polypeptides marked in blue, Figure 3.8A). Some of these merely appeared to be present at lower levels in germinated achenes than in dry achenes, whilst others disappeared completely during the germination process. This group probably represents proteins involved in late embryogenesis and whose expression is switched off following desiccation. A pair of small acidic polypeptides (marked in purple, Figure 3.8A) which were only observed in the dry achenes probably belongs in this class.

A larger group of about 40 polypeptides showed a steady increase in concentration in protein isolations upon imbibition. Of these, a small group of four basic polypeptides (marked in orange, Figure 3.8A) increased up to 12 h of imbibition and then decreased slightly. All others of this class appeared to increase steadily as germination progressed (marked in green, Figure 3.8A). It is likely that this group includes proteins related to the germination process.

The fourth major group, also comprising some 40 polypeptides (marked in yellow, Figure 3.8A), includes those proteins which were not present in dry achenes but were synthesized *de novo* in imbibing achenes. Of these, the vast majority appeared after 12 h of imbibition, shortly before the first visible signs of germination (radicle emergence) became evident in the samples. Relatively few were synthesized between 12 h and 36 h after germination. Two-dimensional electrophoresis of protein isolates from achenes imbibed for 48 h at 25°C was conducted, but, as was the case with the single-dimension gels, background staining was sufficiently severe

Figure 3.7: Two-dimensional polyacrylamide gels showing changes in polypeptide patterns in germinating achenes imbibed at 25°C. Polypeptides were isolated from dry achenes (A) and achenes germinated for 12 h (B), 24 h (C) and 36 h (D). These three time intervals correspond to germination percentages of approximately 0%, 40% and 95% respectively.

Figure 3.8: Changes in polypeptide patterns in achenes germinating at 25°C (A) and during thermoinhibition at 36°C (B). These diagrams were prepared from data obtained from the gels shown in Figures 3.7 and 3.9. Uncoloured polypeptides show no changes in expression when dry achenes are imbibed. Levels of polypeptides marked in green increase during imbibition, whilst levels of polypeptides marked in blue decline upon imbibition. Polypeptides marked in orange show a slight increase up to 12 h of imbibition at 25°C and subsequently decline. Polypeptides marked in yellow are not present in dry achenes and are synthesized *de novo* upon rehydration, at the times indicated on the diagram (12 h, 24 h or 36 h). The polypeptides marked in purple in (A), were specific to the dry achenes at 25°C, but are still present in thermoinhibited achenes, albeit at a lower level. Polypeptides marked in red are specific to thermoinhibited achenes and may play a role in the imposition of thermoinhibition of these achenes.

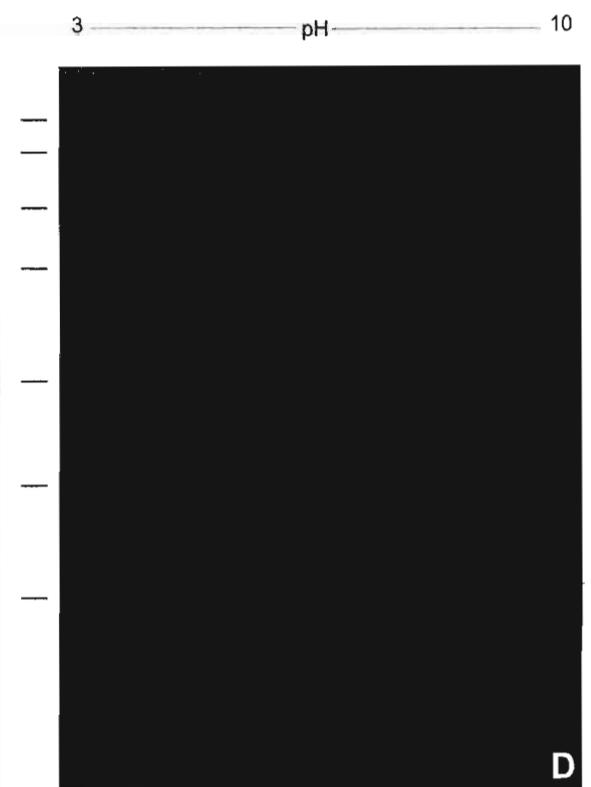
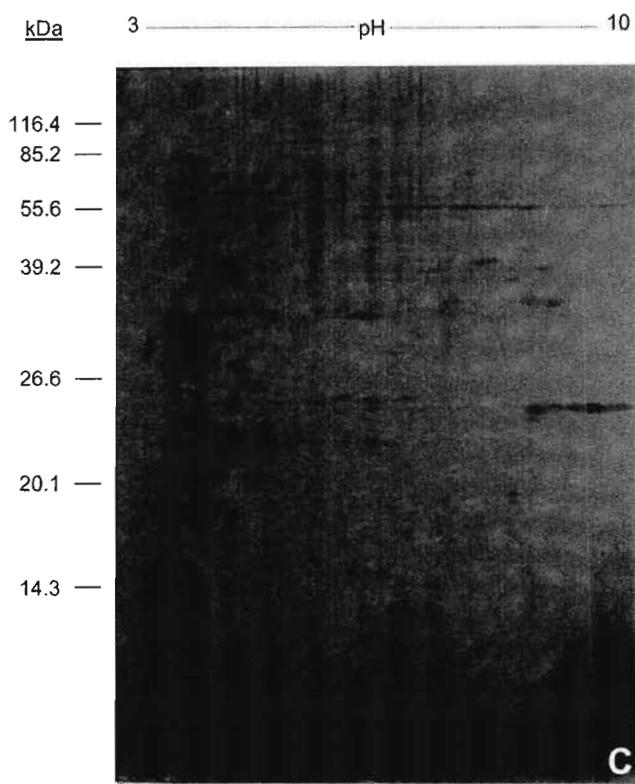
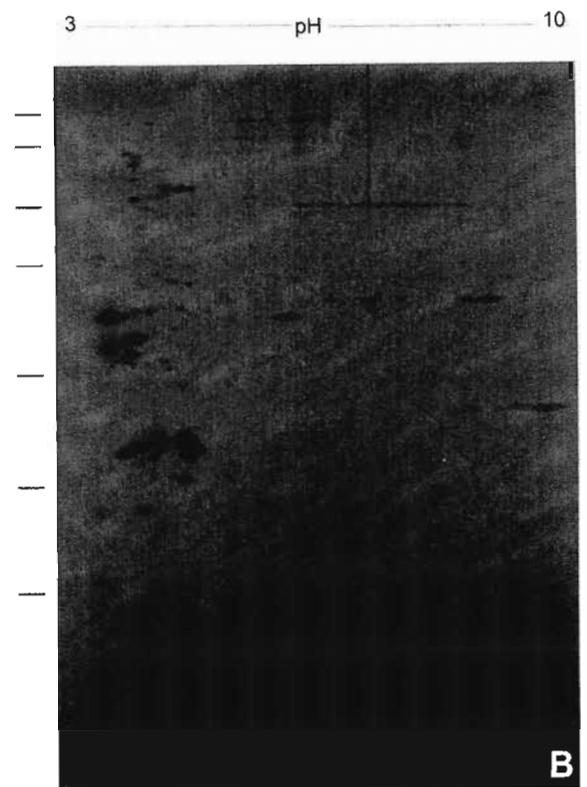
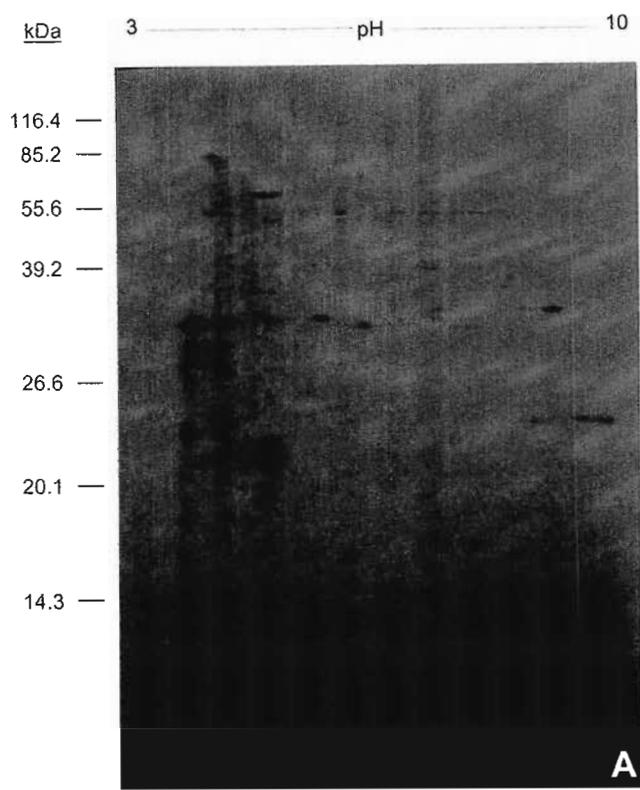


as to make these gels worthless. It is likely that the number of proteins present in the germinating seedlings does increase drastically after 36 h of imbibition, however this is obviously a post-germinative event and beyond the scope of this study.

When achenes were imbibed at 36°C, similar trends were evident (Figure 3.8B, Figure 3.9). No changes were observed amongst those polypeptides which were constitutively expressed during normal germination. Most of the differences observed were in polypeptides from the other three categories. Several of those polypeptides which were observed to decrease during imbibition at 25°C remained at a constant level in thermoinhibited achenes. These may represent heat shock proteins or LEA-like proteins involved in protecting the embryo during heat stress. The pair of polypeptides which were previously only observed in the dry achenes were still visible after 72 h at 36°C, although they appeared to have decreased in abundance. Of those proteins which increased in abundance during imbibition at 25°C, most followed the same pattern at 36°C. Eight large, acidic polypeptides in this category, ranging in size from 56 kDa to 85 kDa did not increase in abundance, however.

Less than half of the polypeptides synthesized *de novo* in germinating achenes were observed in extracts from thermoinhibited achenes. None of those polypeptides synthesized during the later stages of germination were synthesized in the thermoinhibited achenes. However, a new set of 10 polypeptides (marked in red, Figure 3.8B) was synthesized in thermoinhibited achenes. These polypeptides were clustered into two separate groups of polypeptides of similar size. The larger group contained seven polypeptides, ranging in size from 22 kDa to 26.7 kDa and with isoelectric points between pH 3.0 and 4.0. The second group of three polypeptides had molecular weights of between 14 kDa and 15 kDa, with pI values of approximately 3.0. These polypeptides were all extremely specific to the thermoinhibited achenes. Once the temperature inhibition was lifted and the achenes allowed to continue with the germination process at 25°C, these polypeptides declined rapidly, some within 2 h of the temperature shift. The remainder declined to low levels by 6 h and were no longer visible after 20 h at 25°C. These polypeptides are only synthesized in metabolically active, imbibing seeds. Where dry

Figure 3.9: Polypeptide patterns obtained by 2D-PAGE of protein isolates from achenes of *Tagetes minuta* which had been thermoinhibited (A) and subsequently allowed to germinate at 25°C for periods of 2 h (B), 6 h (C) or 20 h (D). Black arrowheads indicate polypeptides which are specific to thermoinhibited achenes.



achenes were incubated at 36°C, no differences in the polypeptide patterns were observed from those found in dry achenes which had not been heat-treated (compare Figure 3.7A and Figure 3.10). Besides these thermoinhibition-associated polypeptides, the pattern of changes in polypeptides during germination after the relief of thermoinhibition (Figure 3.9B-D) was similar to that of achenes germinated under ideal conditions, although the timing of the different events did differ. These temporal differences were, however, in keeping with the differences in the growth curves of control and previously-thermoinhibited achenes.

No molecular analyses have, as far as can be ascertained, been conducted to date in which the high-temperature inhibition of seed germination, or thermoinhibition, has been studied. Previous studies involving thermoinhibition have concentrated on physiological phenomena such as the effects of temperature on the structures enclosing the embryo (CÔME and TISSAOUI, 1973; AGRAWAL, SETHI and MEHRA, 1980; DREW and BROCKLEHURST, 1984) and on endogenous levels of various phytohormones, including gibberellins (BIDDINGTON and THOMAS, 1978; SAINI, BASSI, CONSOLACION and SPENCER, 1986), cytokinins (BIDDINGTON and THOMAS, 1978), abscisic acid (YOSHIOKA, ENDO and SATOH, 1998) and ethylene (GALLARDO, DEL MAR DELGARDO, SANCHEZ-CALLE and MATILLA, 1991; GALLARDO, SANCHEZ-CALLE, MUÑOZ DE RUEDA and MATILLA, 1996). It has also been suggested that thermoinhibition may result from temperature-induced repression of metabolic reactivation from the quiescent state, possibly due to denaturation of key enzymes (PIROVANO, MORGUTTI, ESPEN and COCUCCHI, 1997). TAKEBA and MATSUBARA (1976) have suggested that a thermolabile factor is involved in thermoinhibition in *Lactuca sativa* cv. New York. Although these authors did not establish the nature of this factor, they proposed that it was probably protein based.

This study thus appears to be the first to examine thermoinhibition from a molecular viewpoint. Two-dimensional polyacrylamide gel electrophoresis has shown that a specific subset of polypeptides is differentially-expressed in thermoinhibited achenes. In the absence of any comparable studies involving thermoinhibition, literature on

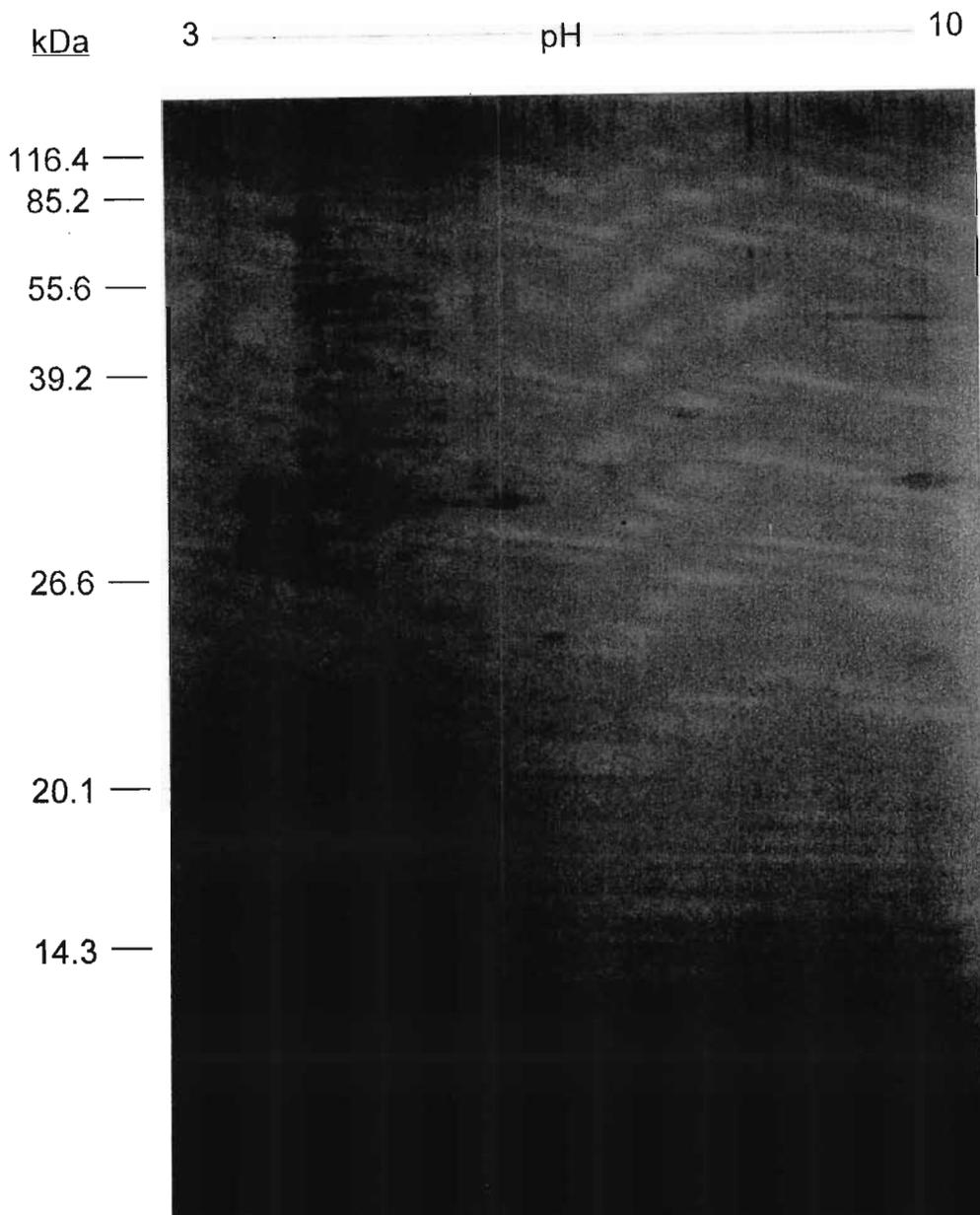


Figure 3.10: 2D-PAGE separation of proteins from achenes incubated in a dry state at 36°C for 72 h.

seed dormancy, another form of arrested germination, was examined. Many features of the pattern of expression of these polypeptides resemble very strongly the patterns of gene expression which have been observed in species where embryo dormancy is believed to be actively imposed by the expression of specific “dormancy-related” genes. Firstly, the thermoinhibition-associated polypeptides identified in *T. minuta*

achenes are only expressed when the achenes are imbibed at 36°C. Incubating dry achenes at this temperature does not lead to their induction. Although several of the dormancy-associated genes which have been identified are expressed in dry, dormant seeds, it has been proposed that seed dormancy is only imposed during early imbibition (DYER, 1993), when these proteins can exert a metabolic effect. This proposal was supported by several lines of physiological evidence which suggests that both dormant and non-dormant seeds are equally active during the first few hours of imbibition. In *Avena fatua*, dormant and non-dormant embryos show similar respiratory quotients during the first 16 h of imbibition. Respiration increased in the germinating, non-dormant embryos after this time but remained constant or declined slightly in the dormant embryos (SIMMONDS and SIMPSON, 1971). Similar results have been reported for *Agrostemma githago* (DE KLERK, 1981). Furthermore, several studies have shown that imbibition rates are similar in dormant and non-dormant seeds during phases I and II of water uptake (DYER, 1993). A similar trend has in fact been shown for water uptake in thermoinhibited and non-thermoinhibited *T. minuta* achenes (DREWES, 1989).

Secondly, the basic pattern of germination under ideal conditions is very closely mirrored in the thermoinhibited achenes, with the only major changes observed being the production of the ten thermoinhibition-associated polypeptides. The fact that only ten proteins specific to the thermoinhibited achenes were identified is consistent with the results of most dormancy studies, where the number of dormancy-associated genes identified is typically very low. DYER (1993) observed two-dormancy-associated soluble proteins and mRNA's in *Avena fatua*. LI and FOLEY (1995) used subtractive hybridization to generate 16 cDNA clones which were differentially-expressed in dormant caryopses of the same species. Only a single dormancy-associated cDNA clone was isolated from *Bromus secalinus* (GOLDMARK, CURRY, MORRIS and WALKER-SIMMONS, 1992). Two dormancy-associated proteins were observed in *Acer saccharum* embryos following 2D-PAGE (HANCE and BEVINGTON, 1992). It has been suggested that the overall similarities between the dormant and the non-dormant conditions is perhaps an indication of the potential significance of the few dormancy-specific changes observed (DYER, 1993).

Thirdly, the appearance and disappearance of these polypeptides is very closely correlated with the imposition of thermoinhibition and its release respectively. These polypeptides are not synthesized at all in achenes germinated at favourable temperatures. Their decline following a reduction in the incubation temperature also correlates very strongly with an increase in the germinability of the achenes as measured by radicle emergence, and with the production of a number of new proteins which are presumably involved in germination and post-germinative events. In several dormancy studies, the dormancy-associated genes were also expressed in non-dormant achenes, although usually at lower levels than in the dormant seeds (MORRIS, ANDERBERG, GOLDMARK and WALKER-SIMMONS, 1991; GOLDMARK, CURRY, MORRIS and WALKER-SIMMONS, 1992; DYER, 1993; LI and FOLEY, 1995). In all cases, in the non-dormant seeds these proteins and mRNAs disappeared within a couple of hours of imbibition. In all cases, however, expression of these dormancy-associated genes is discontinued as dormancy is lost, correlating with an increase in the germinability of the seed lots.

The similarities between the expression of the thermoinhibition-associated proteins observed in this study and the expression of various dormancy-associated genes observed in a variety of species suggest that thermoinhibition in *T. minuta* may also be actively imposed by the expression of a specific subset of genes which are only expressed when achenes are thermoinhibited. Since identification of these polypeptides has not yet been attempted, it is possible that some or all of these polypeptides are either heat shock proteins or LEA (late embryogenic abundant) proteins and that they are involved in protecting the achenes from heat and water stress at 36°C. Several studies thus far have in fact tentatively identified some of the dormancy-associated genes which have been isolated as LEA proteins (DYER, 1993; LI and FOLEY, 1995) such as dehydrin and the wheat Em protein (MORRIS, ANDERBERG, GOLDMARK and WALKER-SIMMONS, 1991). The extremely short-lived nature of the polypeptides isolated in this study following a reduction in the incubation temperature would, however, suggest otherwise.

If thermoinhibition is imposed through the regulation of gene expression in *T. minuta*,

it is not inconceivable that the arrest of germination in imbibing seeds observed in dormant and thermoinhibited seeds may be regulated by the same or at least similar genes. Indeed, in *Avena fatua*, a dormancy-associated cDNA clone was also found to be expressed when non-dormant caryopses were thermoinhibited by imbibition at 35°C (LI and FOLEY, 1995). A cDNA clone for a dormancy-associated gene isolated from dry embryos of *Bromus secalinus* hybridized to total RNA from dissected embryos of *Bromus tectorum*, *Aegilops cylindrica*, *Secale cereale* and *Avena fatua* (GOLDMARK, CURRY, MORRIS and WALKER-SIMMONS, 1992). Taken together, these results, along with the similarity of mRNA and protein expression observed in dormant seeds and thermoinhibited achenes of *T. minuta*, are suggestive of a common underlying mechanism of genetic expression which actively represses germination in dormant or thermoinhibited seeds. It must, however, be remembered that thermoinhibition is a distinctly different phenomenon from embryo-imposed seed dormancy. Thermoinhibited seeds are not dormant, but merely in a state of suspended germination which is quickly released when the temperature is reduced below the thermoinhibitory threshold. Unlike dormancy, no other dormancy-breaking treatment is required.

Therefore, whilst both dormancy and thermoinhibition may possibly operate through the expression of similar genes, the signal transduction pathways involved in eliciting these gene responses are probably different. It is interesting to note that where hormonal studies have been conducted in conjunction with molecular studies of dormancy, all of the dormancy-associated clones identified to date show increased expression in response to applied abscisic acid (ABA) (MORRIS, ANDERBERG, GOLDMARK and WALKER-SIMMONS, 1991; GOLDMARK, CURRY, MORRIS and WALKER-SIMMONS, 1992; LI and FOLEY, 1995) and/or a decline in expression in response to exogenously-applied gibberellic acid (GA) (JOHNSON, CRANSTON, CHAVERRA and DYER, 1995; LI and FOLEY, 1995; BAILEY, LYCETT and ROBERTS, 1996). This provides further evidence of a common signal transduction pathway leading to the expression of these genes during dormancy. No studies have been conducted on the effects of ABA on achenes of *T. minuta*, but exogenously applied GA has been shown to break thermoinhibition in this species

(DREWES and VAN STADEN, 1990).

3.4 CONCLUSION

Incubating achenes of *T. minuta* at 36°C results in the synthesis of a novel set of ten polypeptides which are highly specific to achenes treated in this fashion. These polypeptides can only be separated using 2D-PAGE, as single-dimension SDS-PAGE does not have sufficient resolution to allow their detection. These polypeptides are only synthesized when achenes are metabolically active, that is, during imbibition. They are not synthesized in dry achenes incubated at thermoinhibitory temperatures. When the incubation temperature is reduced and thermoinhibited achenes are allowed to complete the germination process, these polypeptides decline rapidly in a fashion which corresponds with an increase in the germinability of the seed lot. These results are consistent with a role for these polypeptides in the active enforcement of a period of stasis in thermoinhibited achenes.

Similar results have been observed in several plant species which exhibit embryo-imposed seed dormancy. It is therefore proposed that thermoinhibition in *T. minuta* achenes is caused by the expression of a small subset of genes which are specifically expressed at high temperatures and which may allow the respective gene products to repress key metabolic and/or regulatory events during the early stages of imbibition to prevent the completion of germination under unfavourable temperature conditions.

CHAPTER 4

NUCLEIC ACID ANALYSIS

4.1 INTRODUCTION

Differential display was developed by LIANG and PARDEE (1992) as an alternative to subtractive hybridization for identifying differentially expressed genes in mammalian cell populations and then recovering and cloning the cDNA fragments. The basic procedure involves generating cDNA from the messenger RNA (mRNA) molecules in the cells, the subsequent amplification of these fragments via the Polymerase Chain Reaction (PCR) and the separation and visualization of these fragments by polyacrylamide gel electrophoresis. An anchored oligo(dT) primer is used in the reverse transcription reaction. This anneals to the beginning of the poly(A) tails of the mRNA (LIANG, AVERBOUKH and PARDEE, 1993). Consequently, differential display can be carried out using either polyadenylated RNA or total RNA samples. Four sets of degenerate anchored primers, $(T)_{12}VN$ where V is G, A or C and N may be either G, A, T or C, are used to prime the reverse transcription reaction. The specificity of the reaction is mainly provided by the 3' base of the primers (N), which allows for considerable degeneracy at the penultimate position (LIANG, AVERBOUKH and PARDEE, 1993). This allows the subdivision of the mRNA population into 12 separate fractions, as there are 12 different possible combinations of the last two bases, omitting T at the penultimate position (LIANG and PARDEE, 1992). The cDNA subpopulations are then PCR amplified using the degenerate primer set in conjunction with a decameric oligonucleotide of arbitrary but defined sequence, as well as a radioactive nucleotide to allow for visualization of the fragments by autoradiography. The amplification products are separated on a polyacrylamide gel and the banding patterns of the two or more cell lines under study are compared to identify differentially expressed mRNAs. By changing the combinations of the primers used in the amplification reactions, most of the mRNA species present in the cells should be represented. LIANG and PARDEE (1992) suggested that the use of about 20 arbitrary decamers should result in the

identification of all possible mRNA sequences upstream of the anchored primers. BAUER, MÜLLER, REICH, RIEDEL, AHRENKIEL, WARTHOE and STRAUSS (1993) subsequently revised this figure to 25 primers.

Since the publication of the original protocol by LIANG and PARDEE in 1992, the differential display technique has been extensively used. By 1996, well over 100 papers had already been published detailing successful applications of the technique or suggesting modifications to improve the efficiency of the method (WAN, SHARP, POIRIER, WAGAMAN, CHAMBERS, PYATI, HOM, GALINDO, HUVAR, PETERSON, JACKSON and ERLANDER, 1996). The technique has, however, been criticized on a number of accounts. Firstly, a high incidence of false-positives are obtained, although WAN, SHARP, POIRIER, WAGAMAN, CHAMBERS, PYATI, HOM, GALINDO, HUVAR, PETERSON, JACKSON and ERLANDER (1996) have suggested that this rate is equal to or lower than that obtained using subtractive hybridization. Secondly, the technique appears to show a strong bias towards abundant mRNAs (BERTIOLI, SCHLICHTER, ADAMS, BURROWS, STEINBIß and ANTONIW, 1995). Thirdly, the sequences amplified by differential display frequently lie within the 3' untranslated region of the mRNA molecule, so that the usefulness of sequence data obtained directly from these sequences is limited (SOMPAYRAC, JANE, BURN, TENEN and DANNA, 1995). Several modifications to the original technique have been suggested to attempt to overcome some of these problems.

LINSKENS, FENG, ANDREWS, ENLOW, SAATI, TONKIN, FUNK and VILLEPONTEAU (1995) have suggested using longer primers and elevated annealing temperatures in the amplification step to reduce the production of false positives. These authors used 22-mer arbitrary primers and low annealing temperatures (41°C) for a few cycles of PCR. Under these conditions, the primers behave exactly as the 10-mers used in the original protocol and prime synthesis from sequences complementary to the 3' end of the primer. In subsequent cycles of PCR, the annealing temperature was increased to 60°C. This increased stringency, in conjunction with the longer primers, allows for more efficient and accurate replication of the products of the first few PCR cycles. A similar approach was taken by ZHAO,

OOI and PARDEE (1995), who suggested the use of 20-mer primers in conjunction with low-stringency (40°C annealing) amplification for a few cycles, followed by the bulk of the amplification cycles at higher stringency (60°C annealing).

SOKOLOV and PROCKOP (1994) have advocated the use of fully-degenerate 6-mer oligonucleotides to prime reverse transcription, followed by amplification of the cDNA fragments with two or three larger primers of arbitrary but defined sequence. Since the hexamers are able to prime reverse transcription from the internal sequences of the mRNA molecule, this procedure resulted in the formation of a set of cDNAs from which useful sequence data could be extracted. Since random decamers are too short to prime sequencing reactions, direct sequencing of the cDNA fragments generated by differential display is not possible. Consequently, fragments must first be cloned into a suitable vector before they can be sequenced. During the reamplification of differentially expressed fragments isolated from electrophoresis gels, WANG and FEUERSTEIN (1995) used modified random and anchored primers with additional bases at their 5' ends. These additional bases increased the length of the primers whilst maintaining the original sequences used to prime the initial round of cDNA synthesis, so allowing these modified primers to prime sequencing reactions. Since any sequence can theoretically be added to the primers, ease of cloning of these fragments can simultaneously be achieved by constructing these sequences to match restriction sites in the vector molecule.

The aim of this section of the study was to develop and optimize a method for extracting high quality RNA from thermoinhibited and germinating achenes of *Tagetes minuta* as well as optimizing the technique of differential display for use with this RNA. Once optimized, differential display would then be used to isolate differentially-expressed genes which may be involved in thermoinhibition from the thermoinhibited achenes.

4.2 MATERIALS AND METHODS

4.2.1 Isolation of RNA

Total RNA was isolated from treated achenes according to a method modified from the acid guanidinium thiocyanate-phenol-chloroform protocol of CHOMCZYNSKI and SACCHI (1987). All solutions were prepared with water which had been treated with 0.1% (v/v) diethyl pyrocarbonate (DEPC), unless otherwise stated. Glassware was baked at 250°C overnight before use. Glassware containing plastic or rubber seals which could not be baked was treated overnight with 0.1% DEPC and then autoclaved prior to use. Plasticware was used directly from unopened, sterile packs. All surfaces were cleaned with ethanol, followed by treatment with RNase AWAY (Molecular Bio-Products) before any work was done. Clean latex gloves were worn at all times when working with RNA.

Five hundred milligrams of achenes were ground to a fine powder in a pre-chilled mortar and pestle using liquid nitrogen. The powdered achenes were then rapidly transferred to a 15 ml centrifuge tube containing 7 ml denaturing solution (4 M guanidinium thiocyanate; 25 mM sodium citrate, pH 7.0; 0.5% [v/v] Sarkosyl; 0.1 M 2-mercaptoethanol) containing 5% (w/v) PVPP. To this were added, in the following order, 0.75 ml 2 M sodium acetate (pH 4.0), 7 ml water-saturated phenol (equilibrated to pH 7.8-8.2) and 1.5 ml chloroform:isoamyl alcohol mixture (49:1). The centrifuge tube was capped and the contents thoroughly mixed by inversion following the addition of each reagent. After all reagents had been added, the tube was shaken vigorously and incubated on ice for 15 min. Phases were separated by centrifugation at 12 000 *g* at 4°C for 20 min and the aqueous phase transferred to a fresh 15 ml tube. This was subjected to another phenol-chloroform extraction, using an equal volume of phenol: chloroform: isoamyl alcohol (25:24:1). After incubating on ice for a further 15 min, the tube was centrifuged and the aqueous phase removed to a fresh 15 ml tube. RNA was precipitated with an equal volume of isopropanol at -20°C for at least one hour. The precipitate was collected by centrifuging the tube as before. The pellet was completely resuspended in 3.6 ml denaturing solution and the

precipitation step repeated. After centrifugation, all of the supernatant liquid was removed by aspiration and the pellet of nucleic acid resuspended completely in 500 μ l HPLC-grade water.

An equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol was added and the tube contents mixed by gentle inversion. The tube was incubated on ice for 10 min and then centrifuged at maximum speed in a microfuge for 10 min. This step was repeated until the interphase became clear. Once the interphase had cleared, a final extraction with 24:1 chloroform:isoamyl alcohol was performed to remove any remaining traces of phenol. After centrifugation, the supernatant liquid was transferred to a fresh microfuge tube. The RNA was then precipitated by adding 500 μ l 4 M LiCl and storing the tube at -20°C overnight, whereafter the tube was centrifuged to collect the RNA. After removing the supernatant liquid by aspiration, the pellet was dissolved in 250 μ l HPLC-grade water. One tenth volume 5 M NaCl and three volumes ice-cold ethanol were added and the RNA pellet collected by centrifugation in a microfuge for 10 min. After discarding the supernatant liquid, the RNA was washed twice in ice-cold 70% ethanol and then recentrifuged. Finally, the pellet was washed with ice-cold absolute ethanol and dried under vacuum. The RNA was resuspended in 20 μ l sterile HPLC-water. RNA samples were stored at -70°C until further use.

4.2.2 Analysis of RNA

Isolated RNA was quantified by spectrophotometric analysis at wavelengths of 260 nm, 280 nm and 320 nm in clean quartz cuvettes, using an HPLC-water blank. Since nucleic acids do not absorb at 320 nm, measuring absorbance at 320 nm allows for the elimination of background and hence more accurate quantitation of RNA samples. The corrected A_{260} value was thus obtained by subtracting the absorbance at 320 nm from that obtained at 260 nm. The concentration of the sample was then

calculated using the equation:

$$[\text{RNA}] \text{ } \mu\text{g/ml} = \text{Corrected } A_{260} \times \text{dilution} \times 44.19 \text{ } \mu\text{g/ml}$$

where:

Corrected A_{260} = absorbance (in optical densities) at 260 nm,
corrected to eliminate background as outlined above,
dilution = dilution factor, and
44.19 $\mu\text{g/ml}$ = extinction coefficient of RNA (ΔE).

Since measurements giving readings of less than 0.1 OD are generally unreliable (FARRELL, 1993), care was taken to use sufficient RNA to exceed this minimum value. To estimate the purity of the samples, the A_{260}/A_{280} ratio of the sample was calculated. Pure RNA samples have an A_{260}/A_{280} ratio of 2 ± 0.05 (FARRELL, 1993). The integrity of the RNA samples was verified by examining the integrity of the 28S and 18S ribosomal RNA fragments after electrophoretic separation in a 1.5% (w/v) agarose gel under non-denaturing conditions.

4.2.3 Differential display of mRNA

Differential display was performed based on the original method of LIANG and PARDEE (1992). RNA was isolated from achenes imbibed for 72 h at 35°C (thermoinhibition treatment) and from achenes imbibed at 25°C for 24 h (control treatment). For each treatment, four sets of duplicate reactions were prepared using the degenerate (T)₁₂VN primers (T)₁₂VG, (T)₁₂VA, (T)₁₂VT and (T)₁₂VC, where V is G, A or C. These anchored primers were custom synthesized by Roche.

4.2.3.1 Reverse transcription

First-strand cDNA synthesis was conducted in 20 μl volumes. Reactions contained

1 µg total RNA, 10 µM DTT, 20 µM of each dNTP, 0.5 µg (T)₁₂VN primer, 40 U placental Ribonuclease Inhibitor (Roche), 100 U Expand Reverse Transcriptase (Roche) and 1x RT buffer, as supplied by the enzyme manufacturer. Reactions were made up to volume using HPLC-grade water and were incubated at 37°C for 1 h in a Hybaid Thermal Reactor thermocycler, after which the reaction was terminated by a 5 min incubation at 95°C.

4.2.3.2 Amplification of cDNA by the polymerase chain reaction

Amplification of cDNA samples was conducted in 20 µl reaction volumes. PCR was performed in the presence of the same (T)₁₂VN anchored primer as was used in the corresponding reverse transcription reaction, as well as a single random decamer (Operon Technologies) as an upstream primer. Ten different upstream primers were used for each of the cDNA reactions. Sequences of these decamers are given in Table 4.1.

Table 4.1: Random decamer primers used for PCR amplification of cDNA fragments generated by reverse transcription of RNA isolated from *Tagetes minuta* achenes.

| Primer | Sequence (5'-3') | M _r | % GC |
|--------|------------------|----------------|------|
| OPJ-01 | CCCGGCATAA | 2988 | 60 |
| OPJ-02 | CCCGTTGGGA | 3035 | 70 |
| OPJ-03 | TCTCCGCTTG | 2961 | 70 |
| OPJ-04 | CCGAACACGG | 3013 | 70 |
| OPJ-05 | CTCCATGGGG | 3035 | 70 |
| OPJ-06 | TCGTTCCGCA | 2970 | 60 |
| OPJ-07 | CCTCTCGACA | 2939 | 60 |
| OPJ-08 | CATACCGTGG | 3019 | 60 |
| OPJ-09 | TGAGCCTCAC | 2979 | 60 |
| OPJ-10 | AAGCCCGAGG | 3053 | 70 |

The PCR reaction mixture contained 1 μ l cDNA template, 2.5 μ M (T)₁₂VN anchored primer, 0.5 μ M random decamer, 2 μ M of each dNTP, 0.5 μ l (α -³⁵S)dATP (1000 Ci / mMol, ICN), 2.5 U *Taq* DNA-Polymerase (Roche) and 1X PCR buffer as supplied by the enzyme manufacturer. Reactions were made up to volume using HPLC-water and were overlaid with 50 μ l mineral oil before PCR to prevent evaporation. All components of the PCR reaction mixture except the random decamers and the cDNA templates were prepared as a cocktail mix which was divided between the different reactions, to ensure uniformity of the reaction mixtures. Amplification was carried out for 40 cycles at 95°C for 20 sec, 42°C for 20 sec and 72°C for 30 sec, followed by an additional extension period of 5 min at 72°C. To minimize errors due to non-specific amplification, each reaction was performed in duplicate. Control reactions were included for each of the RNA isolates in which 1 μ l of non-reverse transcribed RNA was used in reactions with each of the four (T)₁₂VN primers and a random decamer in order to ensure that any amplification products detected were not due to amplification of contaminating genomic DNA. If reactions were not electrophoresed immediately, they were stored at -20°C until required.

4.2.3.3 Labeling of molecular weight markers for use in differential display gels

The recessed 3' ends of the marker fragments of the 100 bp DNA Ladder from GibcoBRL® (Life Technologies) were labeled, as described in MANIATIS, FRITSCH and SAMBROOK (1982), using the Klenow fragment of DNA polymerase I to allow estimations to be made of the sizes of the various cDNA fragments in the differential display gels. One microgram of marker DNA was incubated with 1 U of Klenow enzyme in a 20 μ l reaction in the presence of 2 μ M each of dCTP, dGTP and dTTP, 2 μ Ci of (α -³⁵S)dATP (1000 Ci / mMol, ICN) in 1X Klenow buffer (50 mM Tris-Cl, pH 7.2; 10 mM MgSO₄; 0.1 mM DTT; 50 μ g/ml BSA (fraction V)). The reaction was allowed to proceed for 30 min at room temperature and was then stopped by the addition of 5 μ l formamide stop solution (95% v/v formamide; 20 mM EDTA; 0.05% w/v bromophenol blue; 0.05% w/v xylene cyanol FF). Labeled marker was stored at -20°C until used. Once labeled, the marker was used within two weeks of

preparation.

4.2.3.4: Electrophoresis of PCR products from differential display reactions

After amplification, samples were separated on a 6% acrylamide (19:1 acrylamide:*bis*acrylamide) / 8 M Urea sequencing gel (Appendix A). Two microlitres of loading buffer were added to 4 μ l aliquots of the amplified cDNAs and labeled molecular weight marker and the samples denatured at 75°C for 2 min immediately before loading. Gels were run at 60 W constant power at 50°C until the xylene cyanol front had moved three quarters of the way down the gel. The gels were dried on Whatman's 3MM chromatography paper without fixing and exposed using Hyperfilm- β max (Amersham) autoradiography film for 5 days. A hypodermic needle was used to punch holes through the gel and the film in three corners to allow for alignment of the gel and the film after development, so that differentially-expressed bands could be located on the gel.

The film was developed for 4 min in Rodinal black and white film developer diluted 1:25 with water with continuous agitation. The film was then transferred to a stop bath containing 3% (v/v) acetic acid and agitated for 30 sec. Once the developing reaction had been stopped, the film was fixed for 2 min in Ilford Hypam rapid fixer with continuous agitation. Finally the film was washed in fresh, running water for 20 min before drying.

4.2.4 Recovery and reamplification of differentially expressed cDNAs

Once the film had been dried, it was examined on a light cabinet and putative differentially expressed fragments were identified. Where fragments were not replicated in both lanes of the duplicate reactions, these fragments were ignored as non-reproducible. Once all of the putative differentially-expressed fragments had been marked, the film was re-aligned with the gel and the marked bands excised

using a scalpel and a number 10 blade. The gel slices were placed into carefully labeled 0.2 ml PCR tubes and 100 μ l HPLC-grade water added to each tube. The samples were incubated for 10 min at room temperature and then boiled for 10 min at 99°C in a Corbett Research PC-960G thermocycler with a heated lid to prevent evaporation. The tubes were then centrifuged for 3 min at maximum speed in a Sigma 113 desktop microfuge and the supernatant liquid removed to fresh, labeled 0.5 ml PCR tubes. Samples were stored at -70°C until further use.

For re-amplification of the cDNA fragments isolated from the gels, 4 μ l of the eluate was PCR amplified in the presence of the same primers as were originally used to generate that cDNA, using a modified version of the original PCR reaction. The radioactively-labeled dATP was left out of the reaction. In order to generate sufficient product to visualize on an agarose gel with ethidium bromide fluorescence, the dNTP concentration was increased to 20 μ M of each nucleotide. The total volume of the reaction was increased to 40 μ l, although the amount of *Taq* polymerase used remained at 2.5 U. The concentrations of all other components were the same as in the original PCR reaction. The same amplification conditions as were used for the original PCR were used (40 cycles at 95°C for 20 sec, 42°C for 20 sec and 72°C for 30 sec, followed by an additional extension period of 5 min at 72°C).

Following reamplification, 25 μ l aliquots of each reaction were electrophoresed on 1.5% TBE agarose gels (Appendix A) alongside Molecular Weight Marker XIV (Roche) to determine whether the cDNA fragment had been reamplified and to ensure that the reamplified product was the same size as the fragment in the differential display gel.

4.3 RESULTS AND DISCUSSION

4.3.1 Isolation of total RNA

Isolation of RNA was initially attempted using the hot phenol method of VERWOERD,

DEKKER and HOEKEMA (1989). The original method involves grinding samples in RNase-free microfuge tubes to reduce the potential for RNase contamination during the transfer of samples from a mortar and pestle to the centrifuge tube and also to increase the number of samples which can be handled at one time. Even using liquid nitrogen to facilitate grinding, this proved impossible as the long, thin achenes would inevitably align themselves vertically around the circumference of the tube and hence could not be ground efficiently. Mortars and pestles were thus used in the grinding procedure. These were acid washed in 0.25 M HCl and then baked at 250°C overnight before use to eliminate RNase. Achenes were then ground to a fine powder using liquid nitrogen and then transferred to pre-chilled microfuge tubes before the powder could thaw. The basic protocol was then followed as described (VERWOERD, DEKKER and HOEKEMA, 1989). This resulted in a large, brown pellet being formed at the end of the procedure which was completely insoluble. After trying to dissolve the pellet, electrophoretic analysis of the supernatant liquid using a 1.5% agarose gel under non-denaturing conditions revealed that no nucleic acids could be dissolved from it. Since the pellet appeared to contain a large amount of phenolics and pigments, presumably from the black fruit walls of the achenes, the same procedure was attempted again, with the addition of 1.5% (w/v) PVPP, a phenolic adsorbent, to the initial extraction buffer. This was done as phenolic compounds have been shown to react with both proteins and nucleic acids during the extraction from tissues rich in phenolics such as pigments (WANG and VODKIN, 1994). Whilst this reduced the colour of the pellet to a pale brown, it did not increase its solubility in any way. Increasing the volume of extraction buffer also had no effect on the solubility of the pellet.

The pellet was distinctly gel-like and appeared to consist mainly of complex polysaccharides. SHIRZADEGAN, CHRISTIE and SEEMANN (1991) modified the procedure of VERWOERD, DEKKER and HOEKEMA (1989) for use with tissues rich in polysaccharides. This protocol was thus attempted as described by the authors and with 1.5% PVPP added to the initial extraction buffer. This technique was also not successful and did not appear to produce any improvement over the original VERWOERD, DEKKER and HOEKEMA (1989) method.

In general, the highest quality RNA preparations are isolated using buffers which contain guanidinium thiocyanate or guanidinium hydrochloride, due to the extremely chaotropic nature of these chemicals (FARRELL, 1993). Early guanidinium-based methods were labour-intensive and required caesium-chloride ultracentrifugation to separate RNA from genomic DNA (FARRELL, 1993). The procedure of CHOMCZYNSKI and SACCHI (1987) obviates this requirement by incorporating an extraction buffer containing water-saturated phenol, sodium acetate and chloroform. This results in RNA partitioning into the aqueous phase and DNA and proteins into the organic phase and interphase. The first attempt at isolating RNA using this method as described (CHOMCZYNSKI and SACCHI, 1987) again resulted in the isolation of an insoluble brown pellet. After attempting to dissolve the pellet in ddH₂O, the supernatant was analyzed by agarose gel electrophoresis. A broad band which co-migrated with the 100 bp band of the molecular weight marker was observed, which presumably represented the 5S rRNA, 5.8S rRNA and tRNA fractions. No other bands or smears were visible on the gel. Adding 1.5% (w/v) PVPP to the extraction buffer resulted in the formation of a light brown pellet, whilst increasing the PVPP concentration to 5% (w/v) resulted in a white pellet being produced. Results of electrophoretic analysis were no different from those obtained using the original protocol. Increasing the volume of extraction buffer (as given in Section 4.2.1) and the addition of a second phenol:chloroform extraction immediately after the first centrifugation step reduced the size of the final pellet and although the pellet *per se* remained insoluble, electrophoresis showed that high quality RNA could be resuspended from this pellet (Figure 4.1). This protocol was employed to isolate the RNA used during the optimization of the differential display technique.

For differential display, two different imbibition treatments were selected to attempt to isolate any genes which may be differentially expressed during thermoinhibition of *T. minuta* achenes. RNA representative of gene expression in thermoinhibited achenes was isolated from achenes subjected to the standard thermoinhibitory treatment of 72 h imbibition at 36°C. As a control, RNA was isolated from achenes imbibed at 25°C for 24 h. Based on the results of 2D-PAGE, 12 h and 24 h of imbibition produced polypeptide patterns most similar to those of thermoinhibited

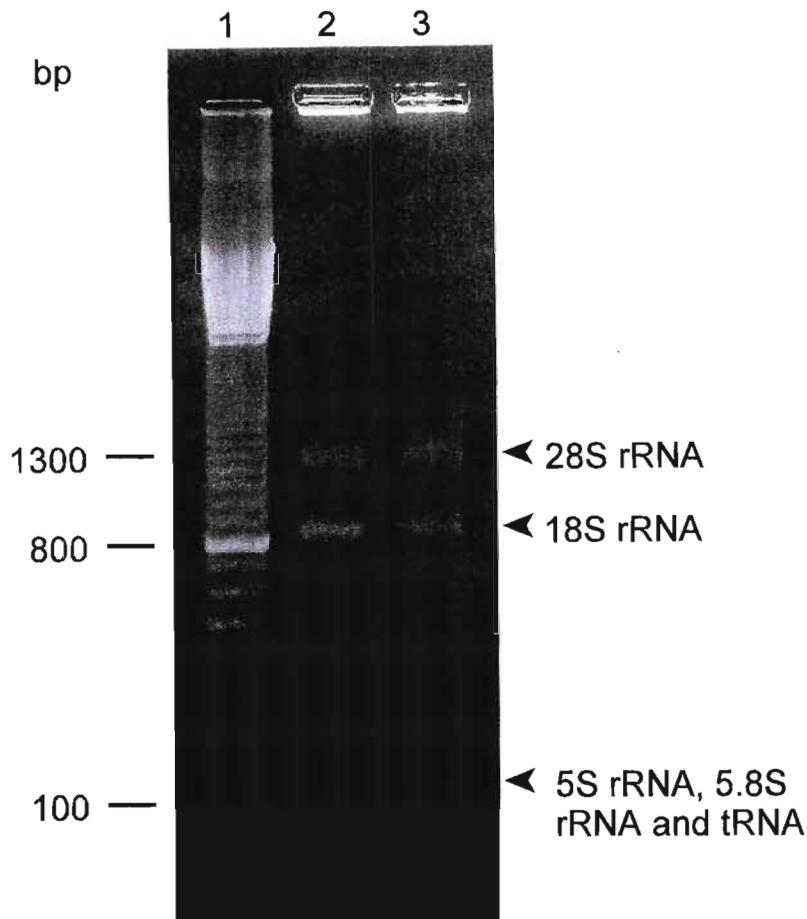


Figure 4.1: RNA isolated from control (24 h imbibition at 25°C) achenes (lane 2) and thermoinhibited (72 h at 36°C) achenes (lane 3). Lane 1 contains the molecular weight marker used in the RNA studies (100 bp Ladder, Pharmacia Biotech). The 800 bp fragment of this marker is twice as intense as the other fragments for ease of recognition.

achenes. The 24 h treatment was selected over the 12 h treatment as no visible signs of germination were observed at 12 h, whilst some radicle emergence had occurred by 24 h. All genes expressed by the germination-arrested achenes of the thermoinhibited treatment should therefore be represented in the latter treatment, whilst a few may not yet have been expressed or have been expressed at low levels in achenes of the control group, increasing the opportunity for error. During the isolation of RNA samples, it was consistently observed that approximately three times

the amount of RNA was isolated from thermoinhibited achenes than from the control group. Thermoinhibited achenes yielded an average (mean) of 65.2 ± 12.2 μg RNA per gram FW achenes, whilst 21.3 ± 11.5 μg RNA was isolated per gram FW of the control group. Further analysis of this phenomenon may be called for, as this may provide some insight into the mechanism by which thermoinhibition is implemented. These increased RNA levels in thermoinhibited achenes may indicate a build up of transcripts required for germination which are transcribed but not translated, which would suggest that thermoinhibition is regulated at the translational level. This seems unlikely, given the similar protein expression patterns of achenes subjected to these two treatments. An alternative and more likely explanation is that this increase in RNA may indicate extreme overexpression of certain genes, the gene products of which actively inhibit germination, regulating thermoinhibition at the transcriptional level.

Once the differential display technique had been optimized (Section 4.3.3), it became obvious that further purification steps were required in the RNA isolation protocol that had been used up until this point. Whilst a fairly good banding pattern was obtained in lanes where cDNA from thermoinhibited achenes was run, very poor results were observed in all of the control lanes. In the majority of cases, no amplification products were observed at all. In those lanes where bands were produced, these were much fainter than the corresponding bands in the 'thermoinhibition' lanes (Figure 4.2). This result was consistently observed in all control lanes, for all combinations of anchored and arbitrary primers. Since an identical amount of total RNA (1 μg) was used in both reverse transcription reactions, this result suggests that germinating achenes produce some factor not present in thermoinhibited achenes which is isolated along with RNA and which is inhibitory to the reverse transcriptase and/or the *Taq* polymerase enzymes used in these reactions. The RNA extraction protocol was thus modified to the procedure described in Section 4.2.1. After the first phenol:chloroform extraction, the RNA was precipitated using isopropanol, resuspended in 500 μl HPLC-grade water and then subjected to numerous rounds

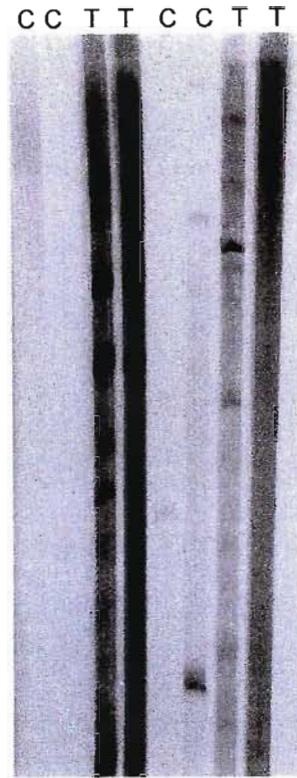


Figure 4.2: Comparison of amplified cDNA from control (24 h at 25°C) achenes (lanes 1-2, 5-6, labeled C) and thermoinhibited (72 h at 36°C) achenes (lanes 3-4, 7-8, labeled T). Note the poor amplification of the control samples.

of phenol:chloroform extractions until the interphase remained completely clear. On average, eight or nine rounds of extraction were necessary to completely remove contaminants from the RNA. A sodium chloride and ethanol precipitation step was also added at the end of the extraction protocol to help remove polysaccharides from the RNA extract. This final protocol resulted in the production of clear RNA pellets after the final round of centrifugation and ethanol washes which was easily resuspended in HPLC water. The RNA thus produced was of good quality (Figure 4.3), was easily reverse transcribed and amplified, and produced good banding patterns when run on the differential display gels.

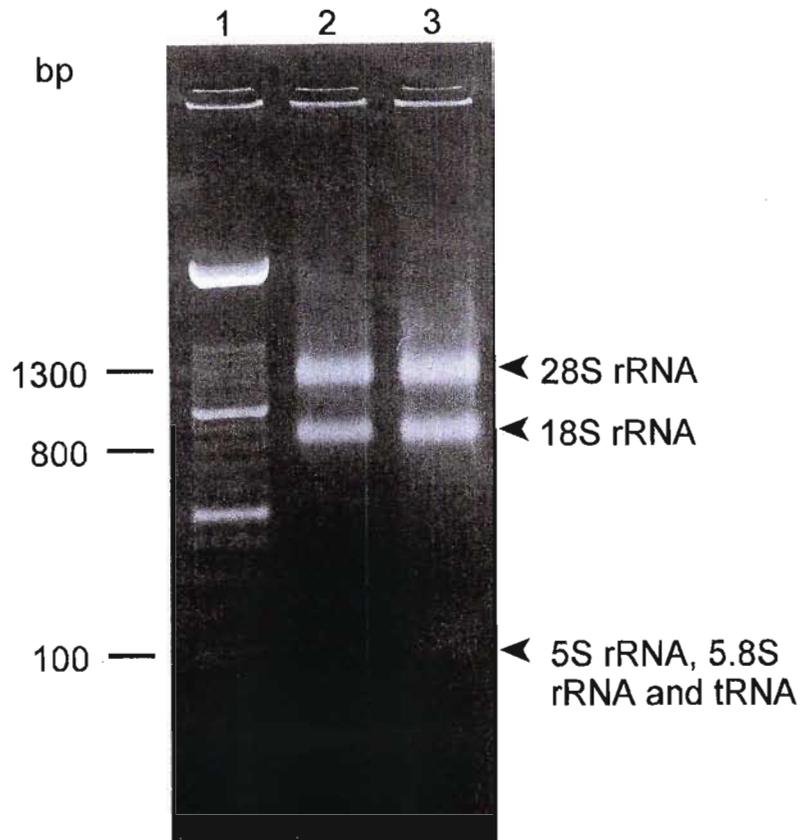


Figure 4.3: RNA isolated from control (24 h imbibition at 25°C) achenes (lane 2) and thermoinhibited (72 h at 36°C) achenes (lane 3) using the modified guanidinium thiocyanate extraction method. Lane 1 contains molecular weight marker XIV (100 bp ladder, Roche). The 500 bp and 1000 bp fragments of this marker stain more intensely than the other fragments for ease of recognition.

4.3.2 Isolation of polyadenylated RNA

Polyadenylated RNA was initially isolated from total RNA for use in differential display, as a means of increasing the efficiency of the reverse transcription reaction. This was conducted using the mRNA purification kit produced by Pharmacia Biotech. Total RNA corresponding to the amount isolated from 1 g of achenes for both of the treatments was used in the isolation, as this is the largest amount that can be used with the kit. Each total RNA sample was processed through two successive oligo(dT)-cellulose columns. Despite following manufacturer's instructions, the final

sample for both treatments contained a large amount of fine oligo(dT)-cellulose dust. When the samples were spectrophotometrically analyzed, insufficient RNA was present to accurately quantify either of the samples. Furthermore, despite several rounds of centrifugation, it proved to be impossible to remove traces of the oligo(dT)-cellulose from the sample. This isolation step was thus abandoned and only total RNA was used in the differential display reactions.

4.3.3 Optimization of differential display reactions

Initial attempts at differential display were conducted using the reverse transcription protocol outlined by KLEBER-JANKE and KRUPINSKA (1997) and with amplification conducted according to WAN, SHARP, POIRIER, WAGAMAN, CHAMBERS, PYATI, HOM, GALINDO, HUVAR, PETERSON, JACKSON and ERLANDER (1996), both of which are based on the original differential display method of LIANG and PARDEE (1992). All reagents were prepared using glass-distilled, deionized water as used in 2D-PAGE. This method failed to produce any bands on the sequencing gels. During the subsequent attempts at optimizing the reaction, the radioactive dATP was not included in the PCR reactions to reduce cost and minimize the production of radioactive waste. All further optimization reactions were electrophoretically analyzed using 1.5% agarose gels. Attempts at optimizing the reactions began with the reverse transcription reaction. An RNase inhibitor protein (Roche) was added to the reaction mixture to ensure that the problem was not due to degradation of the RNA before cDNA synthesis. Since all four of the different (T)₁₂VN anchored primers were tested, it was obvious that primer specificity was not the problem. The amount of primer in the reaction was subsequently increased from 1 μ M, as used by KLEBER-JANKE and KRUPINSKA (1997), to 0.5 μ g to eliminate the possibility that primer availability was causing the problem, although this did not appear likely. A different batch of Reverse Transcriptase enzyme was tested to ensure that the problem was not due to a fault with the enzymes. Since both the RNase inhibitor and the reverse transcriptase enzyme are DTT-dependant, the amount of DTT in the reaction mixture was increased from 10 μ M to 100 μ M, without success. As a final test, a sample of

genomic DNA from *Eucomis zambesiaca* (kindly supplied by Dr J.L.S. Taylor) was substituted for the cDNA sample in the PCR mix. No amplification products were observed following electrophoresis. At this point, it appeared that the problem did not lie in the reverse transcription of the RNA and subsequent attempts to optimize the reaction concentrated on the PCR amplification step.

Fifteen random decamer primers were tested (Table 4.2) for use in PCR amplification. Since none of these produced results, the lack of amplification was definitely not due to problems with primer recognition. In addition, the annealing temperature in the PCR cycles was reduced to 30°C to determine whether or not any problems were being experienced with primer annealing. Despite this extremely low temperature, which was below the melting temperature for all of the primers tested, no results were obtained. This result eliminated problems with primers as the reason for the observed lack of amplification. Increasing the concentration of the nucleotide mixture from 2 µM to 20 µM and 200 µM also had no effect. Finally, a different batch of *Taq* DNA polymerase was tested, again to no avail. Since virtually all parameters of both the reverse transcription and PCR reactions had been modified, the only common denominator remaining in the reactions was the ddH₂O used in preparing the reagents. When a new set of reagents were prepared using HPLC-grade water, results were immediately obtained. Although no amplification was observed in the agarose gels at a dNTP concentration of 2 µM, subsequent reactions using radioactive dATP and run on sequencing gels revealed that this concentration did in fact allow for amplification. The ddH₂O used to prepare the original reagents was prepared by running glass distilled water over an ion-exchange resin column. This can result in organic contaminants from the column being eluted into the water, and these can affect enzymic action. This problem is more pronounced with older columns which are nearing the point at which they need to be replaced. Furthermore, when the pH of this water was tested, it was found to be 5.6. This is relatively far from the pH optima of 8.3 for M-MuLV reverse transcriptase and 9.0 for *Taq* DNA polymerase. Since all of the reagents were made up using this ddH₂O, and a large amount of water was added to the reactions to make them up to the required reaction volumes, the buffering capacities of the RT and PCR buffers may have been

exceeded. The resulting unfavourable pH of the reaction mixtures may therefore have resulted in the inhibition of the enzymic reactions. Once the water had been replaced in all solutions used in the differential display reactions, differential display could be run efficiently and reproducibly. It was therefore determined that the differential display procedure had been sufficiently optimized for use in this study, and an analysis of the differences in the messenger RNA populations of thermoinhibited and germinating achenes was conducted as described in Section 4.2.3.

Table 4.2: Random decamer primers tested during optimization of PCR amplification for differential display.

| Primer | Sequence (5'-3') | M_r | T_m | % GC |
|---------------|-------------------------|----------------------|----------------------|-------------|
| OPA-01 | CAGGCCCTTC | 2995 | 34°C | 70 |
| OPA-11 | CAATCGCCGT | 2979 | 32°C | 60 |
| OPA-19 | CAAACGTCGG | 3028 | 32°C | 60 |
| OPB-09 | TGGGGGACTC | 3075 | 34°C | 70 |
| OPB-19 | ACCCCCGAAG | 2973 | 34°C | 70 |
| OPC-07 | GTCCCGACGA | 3004 | 34°C | 70 |
| OPC-11 | AAAGCTGCGG | 3068 | 32°C | 60 |
| OPC-14 | TGCGTGCTTG | 3041 | 32°C | 60 |
| OPD-03 | GTCGCCGTCA | 2995 | 34°C | 70 |
| OPD-10 | GGTCTACACC | 2979 | 32°C | 60 |
| OPE-05 | TCAGGGAGGT | 3099 | 32°C | 60 |
| OPE-20 | AACGGTGACC | 3028 | 32°C | 60 |
| OPF-02 | GAGGATCCCT | 3019 | 32°C | 60 |
| OPF-12 | ACGGTACCAG | 3028 | 32°C | 60 |
| OPF-16 | GGAGTACTGG | 3099 | 32°C | 60 |

4.3.4 Analysis of changes in gene expression in thermoinhibited achenes of *Tagetes minuta* using differential display of mRNA

Differential display was conducted using ten different random decamers and the four anchored primers for a total of 40 different primer combinations. An average of approximately 70 cDNA fragments, ranging in size from under 100 bp to over 1000 bp were generated from each primer combination (Figure 4.4).

It has been shown that the low stringency PCR conditions typically employed in the differential display technique can allow for mismatching at the priming sites, resulting in the amplification of false-positive fragments and irreproducible banding patterns (BAUER, MÜLLER, REICH, RIEDEL, AHRENKIEL, WARTHOE and STRAUSS, 1993; ZHAO, OOI and PARDEE, 1995). BAUER, MÜLLER, REICH, RIEDEL, AHRENKIEL, WARTHOE and STRAUSS (1993) reported that up to 5% of bands could not be reproduced and that for this reason, duplicate samples should be prepared for all different reactions. In all gels run during the course of this study, duplicate samples were prepared and were run in adjacent lanes. Although several non-reproducible fragments were observed in these gels, the overall number of these was very low. These bands were easily distinguished and were disregarded during the analysis of the gels.

Since differential display relies on the amplification of small amounts of single-stranded cDNA synthesized by reverse transcription of RNA, the presence of even small amounts of contaminating genomic DNA can severely affect the display of amplified fragments. It is thus necessary to include a control reaction in each set of amplification reactions where RNA is amplified without prior reverse transcription. Since no amplification products are expected to be synthesized where the sample of RNA is pure, this procedure will ensure that any DNA contamination in the RNA template will be identified. This procedure is especially important where RNA samples are not subjected to DNase digestion. Throughout the duration of this study, no amplification products were ever seen in lanes loaded with these control reactions.

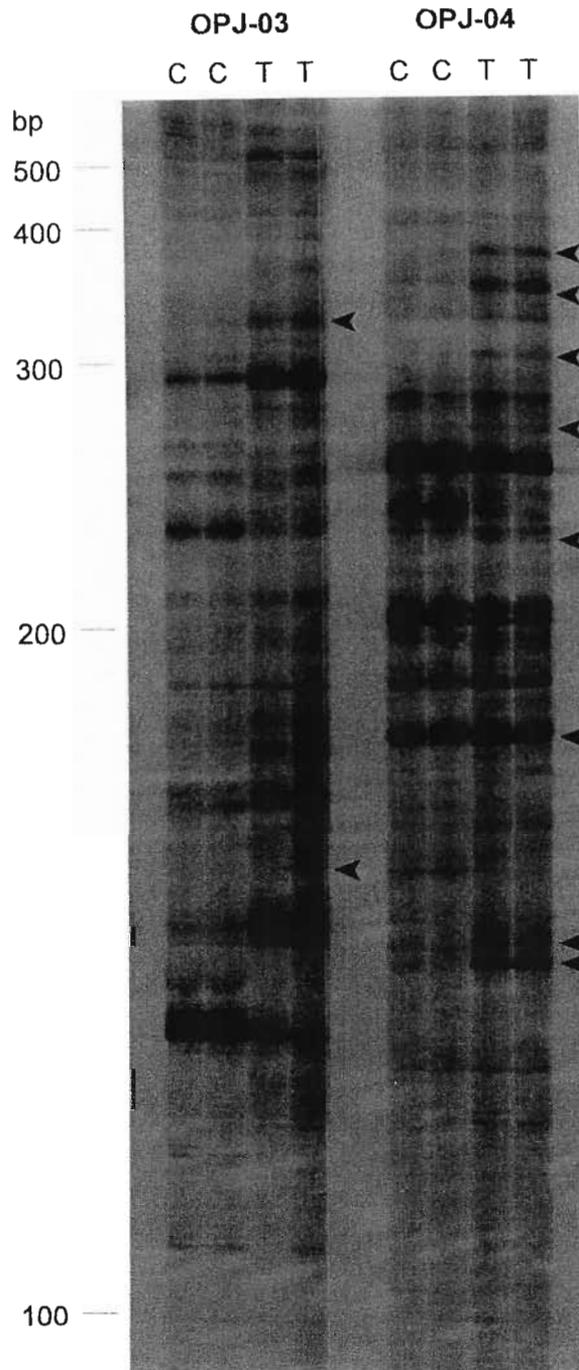


Figure 4.4: Section of a differential display gel showing duplicate control (C) and thermoinhibited (T) samples amplified from the anchored primer $T_{12}VG$ and random primers OPJ-03 (left) and OPJ-04 (right). Black arrowheads represent reproducible cDNA fragments which are specific to the thermoinhibited achenes.

A number of cDNAs, representing genes which were differentially expressed in either the germinating control group achenes of *T. minuta* (imbibed for 24 h at 25°C) or in thermoinhibited achenes (imbibed for 72 h at 36°C), were identified by the fact that they produced bands in both replicates of one treatment, but not in either replicate of the other treatment (Figure 4.4). Sixty two differentially expressed thermoinhibition-associated fragments (approximately 2% of the total number of fragments generated from RNA isolated from thermoinhibited achenes) were identified, as well as 108 cDNA fragments specific to the control achenes (approximately 4% of the total number of fragments) were identified. A summary of the number of differentially expressed cDNA fragments isolated from the control and thermoinhibited treatments using the various primer combinations are presented in Tables 4.3 and 4.4.

Table 4.3: Summary of the numbers of differentially expressed cDNAs synthesized from total RNA isolated from *Tagetes minuta* achenes subjected to the control (24 h at 25°C) treatment using various combinations of primers.

| Primer | T₁₂VA | T₁₂VC | T₁₂VT | T₁₂VG | Total |
|---------------|-------------------------|-------------------------|-------------------------|-------------------------|--------------|
| OPJ-01 | 0 | 1 | 0 | 7 | 8 |
| OPJ-02 | 2 | 0 | 1 | 6 | 9 |
| OPJ-03 | 2 | 2 | 2 | 2 | 8 |
| OPJ-04 | 0 | 2 | 3 | 1 | 6 |
| OPJ-05 | 0 | 0 | 0 | 5 | 5 |
| OPJ-06 | 15 | 1 | 10 | 0 | 26 |
| OPJ-07 | 1 | 1 | 3 | 2 | 7 |
| OPJ-08 | 2 | 0 | 6 | 9 | 17 |
| OPJ-09 | 1 | 5 | 4 | 2 | 12 |
| OPJ-10 | 0 | 0 | 6 | 4 | 10 |
| Total | 23 | 12 | 35 | 38 | 108 |

Table 4.4: Summary of the numbers of differentially expressed cDNAs synthesized from total RNA isolated from *Tagetes minuta* achenes subjected to thermoinhibitory conditions (72 h at 36°C) using various combinations of primers.

| Primer | T ₁₂ VA | T ₁₂ VC | T ₁₂ VT | T ₁₂ VG | Total |
|--------------|--------------------|--------------------|--------------------|--------------------|-----------|
| OPJ-01 | 0 | 0 | 0 | 3 | 3 |
| OPJ-02 | 1 | 0 | 1 | 0 | 2 |
| OPJ-03 | 0 | 0 | 3 | 4 | 7 |
| OPJ-04 | 2 | 6 | 4 | 4 | 16 |
| OPJ-05 | 2 | 0 | 3 | 0 | 5 |
| OPJ-06 | 4 | 0 | 2 | 0 | 6 |
| OPJ-07 | 4 | 3 | 0 | 1 | 8 |
| OPJ-08 | 0 | 0 | 2 | 2 | 5 |
| OPJ-09 | 2 | 1 | 2 | 0 | 5 |
| OPJ-10 | 0 | 3 | 0 | 3 | 6 |
| Total | 15 | 13 | 17 | 17 | 62 |

The fact that a larger number of cDNA fragments were identified as specific to the control group than to the thermoinhibited achenes is consistent with the observations made in the 2D-PAGE protein analysis, where a slightly larger number of proteins were found in the achenes imbibed under the same conditions as used for the control treatment as compared with the thermoinhibited achenes. Furthermore, after 24 h some of the control achenes had germinated and would consequently be expressing some of the genes required for seedling establishment. It would be unwise, however, to view these numbers as representative of the actual number of genes involved, as it is quite likely that several primers would anneal to different sections of the same mRNA molecule, so that several of these fragments would represent different length fragments of the same mRNA transcript.

The number of cDNA fragments generated by the various primer combinations was much more evenly spread between the various primers for the thermoinhibition-

associated fragments than for the germination-associated fragments. An exception to this rule was random decamer OPJ-04, which generated more than twice the average number of thermoinhibition-associated cDNA fragments than the other random decamers.

4.3.5 Reamplification of differentially expressed cDNAs

Once a differentially expressed cDNA was identified, it was cut from the gel using a scalpel and eluted into HPLC water (Section 4.2.4). All differentially expressed cDNAs identified, whether generated from the control or thermoinhibited achenes, were excised, eluted and stored at -70°C. Since the aim of this work was to identify those genes which may be related to thermoinhibition, only the thermoinhibition-associated cDNAs were analysed further. Whilst it may prove valuable to identify some of the germination-associated genes at a later date, this lies beyond the scope of this study. Reamplification of the eluted cDNAs was originally attempted using the same PCR conditions as used in the differential display reactions (Section 4.2.3.2), except that the radioactively labeled dATP was not included and the concentration of the dNTPs was increased from 2 μM to 20 μM , as suggested by LIANG, AVERBOUKH and PARDEE (1993). No amplification products could be seen in the lanes containing the reactions where the lower concentration of dNTPs was used. Reproducible amplification was, however, achieved using dNTPs at 20 μM (Figure 4.5). Whilst amplification products were most likely produced by the reactions containing 2 μM dNTPs, the amount of product generated was probably too small to be visualized using ethidium bromide fluorescence, which is considerably less sensitive than autoradiography. After amplification of several of the cDNA fragments using the revised reaction, it was observed that whilst many of the cDNAs could be reamplified to produce a single band of the correct size, several cDNAs could either not be reamplified or produced multiple bands (Figure 4.6). Of the latter, in all cases, one band, which was of the expected size, was much brighter than the other bands in that lane. In order to determine whether decreasing the dNTP concentration could increase the specificity of the amplification and reduce the generation of multiple

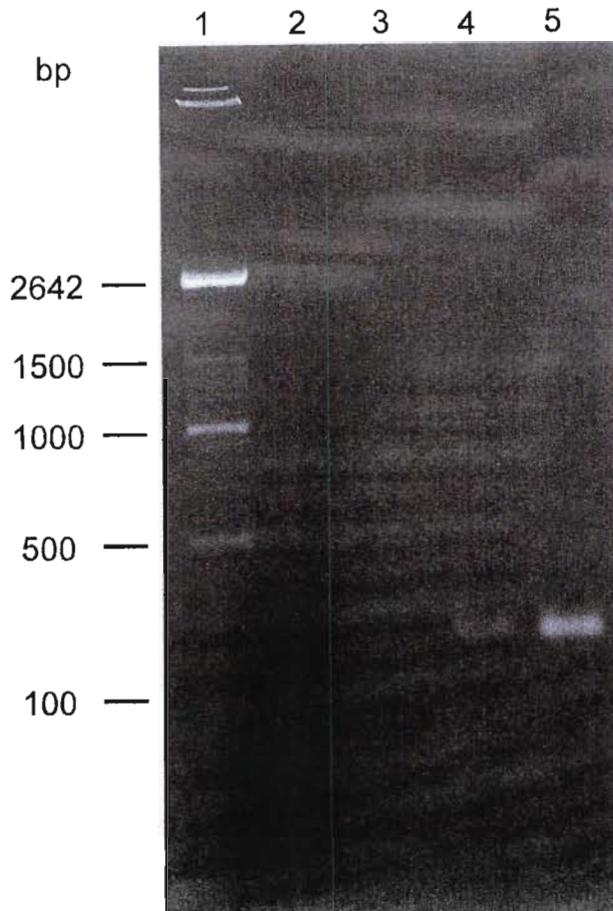


Figure 4.5: Reamplification of differentially expressed cDNAs using a dNTP concentration of either 2 μM (lanes 2-3) or 20 μM (lanes 4-5). Lane 1 contains molecular weight marker XIV (Roche).

bands, the samples were again PCR amplified in the presence of either 2 μM or 20 μM . The reaction volume was, however, increased from 20 μl to 40 μl , whilst maintaining the same concentrations of all of the reaction components except for the *Taq* DNA polymerase, to try and produce sufficient DNA with the lower dNTP concentration to visualize using ethidium bromide fluorescence. The amount of *Taq* used remained at 2.5 U in the larger reactions. Whilst more product did appear to be generated for the reactions containing 20 μM dNTPs, as evidenced by much brighter bands in the gel, the necessity of using this higher concentration was confirmed, as once again no amplification was observed in the lanes containing reactions conducted using 2 μM dNTPs. Attempting to increase the specificity of the reaction

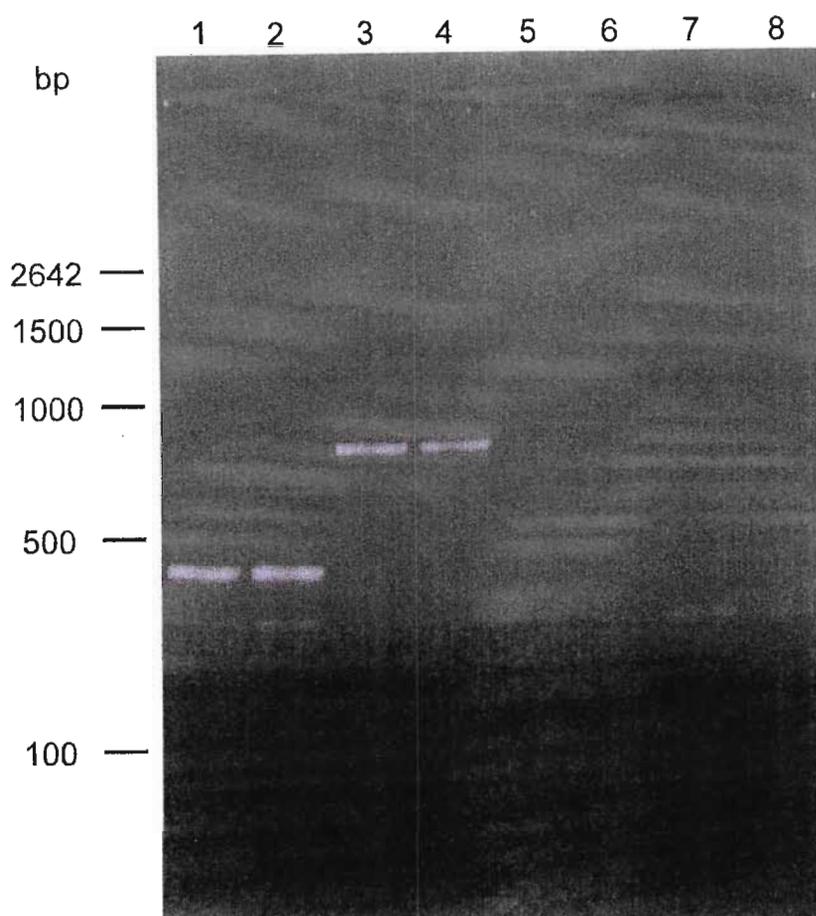


Figure 4.6: Reamplification of four differentially expressed cDNAs showing single (lanes 3-4, 7-8) or multiple (lane 1-2) band synthesis or no reamplification (lanes 5-6). In all cases where multiple bands were generated following reamplification, the brightest band corresponded in size to the original fragment on the differential display gel. As with the differential display reactions, all reamplification reactions were carried out in duplicate and were electrophoresed in adjacent lanes to ensure reproducibility.

by reducing the amount of *Taq* from 2.5 U per reaction to 1 U per reaction or to increase the amount of product generated by increasing the amount of *Taq* to 5 U per reaction proved unsuccessful. When the amount of *Taq* used per reaction was decreased, the amount of product generated was severely reduced for the one replicated reaction and in the duplicate reaction, no amplification was achieved. Increasing the amount of *Taq* also led to non-reproducible results and additional bands in one of the replicates (Figure 4.7). For all further work, therefore,

reamplification was conducted in 40 μ l reactions in the presence of 20 μ M dNTPs and 2.5 U of *Taq* DNA polymerase (Section 4.2.4).

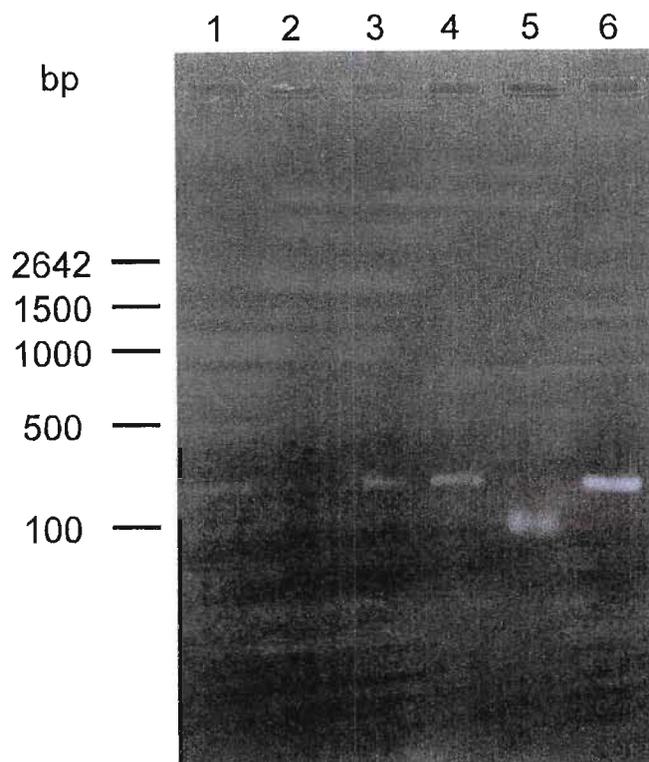


Figure 4.7: Effect of changing the amount of *Taq* DNA polymerase used in the reamplification of differentially expressed cDNAs. All reamplification was eventually conducted using 2.5 U of *Taq* polymerase per reaction (lanes 3-4). Reducing the amount of *Taq* polymerase to 1 U per reaction resulted in a significant decrease in the amount of cDNA produced (lanes 1-2), whilst increasing the amount of the polymerase in an effort to increase the amount of product resulted in the formation of non-reproducible bands (lanes 5-6).

Of the 62 thermoinhibition-associated cDNAs identified from the differential display gels, 25 could be reamplified to generate single bands of the correct size. Eleven cDNAs produced two bands following reamplification and a further 11 generated multiple bands. Fifteen cDNAs could not be reamplified at all. Due to time restrictions, only the 25 cDNA fragments which generated a single band on

reamplification were taken for further analysis. The 22 cDNAs which generated multiple bands on reamplification were not considered at this stage. In future studies, however, these bands may well be used as the correct fragment is readily identified by its size and the fact that it is much brighter than the other bands in the lane. It should be possible to separate the amplification products on either TAE or TBE agarose gels and purify the band of interest from these using the GeneClean[®] II kit (BIO 101) or similar. Such a procedure would also obviate some of the purification steps required before cloning (Chapter 5), such as the removal of unincorporated nucleotides and primers.

4.4 CONCLUSION

The isolation of intact total RNA from *T. minuta* achenes is a difficult process. The achenes contain a large quantity of phenolic compounds in the fruit walls which appear to rapidly bind to and degrade nucleic acids when the achenes are ground. After trying a number of different RNA extraction techniques, a highly modified acid guanidinium thiocyanate-phenol extraction protocol was developed which resulted in the isolation of high quality RNA from both germinating and thermoinhibited achenes. The amount of RNA isolated from control achenes was consistently about one-third of that which can be isolated from thermoinhibited achenes, which may suggest a build-up of RNA in thermoinhibited achenes. Since this may give a clue to the mechanism by which thermoinhibition is implemented, further investigation of this phenomenon may be rewarding.

Although appearing to be technically easy, differential display is recognized as being a difficult technique to establish (DEBOUCK, 1995). Initially, no amplification could be achieved, despite modifying virtually all aspects of the procedure. This was eventually solved by replacing the ddH₂O initially used in the preparation of reagents with HPLC-grade water. Optimization of the differential display procedure took a long time to accomplish, however, once this had been done, results were rapidly achieved. A number of thermoinhibition-specific cDNAs were identified, 25 of which could be

reamplified to produce single bands of a size corresponding with that of the cDNA in the original differential display gels following agarose gel electrophoresis. These fragments were selected for further analysis.

CHAPTER 5

CLONING AND ANALYSIS OF DIFFERENTIALLY EXPRESSED cDNAs

5.1 INTRODUCTION

Once a particular fragment of DNA has been isolated for study, that fragment must generally be cloned into a vector molecule and inserted into a bacterial host strain to allow for further manipulation. Whilst most species of bacteria are able to take up DNA from the medium in which they grow, this is not a very efficient process. In order for *Escherichia coli* to be efficiently transformed, the bacteria have to undergo chemical treatment to enhance their ability to take up DNA. Despite the importance of transformation as a tool in molecular biology, the process is very poorly understood. Efficient transformation of *E. coli* was first described by MANDEL and HIGA (1970). Since then there have been several refinements made to the basic technique to increase transformation efficiency. These include prolonged treatment with calcium chloride, treatment with calcium chloride in combination with other cations such as Mn^{2+} , Rb^+ and K^+ or with other compounds such as dimethylsulfoxide (DMSO), dithiothreitol (DTT) and cobalt hexamine chloride (AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL, 1989). The exact mechanism whereby these treatments prepare a cell to take up plasmid DNA is still unknown, but probably involves a structural alteration of the bacterial cell wall by the calcium ions (OLD and PRIMROSE, 1987).

In preparing the competent cells, three factors are believed to be critical (AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL, 1989):

1. Harvesting the cells in the logarithmic stage of growth,
2. maintenance of cold temperatures (0° - 5° C) throughout the procedure, and
3. prolonged exposure to calcium ions.

A subsequent brief heat-shock at 37° - 45° C has been shown to be beneficial, but is

not strictly required (OLD and PRIMROSE, 1987). Even after treatment, only 3-10% of cells are competent to take up plasmid DNA (AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL, 1989). Transformation efficiency is partially dependant on the size of the DNA fragment to be incorporated and decreases as the size of the plasmid increases (HANAHAN, 1983). Transformation efficiency is also partly affected by the availability of the plasmid. Up to a certain point, which differs according to the protocol used, an increase in the amount of plasmid will result in a linear increase in the number of transformants. Thereafter, further increases in the amount of plasmid DNA will no longer result in an increase in the number of transformed cells (AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL, 1989). In general, most protocols will result in transformation efficiencies of between 1×10^6 and 1×10^9 colony forming units (cfu) per microgram of plasmid DNA. Where higher transformation efficiencies are required, electroporation may be used instead of chemical transformation procedures (AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL, 1989).

Following transformation with most vector plasmids, separation of transformed from untransformed bacterial cells is usually easily accomplished by growing the cells on a medium containing a certain antibiotic, as the vector molecules typically contain one or more antibiotic resistance genes which allow the transformed cells to grow on the selective medium. However, usually, transformation is undertaken to introduce a recombinant vector which contains a specific DNA insert which is of interest to the experimenter into the cells. Determining which transformed colonies contain the insert is usually accomplished by a colour-based selection process which takes advantage of a modification of the bacterial lactose operon. In wild-type *E. coli* strains, the *lac* operon consists of 3 genes, *lacZ*, *lacY* and *lacA* (AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL, 1989). On rich media, and glucose-containing minimal media, the *lac* repressor protein encoded by the *lacI* gene binds to a single operator site upstream of the *lacZ* gene, preventing RNA polymerase from binding to the *lac* promoter and thus inhibiting transcription of the operon (AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL, 1989). On lactose containing media which do not contain glucose, lactose binds to

the *lac* repressor, thereby inactivating it, and a single mRNA transcript containing all three genes is transcribed from the operon (WATSON, 1977). The *lacY* gene encodes a permease enzyme which is required for the uptake of lactose by the cell, whilst *lacZ* encodes the enzyme β -galactosidase, which cleaves lactose into glucose and galactose which are then metabolised by the cell (AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL, 1989). These two genes are critical for lactose metabolism. The third gene, *lacA*, encodes thiogalactoside transacetylase, the function of which is unknown. The enzyme β -galactosidase consists of two peptides, the alpha (α) and omega (ω) peptides (AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL, 1989). When a cell contains a coding region for only one of these peptides in a truncated version of the *lacZ* gene, the cell is unable to utilize lactose. However, if both coding regions are present as separate genes within the bacterial genome, enzyme activity is observed. Such alpha-complementation was first observed by ULLMAN, JACOB and MANOD (1967).

Vectors from the pUC and M13 series and their descendants have utilized alpha-complementation to allow for rapid visual distinction between recombinant and non-recombinant colonies following transformation. These vectors contain the coding region for the α -peptide, *lacZ'* (AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL, 1989). This gene is usually modified to include a multiple cloning site, a short sequence which contains a number of restriction sites into which an insert sequence may be cloned.

Many bacterial strains used in molecular biology carry only a truncated version of the *lacZ* gene, usually on the F' episome, which encodes only the ω -peptide (AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL, 1989). Transformation of such a strain with a vector containing the *lacZ'* gene results in alpha-complementation and the production of active β -galactosidase enzymes. When plated onto media containing the lactose analogue X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactoside) and a non-metabolised inducer of the *lac* operon, such as isopropyl-thiogalactoside (IPTG), β -galactosidase breaks down the X-gal into a blue-coloured compound (BROWN, 1990). If, however, the *lacZ'* gene on the vector is

disrupted by the presence of an insert (insertional inactivation), alpha-complementation is prevented (BROWN, 1990). Colonies containing such recombinant vectors are thus unable to degrade X-gal and remain white (BROWN, 1990). Thus, recombinant and non-recombinant colonies can be distinguished by screening for β -galactosidase activity.

For more detailed analysis of the insert sequence, nucleic acid hybridisation is often used. Any two single-stranded DNA molecules have the potential to form base-pairs with one another. Generally for any two random fragments of DNA or RNA, the number of base pairings will be relatively small and the resulting hybrid molecule unstable (BROWN, 1990). However, when complementary molecules bind to each other, the number of base-pairings is increased dramatically and the molecules are able to form a stable duplex (BROWN, 1990). Since any two single stranded nucleic acid molecules may hybridise, nucleic acid hybridization may be used to identify sections of either DNA or RNA which exhibit regions of complementarity. By altering the conditions under which the hybrids are allowed to form, specifically with respect to the incubation temperature and the concentration of salts, the level of complementarity required to form stable hybrids may be adjusted. As a general rule, stringency is increased when the salt concentration is reduced and/or the incubation temperature increased (GREENLAND, 1987). By altering the stringency of the hybridisation in this way, it is possible to screen for completely homologous sequences only or to allow for hybridisation between related but not identical sections of nucleic acid. This has a wide range of applications in molecular biology, from the identification of specific clones in genomic or cDNA libraries to visualization of the expression patterns of individual genes within tissues or organs.

Once a particular section of DNA, or complementary DNA (cDNA) has been identified, this may be used as a probe to identify similar sequences in bacterial colonies or bacteriophage plaques, as well as in samples of genomic DNA or total RNA. Colony hybridization allows for rapid screening of bacterial colonies for sequences of interest by DNA:DNA or RNA:DNA hybridization. The colonies are transferred onto nylon or nitrocellulose filters and lysed (GREENLAND, 1987). The

DNA thus released is denatured and irreversibly bound to the membrane by either heat or ultra-violet irradiation (GREENLAND, 1987). The filters are then exposed to labelled probe RNA or DNA under specific stringency conditions which allow for binding of the probe to homologous sequences on the filters (GREENLAND, 1987). Once the non-hybridized probe has been washed off the filters, the appropriate detection method can be applied to detect those colonies which contain sequences with homology to the probe (GREENLAND, 1987).

Hybridization analysis can also be used to determine whether or not a particular gene is expressed under the conditions being studied. By transferring the mRNA onto a filter in a northern blot, the mRNA pool can be probed to determine whether or not a specific sequence is present. Such confirmation is particularly important following a procedure such as a differential display or subtractive hybridization, where genes which are specific to a certain set of conditions are being sought. Since differential display does generate a number of false positive sequences (WAN, SHARP, POIRIER, WAGAMAN, CHAMBERS, PYATI, HOM, GALINDO, HUVAR, PETERSON, JACKSON and ERLANDER, 1996), it is necessary to confirm that all sequences identified as being putatively differentially expressed are able to hybridize to sequences in the mRNA pool of the treated population, but not of the control population. This step is usually conducted using traditional northern hybridization analysis, which requires a large amount of RNA, as well as several hybridization membranes and various radioactively-labeled probes. This procedure thus becomes quite costly when large numbers of samples are being tested. MOU, MILLER, LI, WANG and CHALIFOUR (1994) developed an alternative procedure for the simultaneous screening of multiple samples. Known as reverse-northern analysis, this procedure uses labeled first-strand cDNA synthesized from the mRNA pool of total RNA isolated from the two different treatments as the probe. Since this cDNA should contain labeled examples of all mRNAs present in each treatment, all of the samples can be dot or slot blotted onto a single membrane and probed simultaneously. By preparing duplicate membranes and treating one with the cDNA probe prepared from the control group and the other with the probe from the treated group, differential expression of the various cDNAs can be verified.

The most valuable information which can be obtained from a fragment of DNA is the sequence of bases which make up the genetic code of that DNA fragment. Knowledge of the exact sequence of a cloned DNA fragment is vital for the planning of any substantial manipulation of that fragment. Sequencing is usually accomplished using either the MAXAM and GILBERT (1977) chemical sequencing method or the Sanger chain-termination method (SANGER, NICKLEN and COULSON, 1977). The latter method is much simpler and is the more widely used technique. The chain termination method involves enzymatic synthesis of a strand of DNA complementary to a single-stranded template. An oligonucleotide primer, typically about 17 bp in length (BANKIER and BARRELL, 1989), is annealed to the denatured sample and the second strand is synthesized by an enzyme such as the Klenow fragment of DNA polymerase I, T7 DNA polymerase (BROWN, 1990), reverse transcriptase and *Taq* DNA polymerase (BANKIER and BARRELL, 1989). The synthesis reaction includes each of the four deoxynucleotides, as well as one dideoxynucleotide (BROWN, 1990). The dideoxynucleotide is incorporated into the elongating strand in the same manner as any of the deoxynucleotides, but inhibits further elongation of the chain as it lacks the 3' hydroxyl group on the sugar moiety which is necessary for the attachment of the next nucleotide (BROWN, 1990). By careful manipulation of the amount of the dideoxynucleotide present in the mixture, the chain can be terminated at any of the positions along the DNA strand at which the dideoxynucleotide can be incorporated (BANKIER and BARRELL, 1989). Thus, a family of polynucleotides of different lengths is synthesized, each of which ends in the dideoxynucleotide which was added to the reaction mixture (BANKIER and BARRELL, 1989). By running four separate reactions in parallel, each containing a different dideoxynucleotide, four families of polynucleotides are synthesized, each of which ends in a different base (either G, A, T or C)(BROWN, 1990). In order to read the sequence, the four families of nucleotides are separated on a denaturing polyacrylamide gel which is able to separate fragments which differ in length by a single nucleotide (BROWN, 1990). Since each band in the gel generally contains an extremely small amount of DNA, one of the dideoxynucleotides used in the sequencing reactions is usually radiolabeled, so that the banding pattern of the gel may be determined following autoradiography (BANKIER and BARRELL, 1989). The

sequence can then be determined by following the banding pattern from smallest to largest on the autoradiograph and noting from which reaction that band was synthesized (BROWN, 1990). From a conventional sequencing gel, it is usually possible to determine a sequence of between 200 and 400 bases (BROWN, 1990).

This technique has also been adapted for automated sequencing. Instead of radioactive labelling, fluorescent dyes with different emission frequencies are used, usually attached to the dideoxynucleotide (CHEN, 1994). The most widely used dye-terminators contain a rhodamine dye (CHEN, 1994). These dyes work well with *Taq* polymerase in the sequencing reaction and can be used with any primer and a variety of templates, whether single- or double-stranded or PCR-generated (CHEN, 1994). The dGTP nucleotides in the reaction mixture for such sequencing reactions are usually replaced with dITP (deoxyinosine triphosphate), which reduces gel compression and allows longer sequences to be read (CHEN, 1994).

The products of all four reactions are combined and electrophoresed within a thin capillary tube gel (HEINER and HUNKAPILLER, 1989). As the fluorescently-labelled bands migrate past a detector mounted near the base of the gel, the signal spectrum is analysed to determine which dye and hence which base the bands represent (HEINER and HUNKAPILLER, 1989).

The advent of automated sequencing has greatly increased the number of bases which can potentially be sequenced each year, and has allowed for the establishment of genome sequencing projects for organisms ranging from humans to *Drosophila melanogaster* and *Arabidopsis thaliana*. Following the introduction of the first automated sequencers in 1987, the size of the GenBank genetic sequence database almost doubled in size each year for the next five years (CHEN, 1994). By 1992, this database contained over 120 million nucleotides. In conjunction with such databases, electronic hybridisation and homology searches can be used to attempt to identify a DNA fragment based on its homology to known sequences. This allows for rapid determination of whether a gene is completely novel or whether it has been

identified from other species, and greatly facilitates studies where gene expression is being analysed in relation to physiological phenomena.

The aim of this section of the study was to clone the cDNAs identified by differential display as specific to thermoinhibited achenes into the vector plasmid pGEM[®]-T Easy, to verify that these cDNA fragments were indeed thermoinhibition-specific and to determine from sequence data possible roles for these cDNAs in the imposition and maintenance of thermoinhibition in achenes of *Tagetes minuta*.

5.2 MATERIALS AND METHODS

5.2.1 Bacterial strains

Escherichia coli JM109 $\{(r_k^-, m_k^+), relA1, supE44, \lambda^-, \Delta(lac-proAB), [F', traD36, proAB, lacI^qZ\Delta M15]\}$ was obtained from Promega. Cultures were maintained on M-9 minimal medium plates (Appendix A) supplemented with 1 mM thiamine HCl to maintain the F' episome. *Escherichia coli* XL1-Blue $\{endA1, hsdR17 (r_k^-, m_k^+), supE44, thi-1, \lambda^-, recA1, (lac^-) [F', proAB, lacI^qZ\Delta M15, Tn10 (tet^r)]\}$ was obtained from Stratagene. Cultures were maintained on LB medium (Appendix A) supplemented with 12 μ g/ml tetracycline for the maintenance of the F' episome.

5.2.2 Preparation of competent *Escherichia coli* cultures

Competent *Escherichia coli* JM109 cultures were prepared according to INOUE, NOJIMA and OKAYAMA (1990). A 2 l flask containing 250 ml SOB medium (Appendix A) was inoculated with *E. coli* JM109 and the culture grown at 18°C with vigorous shaking. When the culture reached an A_{600} of 0.6, it was removed from the incubator and placed on ice for 10 min and then transferred to a pre-chilled centrifuge bottle. The culture was centrifuged at 6500 g for 10 min at 4°C and the medium poured off. The bacterial pellet was resuspended in 80 ml ice-cold TB buffer

(Appendix A), incubated on ice for 10 min and then pelleted as before. The cells were gently resuspended in 20 ml ice-cold TB buffer and dimethyl sulfoxide (DMSO) was added to a final concentration of 7% (v/v). The culture was again incubated for a further 10 min on ice before being dispensed in 1 ml aliquots into pre-chilled microfuge tubes. The cultures were then flash-frozen by immersion in liquid nitrogen prior to storage at -70°C .

Escherichia coli XL1-Blue cultures were grown to log phase at 37°C in LB medium supplemented with $12\ \mu\text{g/ml}$ tetracycline. Aliquots of 1.5 ml of this culture were transferred to sterile 1.9 ml microfuge tubes and centrifuged at 10 000 rpm in a Sigma 113 desktop microfuge for 2 min at room temperature. The supernatant liquid was decanted and the cells resuspended in $400\ \mu\text{l}$ ice-cold $0.1\ \text{M}\ \text{MgCl}_2$. The cells were centrifuged as before and resuspended in $200\ \mu\text{l}$ ice-cold $0.1\ \text{M}\ \text{CaCl}_2$. The cells were then placed on ice for at least 30 min. Following the ice-incubation step, the cultures were used immediately for transformation.

5.2.3 Preparation of PCR-amplified cDNAs for cloning

Following PCR amplification (Section 4.2.4), the cDNA fragments were purified using the PCR Cleanup Kit from Roche to remove any unincorporated nucleotide bases or primers. All buffers mentioned were as supplied by the manufacturers. The washing buffer had molecular biology grade ethanol (SeccoSolv[®], Merck,) added as per manufacturer's instructions prior to use. Two PCR reactions were always prepared for each cDNA fragment and these were combined following PCR to give a sample of $80\ \mu\text{l}$ volume. An equal volume of chloroform was added to the PCR mixture and the tube briefly vortexed to mix the phases. The tube was then centrifuged at 5000 rpm in a desktop microfuge to separate the organic and aqueous phases. The aqueous phase was transferred into a fresh, sterile microfuge tube and TE buffer added to $100\ \mu\text{l}$. The silica matrix provided in the kit was carefully resuspended until a homogenous suspension was obtained. Ten microlitres of the silica suspension was added to the sample along with $400\ \mu\text{l}$ nucleic acid binding buffer. The tube was

incubated for 10 min at room temperature, with rigorous vortexing every 2-3 min to resuspend the silica matrix. Thereafter the silica was pelleted by centrifugation at maximum speed for 30 sec and the supernatant was discarded. The DNA-containing silica matrix was resuspended by vortexing in 400 μ l nucleic acid binding buffer. The tube was re-centrifuged and the silica pellet washed with 400 μ l of the washing buffer (containing ethanol) by vortexing. The tube was again centrifuged and the washing step repeated. After centrifugation, the washing buffer was carefully removed using a pipette. The tube was inverted over a piece of absorbent paper towel on the laminar flow bench and the silica pellet allowed to dry for 15 min until completely white.

The cDNA was eluted from the silica by adding 25 μ l sterile water and resuspending the pellet by vortexing. The tube was incubated at room temperature for 10 min with vortexing every 2-3 min. The silica was again pelleted by centrifugation for 30 sec and the supernatant was transferred to a fresh, sterile microfuge tube, taking care not to transfer any of the silica. A second elution cycle was then performed as before and the second eluate added to the first to give a final volume of 50 μ l.

Following the purification of the cDNA fragments, the concentration of the samples was estimated by comparison with known amounts of DNA following agarose gel electrophoresis. Aliquots of 10 μ l of each sample were electrophoresed on 1.5% TAE agarose gels alongside 250 ng Molecular Weight Marker III (MWM III) (Roche). The gels were stained using ethidium bromide and the image captured using a Uvitec UVIDOC DOC-008.TFT version-10 Photodocumentation System. The images were analysed using the Uvitec UVIssoft Image Acquisition and Analysis software package. Microsoft Excel was used to perform regression analysis on the peak heights for each band of the marker. The 3530 bp band of MWM III was not included in the regression due to its anomalous staining, and the 5148 bp and 4973 bp bands were combined as they run concurrently in these gels. The regression equation was then used to estimate the amount of DNA present in the cDNA lanes. This value was then divided by 10 to give the concentration of the sample in μ g/ μ l. The same software package was also used to obtain an estimate of the size of each cDNA fragment.

Table 5.1: Amount of DNA present in each of the marker bands for 1 μ l (250 ng) DNA Molecular Weight Marker III (Roche) as used for estimation of concentration of purified cDNA fragments. For regression analysis, the 3530 bp band was disregarded and the 5148 and 4973 bp bands were combined for a total of 52.1 ng.

| Marker fragment (bp) | ng DNA |
|---------------------------------|---------------|
| 21226 | 109.4 |
| 5148 | 26.5 |
| 4973 | 25.6 |
| 4268 | 22.0 |
| 3530 | 18.2 |
| 2027 | 10.4 |
| 1904 | 9.8 |
| 1584 | 8.2 |
| 1375 | 7.1 |
| 947 | 4.9 |
| 831 | 4.3 |
| 564 | 2.9 |
| 125 | 0.7 |
| 48502 | 250 |

5.2.4 Ligation of cDNA fragments into the pGEM[®]-T EASY vector plasmid

The random primers used in differential display are too short for priming DNA sequencing reactions, thus cDNAs isolated using differential display must be cloned into vectors which contain binding sites for longer primers on either side of the insertion site which can then be used to sequence the insert DNA. The differentially expressed cDNA fragments were ligated into the pGEM[®]-T EASY vector (Promega), which is specifically designed for cloning fragments generated by PCR. The vector is cut at the *EcoRV* site of the multiple cloning site of the vector and a 3' terminal

thymidine is added to both ends. This improves ligation by preventing recircularization of the plasmid and by providing compatible overhangs for products generated by *Taq* polymerase which tends to add a single deoxyadenosine to the 3' ends of amplified products (MEZEI and STORTS, 1994). This vector contains binding sites for the T7 RNA Polymerase promoter, the SP6 RNA Polymerase promoter and the pUC/M13 forward and reverse sequencing primers which may be used to synthesize ssDNA for sequencing. The multiple cloning site of the vector lies within the *lacZ'* gene, allowing for blue/white colour screening for recombinant colonies.

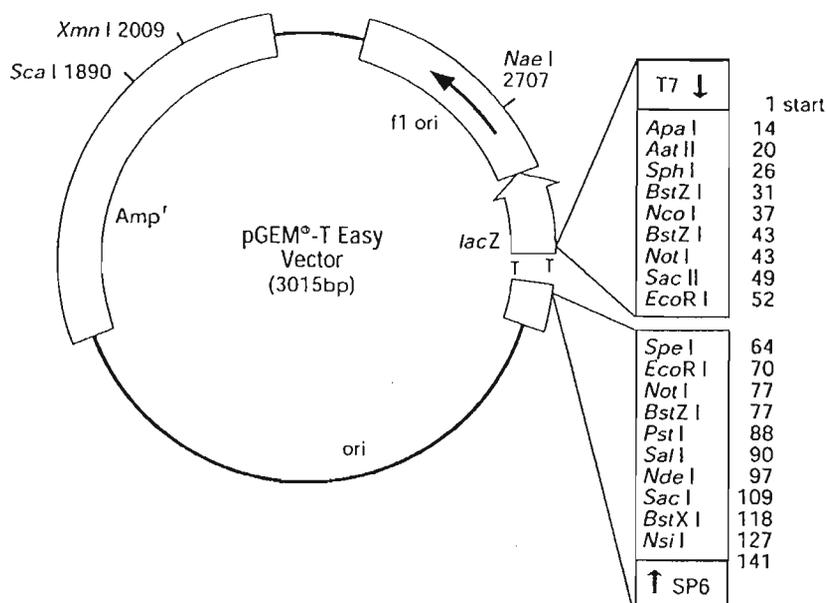


Figure 5.1: pGem®-T EASY vector map and sequence reference points (Promega catalogue).

The insert:vector molar ratios were calculated according to the equation:

$$\frac{\text{ng vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \frac{\text{insert}}{\text{vector}} \text{ molar ratio desired} = \text{ng insert DNA}$$

The vector, which is 3015 bp in size, is optimized for a 1:1 insert:vector ratio and is supplied at a concentration of 50 ng/μl, therefore the equation was used in the

format:

$$\frac{50 \text{ ng vector} \times \text{kb size of insert}}{3.0 \text{ kb vector}} \times \frac{1}{1} = \text{ng insert DNA}$$

Ligation was carried out as per manufacturer's instructions. Reactions were set up as in Table 5.2. These were then mixed by pipetting and incubated overnight at 4°C.

Table 5.2: Ligation reactions for cloning cDNA fragments into the pGEM®-T EASY vector.

| Reaction components | Standard reaction | Positive control | Background control |
|-------------------------------------|-------------------|------------------|--------------------|
| 2X Rapid Ligation buffer | 5 µl | 5 µl | 5 µl |
| pGEM®-T Easy vector (50 ng) | 1 µl | 1 µl | 1 µl |
| PCR product | X µl * | - | - |
| Control insert DNA | - | 2 µl | - |
| T4 DNA Ligase | 1 µl | 1 µl | 1 µl |
| HPLC-grade water to final volume of | 10 µl | 10 µl | 10 µl |

* As calculated using the equations given above.

5.2.5 Transformation of *Escherichia coli* with the pGEM®-T EASY vector

Tubes of competent *Escherichia coli* strains JM109 or XL1-Blue were removed from storage at -70°C and allowed to thaw slowly on ice. Once thawed, the cells were carefully resuspended by gently flicking the tubes. For each ligation reaction which was to be used in the transformations, 2 µl of the reaction was placed in a sterile 1.5 ml microfuge tube on ice. Once the tubes containing the ligation reaction mixture had chilled, 50 µl of competent cells were carefully pipetted into each tube. The tubes were again mixed by gentle flicking and incubated on ice for 20 min. The cells were then heat-shocked at exactly 42°C for 45 sec and then immediately placed on ice for 2 min. After the ice incubation, 950 µl room temperature SOC medium (Appendix A)

was added to the cells and the tubes were incubated at 37°C for 1.5 h with gentle shaking (approximately 150 rpm).

LB/ampicillin/IPTG/X-gal plates were prepared by supplementing LB-agar medium with 100 µg/ml ampicillin. When transforming *E. coli* XL-1 Blue, these plates were further supplemented with 12 µg/ml tetracycline. Once the plates were set, 100 µl of 100 mM IPTG and 20 µl of 50 mg/ml X-gal were spread over the surface and allowed to absorb for at least 1 h at room temperature before use. One hundred microlitres of each transformation culture was then plated out onto duplicate plates, which were then incubated overnight at 37°C. Plates were then stored overnight at 4°C to increase the intensity of the blue colour in the non-recombinant colonies.

5.2.6 Colony hybridisation

Colony hybridisation was performed using a method adapted from AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL (1989) and GREENLAND (1987). BIOTRACE NT 45 µm nitrocellulose filters (Gelman Sciences) were used for all hybridisations. Fifty white colonies were selected from the duplicate transformation plates and grid plated onto LB/ampicillin plates. When using *E. coli* XL-1 Blue, these plates were also supplemented with tetracycline. These plates were incubated at 37°C overnight or until colonies with a diameter of approximately 1 mm had grown. A sterilised nitrocellulose filter was then carefully laid over the plate under sterile conditions on the laminar flow bench. When the filter had become completely moist, it was peeled off the surface of the plate and placed with the bacterial side up on a fresh plate.

5.2.6.1 Preparation of replica filters

Nitrocellulose filters were moistened on a pad of Whatman® Number 1 filter paper discs which had been wetted with sterile LB medium. The original nitrocellulose

filters were removed from the agar plate and placed with the bacterial side up on a pad of sterile Whatman® 3MM chromatography paper in the laminar flow bench. The replica filters were then placed over the original filters, offset by 2 - 3 mm to allow for easier separation of the filters. Care was taken to avoid trapping air bubbles between the two filters. A pad of Whatman® 3MM paper was then placed on top of the filters, followed by a glass plate. Pressure was then applied to the plate to transfer the colonies. The plate and top layer of 3MM paper was then removed. Orientation holes were punched through the paired filters using a 21G hypodermic needle. The filters were then carefully peeled apart and placed onto LB agar plates containing the relevant antibiotics. Plates containing replica filters were incubated overnight at 37°C, whilst the original filters were incubated at room temperature to allow the colonies to re-grow. The original plates were then maintained at 4°C until needed. Once the colonies on the replica filters had re-grown, the filters were transferred to LB agar plates supplemented with 100 µg/ml ampicillin and 50 µg/ml chloramphenicol for plasmid amplification. These plates were then incubated at 37°C for a further 4 - 10 h.

5.2.6.2 Preparation of filters for hybridisation

Filters were removed from the chloramphenicol-supplemented plates and placed with the bacterial side up on a sheet of Whatman® 3MM paper soaked in 0.5 M NaOH for 5 min. They were then carefully transferred to another sheet of 3MM paper soaked in 1M Tris-Cl (pH 7.5) and allowed to neutralize for a further 5 min. Finally the filters were transferred to a third sheet of 3MM paper soaked in 0.5 M Tris-Cl (pH 7.5) / 1.25 M NaCl for 5 min. Filters were then placed onto dry chromatography paper and allowed to dry completely at room temperature before being baked at 80°C for 90 min to fix the plasmid DNA to the filter.

5.2.6.3 Preparation of radioactive probe DNA

In order to prepare probe DNA, the cleaned cDNAs used in the ligation reactions were re-amplified by PCR. The same basic PCR protocol as was used to re-amplify the cDNA fragments (Section 4.2.4) was followed, with slight modifications. Twenty nanograms of cDNA was amplified per reaction. Radioactive ^{35}S -dATP (0.5 μl per 40 μl reaction, 1000 Ci / mMol, ICN)) was added to label the probe DNA. The reaction was then performed in the presence of a mixture of 20 μM each of dGTP, dTTC and dCTP, so that all dATP was supplied in the form of ^{35}S -dATP.

Following PCR, unincorporated nucleotides were removed by column chromatography using Roche Sephadex G50 mini Quick Spin columns. Column chromatography was conducted according to the manufacturer's instructions. The eluate from the Sephadex G50 columns was then used directly as the probe.

5.2.6.4 Hybridisation of filters with radioactive probe DNA

The baked filters were pre-hybridised in 20 ml pre-warmed hybridisation buffer (5X Denhardt's solution in 6X SSC (Appendix A) at 65°C for 3-4 h in a HYBAID mini-hybridisation oven. The pre-hybridisation solution was then removed and replaced with 8 ml fresh, pre-warmed hybridisation buffer. The radiolabeled probe DNA samples were denatured for 5 - 10 min in a waterbath at 100°C and then added to the hybridisation solution. Hybridisation was then allowed to proceed overnight. The hybridisation solution was poured out of the hybridisation bottle and the filters washed in 70 ml pre-warmed 3X SSC / 0.1% SDS for 20 min at 65°C. This wash was repeated three times. The filters were then removed from the hybridisation bottles and dried on a sheet of filter paper at 37°C.

5.2.6.5 Detection of the hybridisation signal and recombinant colonies

Once dry, the filters were mounted onto a sheet of 3MM paper. Three marks were made on the paper with radioactive ink (10 µl India ink with 0.5 µl [α -³⁵S]dATP, [1000 Ci / mMol, ICN]) for later alignment of the autoradiograph. The paper was then covered with clingfilm and exposed to autoradiography film (Hyperfilm- β max, Amersham) for 7 days in an Amersham Hypercassette. After the autoradiograph had been developed (as per Section 4.2.3.4), three clones which produced the strongest signals were selected from each plate and streaked onto LB plates supplemented with the relevant antibiotics.

5.2.7 Isolation of plasmid DNA

Plasmid DNA was isolated from recombinant colonies using the alkaline lysis method of FELLICIELLO and CHINALI (1993). Single colonies from the various culture plates were used to inoculate 10 ml aliquots of LB medium supplemented with the relevant antibiotics. Following overnight growth in an oscillating waterbath at 37°C, the cultures were decanted into sterile 15 ml centrifuge tubes and centrifuged at 1000 *g* for 5 min. The supernatant liquid was then discarded and the bacterial pellets resuspended in 1 ml ice-cold STE buffer (Appendix A) before being transferred into 1.9 ml microfuge tubes. The cell suspensions were next centrifuged at maximum speed in a desktop microfuge at room temperature for 1 min. The supernatant liquid was completely removed by aspiration and the pellets completely resuspended in 250 µl Solution I (Appendix A) by repeated and vigorous vortexing. Five hundred microlitres of Solution II (Appendix A) were added and the tube contents mixed by gently inverting them 3-4 times. The tubes were then held on ice for 3-5 min. Cell debris and genomic DNA was then precipitated out by adding 750 µl ice-cold 4 M potassium-acetate/2 M acetic acid and immediately and vigorously shaking the tubes by hand to mix the contents. The tubes were again incubated on ice for 5-10 min before being centrifuged at maximum speed for 5 min. The supernatant liquid was transferred to fresh microfuge tubes. The pipette tip was placed well into the

supernatant liquid before suction was applied to avoid collecting the floating film of denatured material. All subsequent operations were conducted at room temperature. Seven hundred microlitres of isopropanol were added to the tubes and the contents mixed well by inversion. The tubes were recentrifuged as before and the supernatant removed completely. The plasmid DNA pellet was resuspended carefully in 250 μ l TE buffer containing 10 μ g/ml RNase A (Roche, see Appendix A for preparation). Digestion of the RNA was allowed to proceed at room temperature for 15 min. Thereafter an equal volume of 25:24:1 phenol:chloroform:iso-amyl alcohol was added and the tube contents mixed by inversion. After 3 min of centrifugation at maximum speed, the aqueous phases were removed to fresh microfuge tubes. This phenol:chloroform extraction step was repeated until the interphase between the aqueous and organic phases was clear of any denatured material. The DNA solution was then extracted with an equal volume of 24:1 chloroform:iso-amyl alcohol to remove any remaining traces of phenol, the tubes were centrifuged as before and the aqueous phases transferred to fresh, sterile microfuge tubes. The plasmid DNA was then precipitated by adding one tenth volume 5 M NaCl and 3 volumes ice-cold ethanol to the DNA solution and incubating overnight at -20°C . The plasmid DNA was collected by centrifugation at maximum speed for 10 min and the supernatant liquid was discarded. The DNA pellets were washed twice with 200 μ l ice-cold 70% ethanol, with centrifugation for 5 min after each wash. Finally the pellets were washed with 200 μ l ice-cold 100% ethanol and centrifuged. The supernatant was poured off and the tubes replaced in the centrifuge in the same orientation. The tubes were then spun briefly to collect the residual supernatant which was carefully removed using a pipette tip. The plasmid DNA pellets were dried *in vacuo* and the DNA resuspended in 20 μ l HPLC-grade water. Samples were stored at -20°C until use. Samples which were to be sequenced were quantified prior to freezer storage and were subjected to as few freeze-thaw cycles as possible to reduce damage to the DNA.

The samples were quantified as outlined in Section 4.2.2, with the equation modified

for use with DNA samples as follows:

$$[\text{DNA}] \text{ } \mu\text{g/ml} = \text{Corrected } A_{260} \times \text{dilution} \times 50.00 \text{ } \mu\text{g/ml}$$

where:

Corrected A_{260} = absorbance (in optical densities) at 260 nm,
corrected to eliminate background,

dilution = dilution factor, and

50.00 $\mu\text{g/ml}$ = extinction coefficient of DNA (ΔE).

5.2.8 Restriction analysis

Following the isolation of the putatively recombinant plasmids, these were subjected to restriction analysis to show that they did indeed contain an insert. All samples were digested with the enzyme *SacI*, which has only a single restriction site within the pGEM[®]-T Easy vector. For all reactions, 10 μg of plasmid DNA was digested using 10 U of enzyme in a total reaction volume of 20 μl . Reactions were allowed to proceed for 1 h at 37°C. For comparative purposes, a non-recombinant plasmid isolated from the background control ligation (Section 5.2.4) was also restricted.

Ten microlitres (5 μg) of each sample were then electrophoresed on 1.5% agarose gels alongside the non-recombinant plasmid and DNA molecular weight marker III for comparative size analysis. Following ethidium bromide staining (0.5 $\mu\text{g/ml}$ ethidium bromide in the electrophoresis buffer), the gels were photographed using the Uvitec UVIDOC DOC-008.TFT version-10 Photodocumentation system and the size of the various plasmids calculated using the UVItect UVIsoft Image Acquisition and Analysis software package.

5.2.9 Preparation of freezer stocks

Once colony blotting and restriction analysis had confirmed the presence of a

recombinant plasmid in each of the putative clones, these clones were grown overnight at 37°C in 10 ml LB medium supplemented with the relevant antibiotics to prepare freezer stocks of these clones. Nine hundred microlitres of this bacterial culture were added to 900 µl glycerol freezer store solution (Appendix A) in a sterile cryopreservation vial. The stock was carefully mixed by pipetting with a wide bore pipette tip and then quick frozen by immersion in liquid nitrogen. The vials were then stored at -70°C.

5.2.10 Reverse-northern hybridisation analysis

For the final verification step to ensure that the inserts contained in the vector plasmids are expressed only in thermoinhibited achenes, the reverse-northern blot procedure developed by MOU, MILLER, LI, WANG and CHALIFOUR (1994) was modified for use in this system.

5.2.10.1 Dot-blotting of recombinant plasmids

Duplicate Zeta-probe GT membranes (BIO-RAD) were cut to size and wetted for 5 min in 0.4 M Tris-HCl (pH 7.5). Whatman's 3MM chromatography paper was cut to the size of the manifold and wetted with the same solution. This was placed onto the base of the dot-blotting manifold and the membranes placed on top of it. The top of the manifold was then carefully placed over the membranes and clamped into position.

Ten micrograms of each plasmid to be blotted was placed into a sterile microfuge tube. To this was added 100 µl denaturing solution (0.25 N NaOH, 1 M NaCl). The plasmid DNA was then allowed to denature for 10 min at room temperature. Fifty microlitres of the denatured plasmid solution was loaded into each of the corresponding wells over the duplicate membranes. Once all wells had been loaded, the samples were left for 30 min before vacuum was applied. Once all samples had drained through, the membranes were carefully removed from the dot-blotting

manifold and neutralized by placing them onto filter paper soaked with 0.5 M NaCl/0.5 M Tris-HCl (pH 7.5), taking care not to get any liquid onto the top surface of the membrane to prevent cross-contamination of the DNA. After 5 min, the membranes were air dried and cross linked using the C-L program on a BIO-RAD GS Gene Linker.

5.2.10.2 Preparation of probes for reverse-northern analysis

First strand cDNA was prepared from total RNA isolated from thermoinhibited and control achenes using a modification of the reverse transcription assay previously described (Section 4.2.3.1). For the preparation of probe cDNA, the amount of total RNA used was increased to 5 µg. The RNase inhibitor was not included in the reaction and the (T)₁₂VN primer was replaced with 0.5 µg oligo d(T) primer. All other reagents were as described previously. The reaction was incubated at 37°C for 60 min. The reverse transcriptase was then denatured by heating the reactions to 95°C for 5 min, whereafter 2.5 µg of pdN₆ random hexamer mixture was added to both assays. The reactions were heated at 90°C for 10 min and then held on ice for 5 min.

A 25 µl cocktail mixture for random priming was prepared by mixing 1.25 µl of a 2 mM 3dNTP mix (dGTP, dTTP and dCTP), 7 µl 10X Klenow buffer (Appendix A), 10 µl (α-³⁵S)dATP (1000 Ci / mMol, ICN), 2 µl (2 U) Klenow enzyme and 4.75 µl HPLC-grade water. Half of this was added to each of the RT assays for a final volume of 35 µl each. The reactions were allowed to proceed for 4 h at 37°C. Reactions were terminated by adding 1 µl 0.5 M EDTA (pH 8.0). Each reaction was made up to 50 µl with TE buffer and unincorporated nucleotides were removed by column chromatography using Roche Sephadex G50 mini Quick Spin columns. The eluate from the spin columns was then used directly as the probe.

5.2.10.3 Hybridisation

The membranes were pre-washed in hybridization bottles in a HYBAID mini hybridisation oven for 1.5 h in 25 ml 0.1X SSPE/1% SDS (Appendix A) at 65°C. The pre-wash solution was then poured off and the membranes were pre-hybridised in 25 ml 0.25 M Na₂PO₄ (pH7.2)/7% SDS/10% PEG 6000 for 30 min at 65°C. For hybridization, the pre-hybridization solution was replaced by 8 ml of the same solution. The probe cDNA, which was denatured by boiling at 100°C for 5 min, was added and hybridisation allowed to proceed overnight at 65°C. Following hybridization, the membranes were washed in 70 ml 2X SSC/0.1% SDS for 30 min at 65°. The membranes were then washed twice in 2X SSC/0.5% SDS for 30 min at a time at 65°C. Finally, the membranes were dried at 37°C mounted onto 3MM paper and exposed to autoradiography film (Hyperfilm-βmax, Amersham) for 14 days in an Amersham Hypercassette.

5.2.11 Sequencing of cDNA clones

Automated sequencing of the cDNA clones was performed at the Agricultural Biotechnology Institute in Gödöllő in Hungary. Forward sequences were generated from the pUC/M13 forward primer and reverse strand sequences from the pUC/M13 reverse primer sites on the pGEM T-Easy vector. Sequencing files were analysed using Chromas version 1.45, available as freeware at <http://www.technelysium.com.au/chromas.html> and the sequences exported as text files which were edited manually. Once the insert sequences had been identified, the reverse-complement sequence of the reverse strand sequences was computed using tools available online at "The Sequence Manipulation Suite" at <http://bioinformatics.org/sms/>. This reverse complement sequence was then aligned with the relevant forward sequence using the ClustalX version 1.81 freeware program (THOMPSON, GIBSON, PLEWNIAK, JEANMOUGIN and HIGGINS, 1997) downloaded from <http://www-igbmc.u-strasbg.fr/BioInfo/ClustalX/Top.html> to determine the consensus sequence for each cDNA. The consensus sequence was then used in a computer

search of the non-redundant (nr) database using the BLAST programs BLASTN (nucleotide-nucleotide search) and BLASTX (translated nucleotide-peptide search) (ALTSCHUL, GISH, MILLER, MYERS and LIPMAN, 1990; ALTSCHUL, MADDEN, SCHÄFFER, ZHANG, ZHANG, MILLER and LIPMAN, 1997) at the National Centre for Biotechnology Information website <http://www.ncbi.nlm.nih.gov/>.

5.3 RESULTS AND DISCUSSION

5.3.1 Transformation of competent cultures with the pGEM[®]-T EASY vector plasmid

The pGEM[®]-T EASY vector system was designed for use with competent *Escherichia coli* JM109 cells. This strain was initially used for transformation for this reason. For the preparation of competent cells, the procedure of INOUE, NOJIMA and OKAYAMA (1990) was followed. This protocol results in the production of highly competent cultures and was originally designed for use with a number of *E. coli* strains, including *E. coli* JM109. This strain presents some difficulties when working with blue-white colour screening for recombinants, however. The modified *lacZ* gene, which is missing the *lacZ'* region encoding the α -peptide of the β -galactosidase enzyme, resides on the F'-episome in this strain. The F' episome is maintained through selective pressure by culturing the strain on minimal media supplemented with thiamine-HCl (AUSUBEL, BRENT, KINGSTON, MOORE, SEIDMAN, SMITH and STRUHL, 1989). During the preparation of competent cells, however, the bacteria must be grown on a nutrient rich medium (SOB) at low temperatures, which results in slow growth of the culture. During this period, some of the cells may lose the F' episome, which results in them producing false-positive white colonies on the selection plates following transformation. Rich media are also used during transformation and the selection procedures, which further compounds this problem. As a result, transformation of this strain may lead to problems with the identification of recombinant colonies.

During transformation, a background control reaction was set up to determine the

number of background blue colonies which result from non-digested plasmids and non-T-Tailed vectors which recircularize during the ligation reaction. When using *E. coli* JM109, 12.9% of these background colonies were white, presumably due to the loss of the F' episome from these cells. Following the preparation of the competent cultures and the first transformation experiment, the -70° ultrafreezer in which the competent strains were stored defrosted due to a loss of coolant gas from a leak, destroying these cultures. Since the INOUE, NOJIMA and OKAYAMA (1990) protocol is a fairly lengthy one, taking several days to prepare the cells, as well as the fact that the ultrafreezer remained out of commission for an extended period whilst it was repaired, a different transformation protocol was adopted in which a small volume of cells were made competent as required by treatment with magnesium and calcium salts and then transformed immediately. At the same time, the *E. coli* strain was changed from *E. coli* JM109 to *E. coli* XL1 Blue. This strain has the advantage that the F' episome carrying the modified *lacZ* gene also carries the Tn10 gene encoding tetracycline resistance, allowing for maintenance of the F' episome by culturing the cells on media containing 12 µg/ml tetracycline (SAMBROOK, FRITSCH and MANIATIS, 1989). Following transformation, only 1.7% of the colonies on the background control plates were white. This strain thus proved much better for use with the pGEM®-T EASY vector for this study.

Both methods of preparing competent cells resulted in cultures with a high transformation efficiency. Positive control ligations and transformations were included in the transformation procedure, where ligation was performed using a control DNA insert provided by the manufacturer. Following transformation with this control reaction, competent cultures with a transformation efficiency of 1×10^8 cfu/µg (colony forming units per microgram) DNA should generate approximately 100 colonies on the selection plates, 10-40% of which may be blue. A larger number of blue colonies indicates that the ligation conditions are suboptimal. Competent *E. coli* JM109 cultures produced over 3500 colonies following transformation, of which 8.2% were blue, whilst competent *E. coli* XL1 Blue cultures generated approximately 550 colonies, 3.7% of which were blue. These results thus indicate a high transformation efficiency for both cultures, as well as excellent ligation conditions. The exceptionally high number of colonies obtained when using *E. coli* JM109 also results partially from

the fact that when this strain was used in the first transformation experiment, an optional centrifugation step in the transformation protocol, which increases the concentration of cells in the culture which is plated onto the selection medium, was used. This step proved unnecessary due to the high transformation efficiency of the various competent cell cultures and was not used in subsequent transformations.

Negative controls were also run for all transformation experiments, where competent cells were subjected to all treatments performed on the other cells except that no plasmid DNA was added. None of the negative control cultures produced any colonies when plated onto the selective media.

The pGEM[®]-T EASY vector system is optimised for a 1:1 molar ratio of insert to vector DNA. Following reamplification and preparation of the differentially expressed cDNAs for cloning, 13 cDNAs were sufficiently concentrated to be used in this optimal ratio. These were therefore selected for cloning. Details of these cDNAs are provided in Table 5.3. In naming the cDNA inserts, the following formula was followed. The initial letter of the name indicates the anchored primer used (G = T₁₂VG, A = T₁₂VA etc.), the second letter indicates the treatment (T = thermoinhibited, C = control) and the third letter followed by a numeral indicates the number of the random decamer used (J1 = OPJ-01, J2 = OPJ-02 etc.). The numeral following the hyphen indicates the order in which the fragment was identified on the relevant differential display gel (-01 indicates the first differentially expressed fragment identified in a gel, etc.).

Following transformation, 50 white putatively recombinant colonies were selected for each of these cDNAs and were plated onto LB agar plates supplemented with the relevant antibiotics. Colony hybridization (Figure 5.2) was then performed, using labeled cDNA as the probes, and the three colonies which produced the strongest hybridization signal were selected for each cDNA clone. These were then cultured in liquid media and the plasmids isolated. These plasmids were then named by adding a forward slash and the number of the colony from which the plasmid was isolated after the name of the cDNA insert contained in the plasmid (e.g. ATJ4-02/1).

Table 5.3: List of cDNAs cloned into the pGEM[®]-T EASY vector and their approximate sizes.

| Clone | Estimated length (bp) |
|----------|-----------------------|
| ATJ4-02 | 250 |
| ATJ6-01 | 285 |
| CTJ4-24 | 360 |
| CTJ4-26 | 295 |
| CTJ4-27 | 280 |
| TTJ6-52 | 270 |
| TTJ6-53 | 270 |
| GTJ1-12 | 170 |
| GTJ3-28 | 295 |
| GTJ4-43 | 165 |
| GTJ4-44 | 150 |
| GTJ10-52 | 335 |
| GTJ10-54 | 170 |

Restriction analysis was performed on each plasmid thus isolated and the plasmids electrophoresed alongside a restricted, non-recombinant vector isolated from the background control plate to ensure that each contained an insert of the expected size. All plasmids thus tested were larger than the control plasmid and thus appeared to contain an insert.

The colony blotting protocol described worked well with *E. coli* JM109 clones, but was problematic with *E. coli* XL1 Blue where all colonies produced a hybridization signal (Figure 5.2). With the latter strain, negative controls prepared using non-transformed colonies also produced a hybridization signal for some of the various probe cDNAs. It thus appears that for this strain, the stringency of the hybridization washes needs to be increased. This procedure was not re-optimized for use with this strain, however, as restriction analysis indicated that each clone selected did in fact contain an insert.

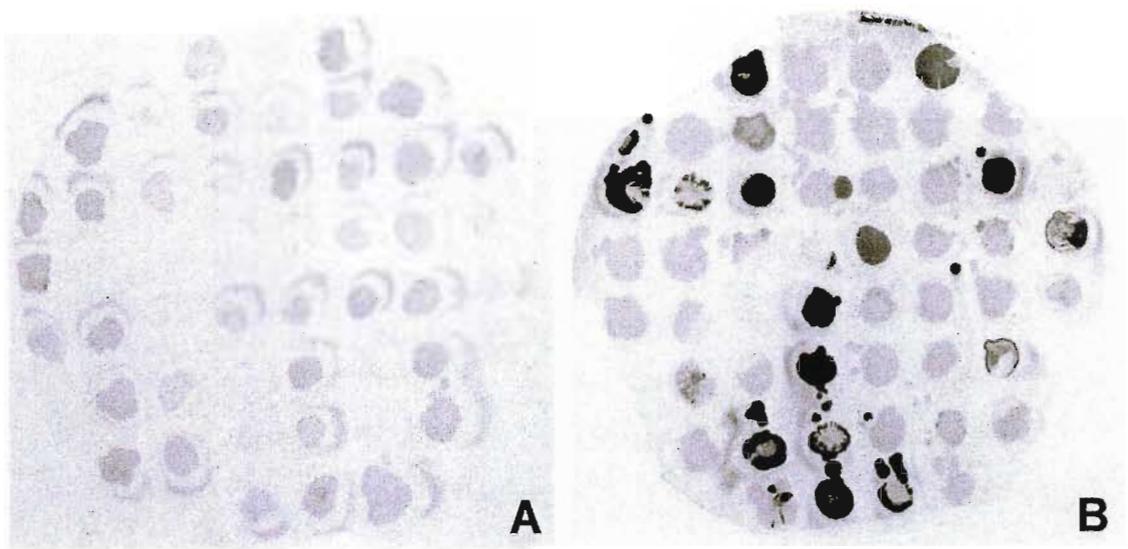


Figure 5.2: Colony blotting of recombinant colonies of *Escherichia coli* JM109 (A) and *E. coli* XL1 Blue (B). The *E. coli* JM109 clones in (A) were transformed with the pGEM[®]-T Easy vector with cDNA insert ATJ4-02, whilst the *E. coli* XL1 Blue colonies (B) contain the vector with cDNA insert GTJ4-43. Whilst *E. coli* XL1 Blue did prove problematic with the colony blotting protocol used as non-recombinant colonies also generated a positive hybridization signal, some colonies produced a much stronger signal than others and were selected for further analysis. Restriction analysis confirmed the presence of a recombinant vector in these colonies.

5.3.2 Reverse-northern blotting of cDNAs

For all 13 of the cDNAs cloned, each of the three clones isolated gave a positive signal when probed with labeled cDNA prepared from thermoinhibited achenes and no signal when probed with cDNA from the germinating control achenes. These 39 cDNAs are thus specific to thermoinhibited achenes.

5.3.3 Sequencing of cDNAs and analysis of sequence data

Automated sequencing was performed on the 39 plasmids at the Agricultural

Biotechnology Centre in Hungary. Initially, all of the plasmids were sequenced from the pUC/M13 reverse sequencing primer. Good sequence data were obtained from 27 plasmids. The remaining 12 plasmids produced low quality sequences with a large number of ambiguities, making the sequence unreadable. This was probably due to degradation of the DNA in transit, as the samples were delayed for several days by customs and excise officials in Hungary. At least one of the three replicate samples for each of the cDNAs produced a readable sequence, with the exception of cDNA TTJ6-52, all three replicates of which produced sequences of low quality. The 27 readable sequences were then analysed for the primer sequences which indicate the beginning and end of the inserts. After checking the vector sequence for at least 20 bp on either side of the insert to ensure that the correct primer sequence had been identified, the sequences were edited to remove the flanking vector sequences. Sequences were aligned using ClustalX version 1.81 (THOMPSON, GIBSON, PLEWNIK, JEANMOUGIN and HIGGINS, 1997) to identify homologous sequences. Sequences are given in Figures 5.3 - 5.10. The original full-length sequences are given in Appendix B.

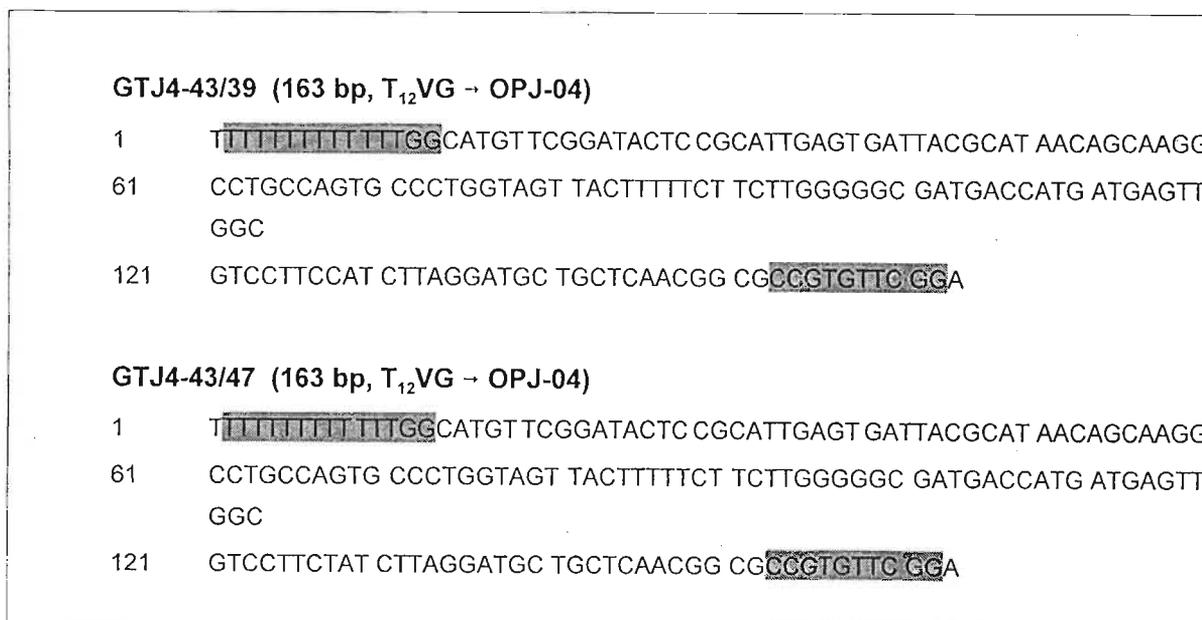


Figure 5.3: Negative strand nucleotide sequences of similar 163 bp cDNA inserts GTJ4-43/39 and GTJ4-43/47 as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. Both inserts are present in the same orientation.

ATJ4-02/1 (249 bp, OPJ-04 → OPJ-04)

1 TCCGAACACG GCCACGTTCC ACTAGTCTTG GTGGTGGATG CACCAAAGTG TCTATAGAAG
61 AGAGTGACTC TGCTTACACA GTTGCTAAGT GATACCCAAG GGTCATGTAG GGTAACATGA
121 AACTGGAATT GAAAAATTAG TCATCCAGGT AATAGTAAGA TCCAGCAAGT TTCTGCTCCC
181 TGTCCTTTGT TGAATTAAGC TTGTTAATCC TTA CTATGTC TGGACCTGGT AAGTTTCCGC
241 GTGTTCCGA

CTJ4-27/30 (249 bp, OPJ-04 → OPJ-04)

1 TCCGAACACG GCCACGTTCC ACTAGTCTTG GTGGTGGATG CACCAAAGTG TCTATAGACG
61 AGAGTGACTC TGCTTACACA GTTGCTAAGT GATACCCAAG GGTCATGTAG GGTAACATGA
121 AACTGGAATT GAAAAATTAG TCATCCAGGT AATAGTAAGA TCCAGCAAGT TTCTGCTCCC
181 TGTCCTTTGT TGAATTAAGC TTGTTAATCC TTA CTATGTC TGGACCTGGT AAGTTTCCGC
241 GTGTTCCGA

ATJ4-02/2 (249 bp, OPJ-04 → OPJ-04)

1 TCCGAACACG GGGAACTTA CCAGGTCCAG ACATAGTAAG AATTAACAAG CTTAATTCAA
61 CAAAGGACAG GGAGCAGAAA CTTGCTGTAT CTTACTATTA CCTGGATGAC TAATTTTCA
121 ATTCCAGTTT CATGTTACCC TACATGACCC TTGGGTATCA CTTAGCAACT GTGTAAGCAG
181 AGTCACTCTC TTCTATAGAC ACTTTGGTGC ATCCACCACC AAGACTAGTG GAACGTGGCC
241 GTGTTCCGA

ATJ4-02/3 (249 bp, OPJ-04 → OPJ-04)

1 TCCGAACACG GGGAACTTA CCAGGTCCAG ACATAGTAAG AATTAACAAG CTTAATTCAA
61 CAAAGGACAG GGAGCAGAAA CTTGCTGGAT CTTACTATTA CCTGGATGAC TAATTTTCA
121 ATTCCAGTTT CATGTTACCC TACATGACCC TTGGGTATCA CTTAGCAACT GTGTAAGCAG
181 AGTCACTCTC TTCTATAGAC ACTTTGGTGC ATCCACCACC AAGACTAGTG GAACGTGGCC
241 GTGTTCCGA

Figure 5.4: Negative strand nucleotide sequences of similar 249 bp cDNA inserts ATJ4-02/1, CTJ4-27/30, ATJ4-02/2 and ATJ4-02/3 as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. Inserts ATJ4-02/1 and CTJ4-27/30 are cloned in the opposite orientation to ATJ4-02/2 and ATJ4-02/3.

TTJ6-53/2 (267 bp, OPJ-06 → OPJ-06)

1 TCGTCCGCAGGCCAAGGA TTTGGATATT GGAATTTAAT ATCNCTAATA ATTACAACAG
61 GAAC TTTGTA CGAAACAGCA GCTCTGAGCT CTTGTAAGCA CCAAAAACTA TCAAAAATGC
121 CTTCCGGAAG CATAAAAACA AACAATTTAG TATGTTTTAC GATCTCTNTC AAATCGTGTA
181 AATCGAAATC GGCTTCTGNG TCNAAGAAAA TACGATAGCC TTNNAGAGCT TGATATAAGC
241 GACCAGCCAT TCCTTGNCGGGAACGAA

TTJ6-53/3 (267 bp, OPJ-06 → OPJ-06)

1 TCGTCCGCACAAGGAATG GCTGGTCGCT TATATCAAGC TCTAAAAGGC TATCGTATTT
61 TCTTAGACAC AGAAGNCGAT TTCGATTTAC ACGATTTGAN NGAGATCGTA AAACATACTA
121 AATTGTTTGT TTTTATACTT ACGGAAGGCA TTTTGTATAG TTTTGGTGC TNACAAGAGC
181 TCAGAGCTGC TGTTTCGTAC AAAGTTCCTG TTGTAATTAT TAGAGATATT AAATTCCAAT
241 ATCCAAATCC TTGGCCGCGGAACGAA

TTJ6-53/50 (267 bp, OPJ-06 → OPJ-06)

1 TCGTCCGCACAAGGAATG GCTGGTCGCT TATATCAAGC TCTAAAAGGC TATCGTATTT
61 TCTTAGACAC AGAAGCCGAT TTCGATTTAC ACGATTTGAA AGAGATCGTA AAACATACTA
121 AATTGTTTGT TTTTATACTT ACGGAAGGCA TTTTGTATAG TTTTGGTGC TTACAAGAGC
181 TCAGAGCTGC TGTTTCGTAC AAAGTTCCTG TTGTAATTAT TAGAGATATT AAATTCCAAT
241 ATCCAAATCC TTGGCCGCGGAACGAA

ATJ6-01/38 (267 bp, OPJ-06 → OPJ-06)

1 TCGTCCGCACAAGGAATG GCTGGTCGCT TATATCAAGC TCTAAAAGGC TATCGTATTT
61 TCTTAGACAC AGAAGCCGAT TTCGATTTAC ACGATTTGAA AGAGATCGTA AAACATACTA
121 AATTGTTTGT TTTTATACTT ACGGAAGGCA TTTTGTATAG TTTTGGTGC TTACAAGAGC
181 TCAGAGCTGC TGTTTCGTAC AAAGTTCCTG TTGTAATTAT TAGAGATATT AAATTCCAAT
241 ATCCAAATCC TTGGCCGCGGAACGAA

ATJ6-01/48 (267 bp, OPJ-06 → OPJ-06)

1 TCGTCCGCACAAGGAATG GCTGGTCGCT TATATCAAGC TCTAAAAGGC TATCGTATTT
61 TCTTAGACAC AGAAGCCGAT TTCGATTTAC ACGATTTGAA AGAGATCGTA AAACATACTA
121 AATTGTTTGT TTTTATACTT ACGGAAGGCA TTTTGTATAG TTTTGGTGC TTACAAGAGC
181 TCAGAGCTGC TGTTTCGTAC AAAGTTCCTG TTGTAATTAT TAGAGATATT AAATTCCAAT
241 ATCCAAATCC TTGGCCGCGGAACGAA

Figure 5.5: Negative strand nucleotide sequences of similar 267 bp cDNA inserts TTJ6-53/2, TTJ6-53/3, TTJ6-53/50, ATJ6-01/38 and ATJ6-01/48 as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. Insert TTJ6-53/2 is present in the opposite orientation to the other sequences.

CTJ4-24/15 (347 bp, OPJ-04 → OPJ-04)

```
1   TCGGAACACGGCCACGTCCA CCTCCACCAG AAGGTAGT TCTTTCACCA GCAGCAGGTT
61  TCAAAAACCTC ATTGATGCTG AGTGACTTCT TGGCTCTCTC TTCTTTCTCA GCAAGCTCTT
121 TACGCTTATC CTTGTCAGTC CCCAGTTTAG CAAAGATTTT ATCGTTGCTC TTCTTGTTGG
181 AGAGTTGTTG CATAGATGCA AGCTCTTGT CAACTTCCAC TTTTCTTCTC TCAGTCTTGA
241 GTGTCTCTAA AGCTTTCCTC TTTTCTCAA GAACCTTTTG ATATTCTTCA AGAGTCATCT
301 CCTATCCTC AGGCTCCTTC TCCTCAGTTT CATTACCGTGTTCGGGA
```

CTJ4-24/43 (347 bp, OPJ-04 → OPJ-04)

```
1   TCGGAACACGGCCACGTCCA CCTCCACCAG AAGGTAGT TCTTTCACCA GCAGCAGGTT
61  TCAAAAACCTC ATTGATGCTG AGTGACTTCT TGGCTCTCTC TTCTTTCTCA GCAAGCTCTT
121 TACGCTTATC CTTGTCAGTC CCCAGTTTAG CAAAGATTTT ATCGTTGCTC TTCTTGTTGG
181 AGAGTTGTTG CATAGATGCA AGCTCTTGT CAACTTCCAC TTTTCTTCTC TCAGTCTTGA
241 GTGTCTCTAA AGCTTTCCTC TTTTCTCAA GAACCTTTTG ATATTCTTCA AGAGTCATCT
301 CCTATCCTC AGGCTCCTTC TCCTCAGTTT CATTACCGTGTTCGGGA
```

Figure 5.6: Negative strand nucleotide sequences of similar 347 bp cDNA inserts CTJ4-24/15 and CTJ4-24/43 as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. Both inserts are in the same orientation.

GTJ10-54/1 (154 bp, OPJ-10 → OPJ-10)

```
1   TAAGCCCGAGGGTAAGTAAT TCATCAATTA GCTTCGTAAT TTAAATTGAT ATGCTTGCAA
61  CTATTTTATT TTGTGTGATC ACCGCTTGAT TCTCACTGGA TTAGCAGCCT TCCGTTAGCT
121 ACCACACAAG CAACTGTTGT TCGCCTCGGGCTTA
```

GTJ10-54/2 (154 bp, OPJ-10 → OPJ-10)

```
1   TAAGCCCGAGGGTAAGTAAT TCATCAATTA GCTTCGTAAT TTAAATTGAT ATGCTTGCAA
61  CTATTTTATT TTGTGTGATC ACCGCTTGAT TCTCACTGGA TTAGCAGCCT TCCGTTAGCT
121 ACCACACAAG CAACTGTTGT TCGCCTCGGGCTTA
```

GTJ10-54/3 (154 bp, OPJ-10 → OPJ-10)

```
1   TAAGCCCGAGGGTAAGTAAT TCATCAATTA GCTTCGTAAT TTAAATTGAT ATGCTTGCAA
61  CTATTTTATT TTGTGTGATC GCGCTTGAT TCTCACTGGA TTAGCAGCCT TCCGTTAGCT
121 ACCACACAAG CAACTGTTGT TCGCCTCGGGCTTA
```

Figure 5.7: Negative strand nucleotide sequences of similar 154 bp cDNA inserts GTJ10-54/1, GTJ10-54/2 and GTJ10-54/3 as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. All inserts are present in the same orientation.

GTJ10-52/17 (318 bp, OPJ-10 → OPJ-10)

```
1   TAAGCCCGAGGGACAATACA CCAACTTATA GTCATCCTCA TACTTCTCAA TCTTGAACCA
61  GTTATCTATG GTTCCTTGCC CGGGCCTTCC TAATATTCCG CGGGTATACA AATAACGTTG
121 CCCGTAAAC TGGTTGATCT GCCAAACCGT GTTGGGTGCA CATGTGACTT GCTCATTGAA
181 CTTAATGTTA AGATCAACAT TTTCGCGAAT GATGCCATCT TTGTTGGCGG ATATAAAATT
241 CAAGGGTGAC CCGTTTCTCC CTTCACTAAC TTCTTGACC ACTTCACGTG GGCATAGTTC
301 GCCTGGT CCTCGGGCTTA
```

GTJ10-52/25 (318 bp, OPJ-10 → OPJ-10)

```
1   TAAGCCCGAGGGACAATACA CCAACTTATA GTCATCCTCA TACTTCTCAA TCTTGAACCA
61  GTTATCTATG GTTCCTTGCC CGGGCCTTCC TAATATTCCG CGGGTATACA AATAACGTTG
121 CCCGTAAAC TGGTTGATCT GCCAAACCGT GTTGGGTGCA CATGTGACTT GCTCATTGAA
181 CTTAATGTTA AGATCAACAT TTTCGCGAAT GATGCCATCT TTGTTGGCGG ATATAAAATT
241 CAAGGGTGAC CCGTTTCTCC CTTCACTAAC TTCTTGACC ACTTCACGTG GGCATAGTTC
301 GCCTGGT CCTCGGGCTTA
```

GTJ10-52/33 (318 bp, OPJ-10 → OPJ-10)

```
1   TAAGCCCGAGGGACAATACA CCAACTTATA GTCATCCTCA TACTTCTCAA TCTTGAACCA
61  GTTATCTATG GTTCCTTGCC CGGGCCTTCC TAATATTCCG CGGGTATACA AATAACGTTG
121 CCCGTAAAC TGGTTGATCT GCCAAGCCGT GTTGGGTGCA CATGTGACTT GCTCATTGAA
181 CTTAATGTTA AGATCAACAT TTTCGCGAAT GATGCCATCT CTGTTGGCGG ATATAAAATT
241 CAAGGGTGAC CCGTTTCTCC CTTCACTAAC TTCTTGACC ACTTCACGTG GGCATAGTTC
301 GCCTGGT CCTCGGGCTTA
```

Figure 5.8: Negative strand nucleotide sequences of similar 318 bp cDNA inserts GTJ10-52/17, GTJ10-52/25 and GTJ10-52/33 as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. All inserts are present in the same orientation.

Figure 5.9: Negative strand nucleotide sequences of cDNA inserts, as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector, which did not align with any other sequences. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted.

CTJ4-24/5 (347 bp, ? → OPJ-04)

1 TTTAACGACG GCCAGTCCAT GATACACAGC AGAAGGAACT TCCCTCCCAC TATGTAACGA
61 TTGTTGTGTC CTCACATTAT CTTCTAGATC TTCTAACTTC TTCNGTAAT ACTCCACGAG
121 ATCAACTTTG CGCCCCAAAA TCCCAAGAAA TCCAACACGC CCAAACCTTT GAGAAGTCTT
181 TTGTGTTGAC TTTAGATGAG CAAGCCTTCT ATTCAGCTTA TCAGCATCNC TAATTAGCTT
241 CTGAATCTTG GTAGTGTGAT GCACCATGTA ATGTGACAGA TATAATGAAG GATAATATTC
301 TGAGAAAAAA TTCTCAACAC TTGCACTTAA GCTGGCCCGTGTTCGGA

CTJ4-26/3 (282 bp, OPJ-04 → OPJ-04)

1 TCCGAACACG GCAGATAAT CAACAGAAAT GTCTGGGATG ATGTAAAATG AAATCAAATA
61 TTTATATTCA TACATTTTGA CACATGTAAA CTCTTCTTAA AATCTCCAGG TTGACCTTTG
121 ACTTTTCAGT TAACTTGACT GATGCCTGGT TTTCTTTGAA AGAGAAGTGA CCCAATAAGA
181 ATGGCAATGC CAGTTATTGC TATGGCGAAA AGAAGCCGCC TCGTAGAGAT GTCTTGTCAA
241 TCCTTACTAT GTCTGGACCT GGNAAGTTTC CCGTGTTCGGA

CTJ4-27/16 (262 bp, T₁₂VC → OPJ-04)

1 TTTTTTTTTTTTGGCAATT GAAAAGAGAT AGCAAGGAGA AACATTGTT CTAGTCAACA
61 TCCAAAGGTC TTGAAAGCAA AATGGGTCGG GTACAAATCA TGCAAACAAT ATCCATACTC
121 AAAACCCTAA CCACAATGAC CAAGTCAAAC CACCACACAT GTTGACTAAA TAATCTATGT
181 GCTGCTAGCT GCTGCTGCTG CAACTTCTGC TTTCTCCTCA GAACCTCCTT CTGCTGTGTA
241 TCATGACTGG CCGTGTTCGGA

GTJ1-12/34 (143 bp, OPJ-01 → T₁₂VG)

1 TCCCGGCATA AATAATGTAG ATGGTGAAAA CTAGAAAGTC AAAGTTTTAT TTATAATTTA
61 AAAGTCAAAG CTTTACATGA GAAATGTAGG ATATTTTACG TTATAATTGT ATTTATAAAT
121 CTTTCTGCC AAAAAAAAAAAA AAA

GTJ3-28/5 (233 bp, T₁₂VG → OPJ-03)

1 TTTTTTTTTTTTGGCACCT CAACCCTGTA ATAGTTTTTT TTTTAACTG AAGCCATATA
61 TGATATTTAA ATGTTTCAA CTACTAACCA CACCGTCACT TGAAACAAAC ATGGATAACA
121 ACTAACAACA TTCGTTTTAA GTAACATGAA GTAGATAAAG AGGAGGGATA AATATGACAT
181 GAATTTGCTC ATAAAGACAT GTTACCTTG CTCGTTTTGG GCDAAGCGGA GAA

GTJ4-44/16 (87 bp, T₁₂VG → ?)

1 TTTTTTTTTTTTGGGGTGA TAGTGAAAA GTTGACAGAT AAGTCACCCT TACTGCCACT
61 CTACAGAACC GTACATGAGA TTTTCAA

GTJ4-44/25 (152 bp, OPJ-04 → OPJ-04)

1 TCCGAACACG GCAAAAAAGC AAGACCAGAC ATAACGTGAA TCATCCTCAA TATGGAACCA
61 AATTTAAACC TAATACATGC CATTACCAAC CACAGTTATT CCATCACTGT TGTCACATAA
121 AATCTTTCCC ACCATATCTT CCGTGTTCGGA

GTJ4-43/2 (310 bp, OPJ-04 → T₁₂VG → OPJ-04)

```
1  TCCGAACAGG GGAAGATATG GTGGGAAAGA TTTTATGTGA CAACAGTGAT GGAATAACTG
61  TGGTTGGTAA TGGTATGTAT TAGGTTTTAA TTTGGTTCCA TATTGAGGAT GATTCACGTT
121 ATGTCTGGTC TTGCTTTTTT GGGCGTTTTT TTTTTTTTGG GACCACTTGA AAAGAAATTA
181 TAAAGTGTCT TTTTATAGGA ATTTGAATTT TAGATGTTAT ATACTCATAT AGCACTGTAA
241 ATTTACAAC TGCAGCAGC AGCAGCAGCA TATGTAAAAA ACCTCCCAGC TTAAGTAGGG
301 CGTCTCGGA
```

Figure 5.10: Negative strand nucleotide sequences of cDNA insert GTJ4-43/2, as sequenced from the pUC/M13 reverse sequencing primer on the pGEM[®]-T EASY vector, which did not align with any other sequences. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted. This insert contains an oligo(dT) section (highlighted in the centre of the sequence) which could have resulted from the ligation of two different sequences, both with T₁₂VG ends and an A/T tail from PCR, in opposite orientations.

Alignment of the sequences indicated that for a number of the differential display gel bands, several sequences were present and reamplified, leading to different sequences being cloned into the vector plasmid from the same reamplified cDNA sample. An example of this was cDNA CTJ4-27, where clone CTJ4-27/16 contained a novel sequence whilst clone CTJ4-27/30 aligned with the sequences for clones prepared from cDNA ATJ4-02. The third clone from the CTJ4-27 cDNA produced a non-readable sequence. Such contamination has previously been reported as occurring during differential display (BAUER, MÜLLER, REICH, RIEDEL, AHRENKIEL, WARTHOE and STRAUSS, 1993; CALLARD, LESCURE and MAZZOLINI, 1994; LINSKENS, FENG, ANDREWS, ENLOW, SAATI, TONKIN, FUNK and VILLEPONTEAU, 1995). In this case, since all recombinant plasmid samples were analysed using the reverse-northern blot, all sequences were nonetheless specific to thermoinhibited achenes.

A further point of interest was that several different cDNAs produced clones which contained homologous sequences. Clone CTJ4-27/30 and the ATJ4-02 clones have already been mentioned. All three clones from cDNA TTJ6-53 aligned with the two clones from cDNA ATJ6-01 which gave readable sequences. This phenomenon may

be explained by the fact that in these sequences, the cDNA was reamplified from the random decamer primer at both ends. Reamplification from the random primer alone was relatively commonly observed for many of the cDNAs in addition to those mentioned above (Figures 5.3 - 5.10). Since non-reverse transcribed total RNA did not produce any bands on the differential display gels following PCR, it may be assumed that these fragments were synthesized from first strand cDNA, such that a family of fragments may have been synthesized: a short fragment between the anchored primer and the first binding site of the random decamer, the sequences represented here which were generated between the random decamer binding sites and a series of longer fragments from the anchored primer to the further binding site for the random decamer. None of these other fragments were among those sequenced, but may be among the other fragments identified but not yet cloned. This suggests that the sequences of primers in the OPJ-series may be relatively common in the genes represented by these cDNAs and in any future studies, a different series of random primers should possibly be used with this system.

For ease of further analysis, once similar sequences had been grouped by alignment, these groups were then identified by the epithet THB- followed by a number. These alignment groups are shown in Table 5.4.

Table 5. 4: Alignment groups for sequences grouped by homology following sequencing.

| Alignment group | Clones represented |
|------------------------|--|
| THB-1 | GTJ4-43/39; GTJ4-43/47 |
| THB-2 | ATJ4-02/1; ATJ4-02/2; ATJ4-02/3; CTJ4-27/30 |
| THB-3 | ATJ6-01/38; ATJ6-01/48; TTJ6-53/2; TTJ6-53/3; 50 |
| THB-4 | CTJ4-24/15; CTJ4-24/43 |
| THB-5 | GTJ10-54/1; GTJ10-54/2; GTJ10-54/3 |
| THB-6 | GTJ10-52/17; GTJ10-52/25; GTJ10-52/33 |

Homology searches were conducted by comparing the insert sequences with known sequences in the NCBI database using the BLAST programs (ALTSCHUL, GISH, MILLER, MYERS and LIPMAN, 1990; ALTSCHUL, MADDEN, SCHÄFFER, ZHANG,

ZHANG, MILLER and LIPMAN, 1997) as described in Section 5.2.11. One sequence from each of the alignment groups was selected for the homology search. If the insert was cloned in different directions within the group, one example from each orientation was used. Results are given in Table 5.5.

In general, relatively short sequences were aligned by the BLAST tool. Results generated from BLASTX (translated nucleotide-amino acid) searches may therefore be more valuable than those generated by BLASTN (nucleotide-nucleotide) searches, as generally longer sequences were compared. Four sequences, ATJ4-02/1 - ATJ4-02/2 (identical sequences in opposite orientations), CTJ4-24/15, GTJ4-43/39 and GTJ10-52/25, showed relatively strong homology (bit score of greater than 45 and extremely low E-value of less than 0.01 for both BLASTN and BLASTX searches). With the exception of the THB-1 clones, these clones, as well as the other cDNA sequences within their alignment group, were sequenced in the opposite orientation along the positive strand of the plasmid vector from the pUC/M13 forward primer. These positive sequences were also edited to remove vector sequences and then aligned with the reverse-complement sequences from the corresponding reverse sequences to obtain a consensus sequence (positive strand) for each of the remaining three groups of cDNAs (Figure 5.11). The original full-length sequences are given in Appendix B. The THB-1 sequences were not re-sequenced as they almost certainly encode contaminating bacterial genes. The presence of such microflora amongst the plant material, especially with field grown plant material, is a hazard associated with differential display as it is such a sensitive procedure (HECK and FERNANDEZ, 1997). Confirmation of whether or not this sequence does result from bacterial contamination could be determined by Southern blotting, however this was not done due to time considerations.

Of the four cDNA alignment groups which showed homology to known genes, three were generated from the random primers only, and not from the anchored primers. The cDNAs isolated during differential display are typically short and are from the 3' end of the mRNA sequence. As a result, they are often from the untranslated region of the mRNA and often lack homology to known protein sequences (WAN, SHARP,

Table 5.5: Summary of results from BLASTN and BLASTX searches for sequences showing homology to the thermoinhibition-associated cDNAs isolated from *Tagetes minuta* achenes. Only the three highest scoring alignments are listed for each search, in descending order of significance.

| | Sequence Description | % | length | Score (bits) | E |
|------------------|---|-----|--------|--------------|-------|
| ATJ4-02/1 | | | | | |
| BLASTN | 1. <i>Megalodonta beckii</i> small subunit ribosomal RNA gene, partial sequence | 100 | 40 bp | 79.8 | 2e-12 |
| | 2. Cryptoperidiniopsoid sp. Brodyi 18S ribosomal RNA gene, partial sequence | 100 | 40 bp | 79.8 | 2e-12 |
| | 3. <i>Cabomba caroliniana</i> small subunit ribosomal RNA gene, partial sequence | 100 | 40 bp | 79.8 | 2e-12 |
| BLASTX | 1. RNase L inhibitor-like protein (<i>Oryza sativa</i> japonica cultivar-group) | 91 | 24 aa | 48.5 | 2e-5 |
| | 2. ABC-type transport protein T18B16.180 (<i>Arabidopsis thaliana</i>) -RNase L inhibitor-like protein (<i>Arabidopsis thaliana</i>) -RNase L inhibitor like protein (<i>Arabidopsis thaliana</i>) | 91 | 24 aa | 48.5 | 3e-5 |
| | 3. RNase L inhibitor protein, putative, protein id At4g19210.1, supported by cDNA gi_20466461 (<i>Arabidopsis thaliana</i>) -RNase L inhibitor-like protein (<i>Arabidopsis thaliana</i>) -RNase L inhibitor like protein (<i>Arabidopsis thaliana</i>) | 91 | 24 aa | 48.1 | 3e-5 |
| ATJ4-02/2 | | | | | |
| BLASTN | 1. <i>Marah macrocarpus</i> 18S ribosomal RNA gene, complete sequence | 100 | 37 bp | 73.8 | 1e-10 |
| | 2. <i>Megalodonta beckii</i> small subunit ribosomal RNA gene, partial sequence | 100 | 36 bp | 71.9 | 4e-10 |
| | 3. Cryptoperidiniopsoid sp. Brodyi 18S ribosomal RNA gene, partial sequence | 100 | 36 bp | 71.9 | 4e-10 |

| | Sequence Description | % | length | Score (bits) | E |
|-------------------|---|-----|--------|-----------------|-------|
| BLASTX | 1. RNase L inhibitor-like protein (<i>Oryza sativa</i> japonica cultivar-group) | 87 | 24 aa | 45.1 | 2e-4 |
| | 2. RNase L inhibitor protein, putative, protein id At4g19210.1, supported by cDNA gi_20466461 (<i>Arabidopsis thaliana</i>) -RNase L inhibitor-like protein (<i>Arabidopsis thaliana</i>) -RNase L inhibitor like protein (<i>Arabidopsis thaliana</i>) | 87 | 24 aa | 44.7 | 3e-4 |
| | 3. ABC-type transport protein T18B16.180 (<i>Arabidopsis thaliana</i>) -RNase L inhibitor-like protein (<i>Arabidopsis thaliana</i>) -RNase L inhibitor like protein (<i>Arabidopsis thaliana</i>) | 87 | 24 aa | 44.7 | 3e-4 |
| ATJ6-01/48 | | | | | |
| BLASTN | 1. <i>Mycoplasma penetrans</i> genomic DNA, complete sequence, section 3/5 | 93 | 31 bp | 46.1 | 0.025 |
| | 2. <i>Mus musculus</i> strain C57BL6/J Chromosome 5 BAC, RP23-72621, complete sequence | 100 | 21 bp | 42.1 | 0.39 |
| | 3. <i>Mus musculus</i> strain C57BL6/J Chromosome 5 RP23-319P12, complete sequence | 100 | 21 bp | 42.1 | 0.39 |
| BLASTX | 1. SAM (Sterile Alpha Motif) and TIR domain-containing protein (<i>Caenorhabditis elegans</i>) - Hypothetical protein F13B10.1a (<i>Caenorhabditis elegans</i>) - Hypothetical protein F13B10.1a (<i>Caenorhabditis elegans</i>) | 30 | 80 aa | 37.4 | 0.050 |
| | 2. SAM domain and TIR domain-containing protein (66.0 KDa) (<i>Caenorhabditis elegans</i>) - Hypothetical protein F13B10.1a (<i>Caenorhabditis elegans</i>) - Hypothetical protein F13B10.1a (<i>Caenorhabditis elegans</i>) | 30 | 80 aa | 37.4 | 0.050 |
| | 3. Resistance protein MG23 (<i>Glycine max</i>) | 30 | 68 aa | 36.6 | 0.085 |
| TTJ6-53/2 | | | | | |
| BLASTN | 1. <i>Mus musculus</i> strain C57BL6/J Chromosome 5 BAC, RP23-72621, complete sequence | 100 | 21 bp | 42.1 | 0.39 |
| | 2. <i>Mus musculus</i> strain C57BL6/J Chromosome 5 RP23-319P12, complete sequence | 100 | 21 bp | 42.1 | 0.39 |
| | 3. <i>Homo sapiens</i> BAC clone RP11-638I8 from 7, complete sequence | 100 | 21 bp | 42.1 | 0.39 |
| BLASTX | 1. Flax rust resistance protein (<i>Linum usitatissimum</i>) | 32 | 61 aa | 32.7 | 1.2 |
| | 2. Flax rust resistance protein L6 | 32 | 61 aa | 32.7 | 1.2 |
| | 3. Large protein (Flanders virus) | 32 | 49 aa | 32.7 | 1.2 |

| | Sequence Description | % | length | Score (bits) | E |
|-------------------|---|-----|--------|--------------|-------|
| CTJ4-24/5 | | | | | |
| BLASTN | 1. <i>Homo sapiens</i> chromosome 20 clone h119, complete sequence | 100 | 21 bp | 42.1 | 0.51 |
| | 2. <i>Homo sapiens</i> nuclear receptor coactivator 3 (NCOA3) mRNA | 100 | 21 bp | 42.1 | 0.51 |
| | 3. <i>Homo sapiens</i> Amplified in Breast Cancer (AIB1) mRNA, complete cds | 100 | 21 bp | 42.1 | 0.51 |
| | | | | | |
| BLASTX | 1. Unknown protein (<i>Arabidopsis thaliana</i>) | 48 | 114 aa | 107 | 5e-23 |
| | 2. Hypothetical protein HYP1 (<i>Arabidopsis thaliana</i>) - HYP1 (<i>Arabidopsis thaliana</i>) | 45 | 79 aa | 68.6 | 2e-11 |
| | 3. Unknown protein At3g01100.1 (<i>Arabidopsis thaliana</i>) - HYP1 (<i>Arabidopsis thaliana</i>) - Unknown protein (<i>Arabidopsis thaliana</i>) | 45 | 79 aa | 68.6 | 2e-11 |
| CTJ4-24/15 | | | | | |
| BLASTN | 1. <i>Vicia faba</i> mRNA for putative transcription factor (1556 bp) | 83 | 102 bp | 67.9 | 9e-9 |
| | 2. <i>Zea mays</i> PCO063451 mRNA sequence | 86 | 72 bp | 63.9 | 1e-7 |
| | 3. <i>Arabidopsis thaliana</i> nuclear antigen homologue mRNA, complete cds | 94 | 39 bp | 61.9 | 6e-7 |
| BLASTX | 1. P0681B11.2 <i>Oryza sativa</i> japonica cultivar-group -Putative nuclear RNA binding protein A (<i>Oryza sativa</i> japonica cultivar-group) | 60 | 94 aa | 102.0 | 9e-22 |
| | 2. Putative nuclear RNA binding protein A (<i>Oryza sativa</i> japonica cultivar-group) | 56 | 92 aa | 94.4 | 3e-19 |
| | 3. Probable transcription factor - Fava bean -transcription factor (<i>Vicia faba</i>) | 56 | 95 aa | 90.1 | 6e-18 |
| CTJ4-25/3 | | | | | |
| BLASTN | 1. <i>Megalodonta beckii</i> small subunit ribosomal RNA gene, partial sequence | 97 | 44 bp | 81.8 | 5e-13 |
| | 2. Cryptoperidiniopsis sp. Brodyi 18S ribosomal RNA gene, partial sequence | 97 | 44 bp | 81.8 | 5e-13 |
| | 3. <i>Cabomba caroliniana</i> small subunit ribosomal RNA gene, partial sequence | 97 | 44 bp | 81.8 | 5e-13 |

| | Sequence Description | % | length | Score (bits) | E |
|-------------------|---|-----|--------|--------------|-------|
| BLASTX | 1. Hypothetical protein (imported) (<i>Arabidopsis thaliana</i>) | 51 | 37 aa | 34.7 | 0.32 |
| | 2. Hypothetical protein (<i>Desulfotobacterium hafniense</i>) | 62 | 24 aa | 33.1 | 0.92 |
| | 3. Hypothetical protein (<i>Yersinia pestis</i> KIM) | 27 | 81 aa | 30.4 | 6.0 |
| CTJ4-27/16 | | | | | |
| BLASTN | 1. <i>Homo sapiens</i> nuclear receptor coactivator 3 (NCOA3) mRNA | 100 | 22 bp | 44.1 | 0.096 |
| | 2. <i>Homo sapiens</i> Amplified in Breast Cancer (AIB1) mRNA, complete cds | 100 | 22 bp | 44.1 | 0.096 |
| | 3. <i>Homo sapiens</i> nuclear receptor coactivator (ACTR) mRNA, complete cds | 100 | 22 bp | 44.1 | 0.096 |
| BLASTX | 1. Hypothetical protein (<i>Plasmodium falciparum</i> 3D7) - Hypothetical protein (<i>Plasmodium falciparum</i> 3D7) | 41 | 39 aa | 33.5 | 0.73 |
| | 2. Similar to ORF2 (<i>Mus musculus domesticus</i>) | 33 | 45 aa | 30.4 | 6.2 |
| | 3. Hypothetical protein, conserved (<i>Plasmodium falciparum</i> 3D7) -Hypothetical protein, conserved (<i>Plasmodium falciparum</i> 3D7) | 44 | 29 aa | 30 | 8.0 |
| GTJ1-12/34 | | | | | |
| BLASTN | 1. Human DNA sequence from clone RP11-490F3 on chromosome 9, contains a cytochrome C pseudogene, 2 isoforms of the 5' end of a gene for a novel protein and a CpG island, complete sequence | 100 | 25 bp | 50.1 | 8e-4 |
| | 2. <i>Homo sapiens</i> chromosome 5 clone CTD-2017K17, complete sequence | 93 | 31 bp | 46.1 | 0.012 |
| | 3. <i>Homo sapiens</i> chromosome 5 clone RP11-305P14, complete sequence | 93 | 31 bp | 46.1 | 0.012 |
| BLASTX | No significant similarities found | - | - | - | - |
| GTJ3-28/5 | | | | | |
| BLASTN | 1. <i>Homo sapiens</i> chromosome 5 clone CTD-2198K16, complete sequence | 95 | 24 bp | 40.1 | 1.3 |
| | 2. <i>Homo sapiens</i> chromosome 5 clone CTD-2004c12, complete sequence | 95 | 24 bp | 40.1 | 1.3 |
| | 3. <i>Ciona intestinalis</i> cDNA, clone cicl048a23, full insert sequence | 100 | 20 bp | 40.1 | 1.3 |

| | Sequence Description | % | length | Score (bits) | E |
|-------------------|---|-----|--------|--------------|-------|
| BLASTX | 1. Hypothetical protein (<i>Desulfitobacterium hafniense</i>) | 26 | 52 aa | 30.8 | 4.9 |
| | 2. Transforming growth factor - beta-inducible early growth response protein 2 (TGFB-inducible early growth response protein 2) (TIEG-2) (Krueppel-like factor II) -TGFB inducible Spl-like protein (<i>Mus musculus</i>) | 39 | 41 aa | 30.0 | 8.3 |
| | 3. Alpha-1 calcium channel subunit Cch1 (<i>Candida albicans</i>) | 32 | 55 aa | 30.0 | 8.3 |
| GTJ4-43/2 | | | | | |
| BLASTN | 1. <i>Homo sapiens</i> chromosome 18, clone RP11-2503, complete sequence | 96 | 26 bp | 44.1 | 0.11 |
| | 2. <i>Homo sapiens</i> genomic DNA, chromosome 11q, clone RP11-299, complete sequence | 100 | 22 bp | 44.1 | 0.11 |
| | 3. <i>Mus musculus</i> chromosome 16 clone RP24-boN23, complete sequence | 100 | 20 bp | 40.1 | 1.8 |
| BLASTX | No significant similarities found | - | - | - | - |
| GTJ4-43/39 | | | | | |
| BLASTN | 1. <i>Pseudomonas aeruginosa</i> PA01, section 263 of 529 of the complete genome | 93 | 143 bp | 212 | 1e-52 |
| | 2. <i>Pseudomonas syringae</i> pv. <i>syringae</i> B728a Psyr_1, whole genome shotgun sequence | 87 | 129 bp | 121 | 3e-25 |
| | 3. <i>Pseudomonas syringae</i> pv. <i>syringae</i> threonyl-tRNA synthetase (thrS) gene, partial cds, initiation factor IF3 (infC), ribosomal protein L53 (rpmI) and ribosomal protein L20 (rplT) genes, complete cds | 87 | 129 bp | 121 | 3e-25 |
| BLASTX | 1. Hypothetical protein (<i>Pseudomonas syringae</i> pv. <i>syringae</i> B728a) | 92 | 26 aa | 53.1 | 1e-6 |
| | 2. Hypothetical protein (<i>Pseudomonas fluorescens</i>) | 92 | 26 aa | 53.1 | 1e-6 |
| | 3. Translation initiation factor IF3 - Initiation factor IF3 (<i>Pseudomonas fluorescens</i>) | 92 | 26 aa | 53.1 | 1e-6 |

| | Sequence Description | % | length | Score (bits) | E |
|--------------------|--|-----|--------|--------------|-------|
| GTJ4-44/16 | | | | | |
| BLASTN | 1. <i>Lactuca sativa</i> ribosomal protien S7 (rps7) and ribosomal protein S12 (rps12) genes, partial cds; hypothetical protein gene, complete cds; tRNA-Val gene, complete sequence; and 16S ribosomal RNA gene, partial sequence; chloroplast genes for chloroplast products | 100 | 73 bp | 145 | 1e-32 |
| | 2. <i>Arabidopsis suecica</i> isolate S361_9 ribosomal protein S12 (rps12) gene, partial cds; chloroplast genes for chloroplast products | 100 | 73 bp | 145 | 1e-32 |
| | 3. <i>Arabidopsis suecica</i> isolate S354_9 ribosomal protein S12 (rps12) gene, partial cds; chloroplast genes for chloroplast products | 100 | 73 bp | 145 | 1e-32 |
| BLASTX | No significant similarities found | - | - | - | - |
| GTJ4-44/25 | | | | | |
| BLASTN | 1. Human DNA sequence from clone RP11-444D13 on chromosome 1, complete sequence | 100 | 20 bp | 40.1 | 0.82 |
| | 2. <i>Homo sapiens</i> Franconi anaemia complementation group D2 protein (FANCD2) gene, exon 18 | 100 | 19 bp | 38.2 | 3.2 |
| | 3. <i>Homo sapiens</i> BAC clone RP11-204o24 from 2, complete sequence | 100 | 19 bp | 38.2 | 3.2 |
| BLASTX | No significant similarities found | - | - | - | - |
| GTJ10-52/25 | | | | | |
| BLASTN | 1. <i>Youngia japonica</i> gene for miraculin homologue, partial cds, isolate 2 | 94 | 39 bp | 61.9 | 5e-7 |
| | 2. <i>Youngia japonica</i> gene for miraculin homologue and insertion of group I intron sequence, isolate 1 | 92 | 42 bp | 60.0 | 2e-6 |
| | 3. <i>Nicotiana tabacum</i> tumour-related protein mRNA, complete cds | 85 | 54 bp | 44.1 | 0.12 |
| BLASTX | 1. Miraculin homologue, <i>Youngia japonica</i> | 51 | 99 aa | 108 | 2e-23 |
| | 2. Miraculin homologue, <i>Taraxacum officianale</i> | 48 | 99 aa | 105 | 1e-22 |
| | 3. Miraculin homologue, <i>Youngia japonica</i> | 48 | 99 aa | 105 | 1e-22 |

| | Sequence Description | % | length | Score (bits) | E |
|-------------------|--|-----|--------|--------------|------|
| GTJ10-54/2 | | | | | |
| BLASTN | 1. Mouse DNA sequence from clone RP23-215P8 on chromosome 2, complete sequence | 100 | 20 bp | 40.1 | 0.83 |
| | 2. <i>Oryza sativa</i> (japonica cultivar-group) genomic DNA, chromosome 7, BAC clone OJ1047_A06 | 100 | 20 bp | 40.1 | 0.83 |
| | 3. <i>Daucus carota</i> antifreeze protein (AFP) mRNA, complete cds | 100 | 20 bp | 40.1 | 0.83 |
| BLASTX | No significant similarities found | - | - | - | - |

Sequence description: Summary description of database sequences.

%: Percentage homology for the specific fragments compared in the search.

Length: Number of base pairs (bp) or amino acids (aa) which were compared in each specific search by BLASTN and BLASTX respectively.

Score (bits): Scoring system which allows for comparisons of significance of scores from different alignments and scoring systems. The bit score is calculated from a raw score, where the individual scores of all letter-letter and letter-null positions are summed, by normalizing the statistical variables which define a scoring system..

E: Expect-value, represents the number of hits with a score equal to or better than the alignment score which would be expected by chance when searching a database of a particular size. Alignments with E-values of greater than approximately 0.1 are unlikely to reflect true sequence relatives.

POIRIER, WAGAMAN, CHAMBERS, PYATI, HOM, GALINDO, HUVAR, PETERSON, JACKSON and ERLANDER, 1996). These cDNAs are thus most often used as probes to identify longer sequences which are more readily identified in homology searches from cDNA or genomic libraries. One way around this problem lies in following a procedure such as that described by SOCKOLOV and PROCKOP (1994), in which a similar approach to differential display is followed except that the cDNAs are generated from arbitrary primers and so contain internal sequences from the mRNA molecules which may be more readily identified. The manner in which these three cDNA groups showed internal priming is analogous to such an approach and may explain why significant homology to known genes was identified for these sequences. Unfortunately, although other of the differentially expressed cDNAs were similarly generated from the random primers only, such fortuitous results were not obtained with all of these. The fact that the majority of the sequences did not show any highly significant homology to known genes may also indicate that they represent

novel genes related to thermoinhibition.

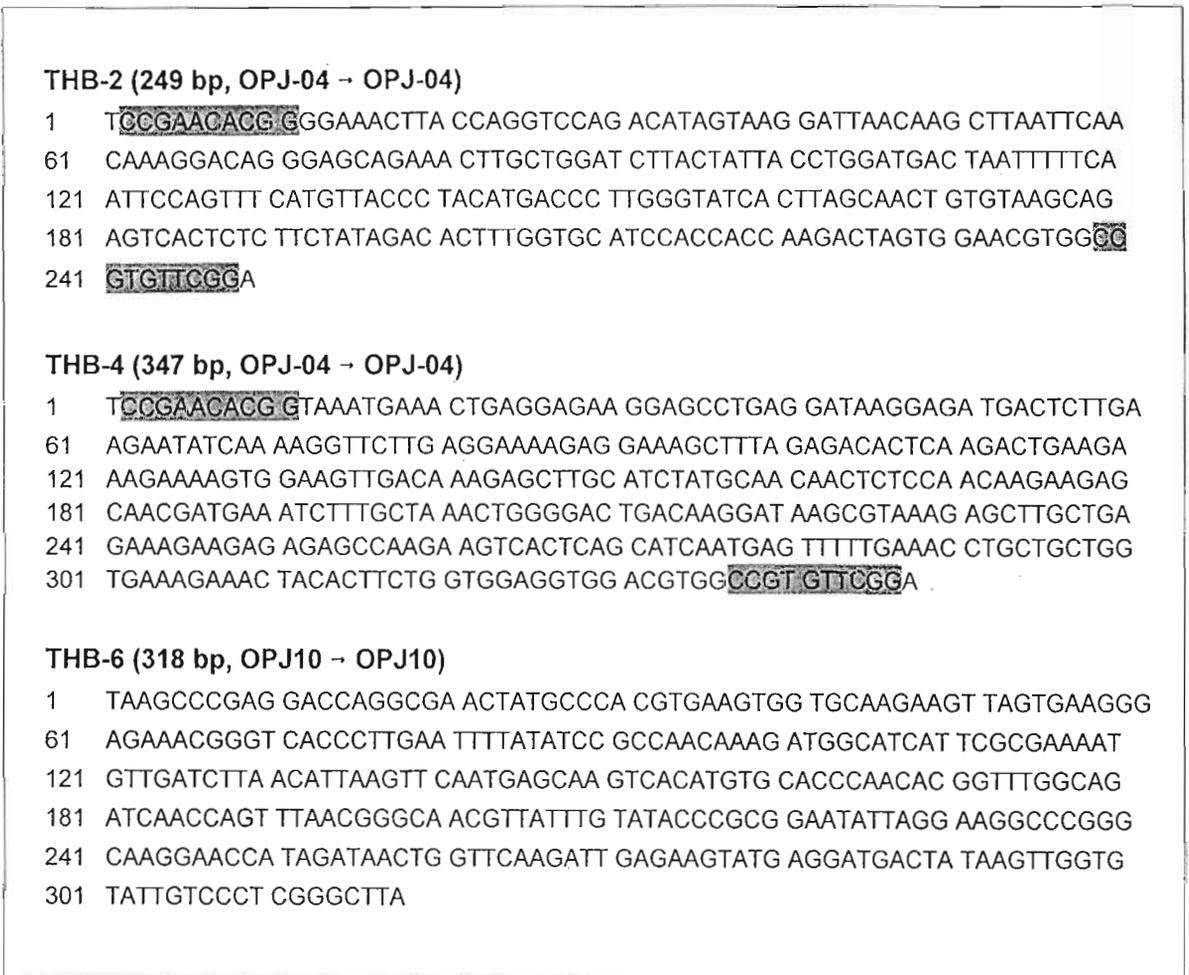


Figure 5.11: Consensus sequences (positive strand) for the three sequence groups which showed some significant homology to known genes. The size of the insert as well as the primers from which it was generated are given in parentheses. Primer sequences are highlighted.

The most interesting result from the homology searches is the similarity between the THB-4 alignment group sequence and a number of plant transcription factors and nuclear RNA binding proteins. As described in Section 1.2.2, the majority of the genes thus far identified as playing a role in genetically-imposed seed dormancy, including *ABI3 / VP1*, *LEC1* and *DAG1*, encode transcription factor-like proteins. These genes appear to exert their effect by either activating or repressing gene expression programs which are instrumental in the normal functioning of seed development or seed germination (HOLDSWORTH, LENTON, FLINTHAM, GALE, KURUP, MCKIBBIN, BAILEY, LARNER and RUSSEL, 2001). The ability of

transcription factors to regulate transcription by RNA polymerase is thought to involve protein-protein interactions (LAM, 1995).

Many RNA binding proteins which associate with nuclear pre-mRNA (heterogenous nuclear RNA or hnRNA) molecules to form heterogenous nuclear ribonucleoprotein particles (hnRNPs) have been identified in plants, yet the functional significance of these proteins is poorly understood and has only been described for a few of these proteins (LANSBERGER, LORKOVIC and OELMÜLLER, 2002). One of the classes of RNA binding proteins consists of small polypeptides of 15-17 kDa, each containing an N-terminal RNA-binding RNA recognition motif (RRM) and a C-terminal RGG (arginine-glycine-glycine) box which is also an RNA-binding motif (LANSBERGER, LORKOVIC and OELMÜLLER, 2002). The first of these small RNA-binding proteins to be identified was expressed in response to an ABA stimulus and water stress in *Zea mays* (GÓMEZ, SÀNCHEZ-MARTINEZ, STIEFEL, RIGAU, PUIGDOMÈNECH and PAGÈS, 1988). Whilst the function of these proteins has thus far not been elucidated, it is interesting that at least one member of this class of proteins is expressed under stress conditions, as thermoinhibited achenes are undoubtedly in a state of heat stress. Furthermore, of the ten thermoinhibition-specific polypeptides identified during this study, three had molecular weights of approximately 14-15 kDa and thus fall within the size range of this class of protein (Section 3.3.5). Although functions for most RNA-binding proteins have yet to be determined, various RNA-binding proteins have been reported to be developmentally expressed or to have regulated as well as tissue or organ specificity (LANSBERGER, LORKOVIC and OELMÜLLER, 2002). This regulated expression may reflect different functions in post-transcriptional regulation of tissue- or cell-specific gene expression (LANSBERGER, LORKOVIC and OELMÜLLER, 2002). As RNA-binding proteins show high-affinity sequence-specific RNA binding, it is believed that they could influence the binding of other trans-acting factors to these mRNAs, thereby functioning in a manner analogous to transcription factors (KILEDJIAN, BURD, GÖRLACH, PORTMAN and DREYFUSS, 1994). These hnRNPs may either recruit other proteins to the binding site via protein-protein interactions, or they may prevent the binding of these factors (KILEDJIAN, BURD, GÖRLACH, PORTMAN and

DREYFUSS, 1994). Several of these hnRNP proteins may also have an affinity for specific DNA sequences and may therefore play a role in transcription and DNA metabolism (KILEDJIAN, BURD, GÖRLACH, PORTMAN and DREYFUSS, 1994).

Although the sequences used in the homology search are relatively short so that the identification of these sequences must be viewed as tentative, homology of the sequences from this alignment group to such genes gives further credence to the hypothesis that thermoinhibition in this species is an actively-controlled process, rather than a passive process where germination is prevented by environmentally-induced changes in seed ultrastructure. These results suggest two tentative hypotheses for the imposition of thermoinhibition in *T. minuta*. Firstly, it may be hypothesized that thermoinhibition is controlled by one or several transcription-factor like proteins, or hnRNP-like proteins acting in the manner of transcription factors, which activate germination-repressive genes and / or directly repress the expression of a gene or program of genes which is required for germination. Alternatively, certain genes which are critical for the normal progression of germination may be regulated at a post-transcriptional level in an undetermined manner. RNA-binding proteins are involved in a number of post-transcriptional processes, including 3'-end processing, editing of the pre-mRNA (hnRNA) sequences to remove intron sequences, mRNA transport and translation (LANSBERGER, LORKOVIC and OELMÜLLER, 2002). Interference by thermoinhibition-specific RNA-binding proteins which affect normal RNA processing at any of these steps could prevent the expression of germination-associated genes and so interrupt the germination process. The most likely genes to be targeted for repression under thermoinhibitory conditions are arguably those relating to the digestion of storage reserves and to respiration, such that the increase in respiration required to complete the germination phase and especially for seedling establishment is prevented. Indeed, the VP1 protein has been linked to the repression of key glyoxalate cycle genes such as isocitrate lyase and malate synthase, consequently restricting reserve lipid breakdown prior to germination (PAEK, LEE, BAI and SMITH, 1998). This gene and its homologues may thus also play a more general role in repressing the transcription of germination-specific genes (FOOTITT and COHN, 2001). Proteins and lipids

constitute the major storage reserves in *T. minuta* achenes and it has been postulated that these lipids were converted to carbohydrate and stored as starch through the glyoxalate pathway at 35°C (DREWES, 1989). Thus, repression of the glyoxalate cycle does not appear to occur in thermoinhibited *T. minuta* achenes. However, no starch accumulation was observed in germinating achenes, indicating that the carbohydrates released from breakdown of the lipid reserves are rapidly utilised. This suggests that there may be some form of repression of respiratory metabolism in the thermoinhibited achenes. Since thermoinhibited achenes are able to germinate within a few hours of the incubation temperature being reduced below the thermoinhibitory threshold of 35°C, this would make sense as thermoinhibited achenes must be metabolically active to reach the stage where they are primed for rapid radicle emergence in response to a drop in temperature. Thus, energy from their stored reserves must be available to the embryos, but in amounts which prevent the increase in respiration which is required with radicle emergence and the onset of seedling establishment.

A second homology grouping, THB-2, showed sequence homology with genes encoding RNase L inhibitor-like proteins. RNase L is a central pathway of interferon action in mammalian tissues, but may also play a more general role as a regulator of RNA stability (BISBAL, MARTINAND, SILHOL, LEBLEU and SALEHZADA, 1995). As far as can be established, no data has been published on either the existence or possible role of this ribonuclease in plant systems. In mammalian systems, the enzyme inhibits protein synthesis by cleavage of mRNAs at the 3' end of UpNp sequences (BISBAL, MARTINAND, SILHOL, LEBLEU and SALEHZADA, 1995). Interferons are produced in mammalian cells in response to various inducers, including viral infection and induce the expression of a range of genes involved in defence against viral attack, the control of cell proliferation and differentiation and the modulation of immune responses. Variations in RNase L activity have been observed during cell growth and differentiation, even in the absence of interferon treatment (BISBAL, MARTIAND, SILHOL, LEBLEU and SALEHZADA, 1995). In mammalian systems, the control of the turnover rate of mRNAs is a critical element in the regulation of gene expression (BISBAL, MARTINAND, SILHOL, LEBLEU and

SALEHZADA, 1995). RNase L may be involved in the control of RNA metabolism by regulating mRNA degradation (BISBAL, MARTIAND, SILHOL, LEBLEU and SALEHZADA, 1995). The RNase L inhibitor (RLI) protein neutralises this ribonuclease activity and may function to regulate the activity of this enzyme (BISBAL, MARTIAND, SILHOL, LEBLEU and SALEHZADA, 1995). Thermoinhibition is very rapidly imposed and then released in *T. minuta* achenes in response to temperature changes. The thermoinhibition-associated polypeptides identified within this study disappeared rapidly from the achenes in response to a drop in ambient temperature. Levels of these polypeptides decreased noticeably in 2D-PAGE gels within 2 h of a reduction in temperature to 25 °C (Section 3.3.5). This rapid response may be explained by a tight regulation on the turnover of both the mRNAs encoding these polypeptides and the polypeptides themselves. The thermoinhibition-associated mRNA molecules may have a rapid turnover, determined by susceptibility to RNase degradation, which is inhibited under thermoinhibitory conditions by the expression of an RNase inhibitor protein under the regulation of the thermoinhibition-program. When the achenes are exposed to thermoinhibitory temperatures, this inhibitor would protect the thermoinhibition-associated transcripts from degradation by the ribonuclease. When the temperature was reduced and the inhibitor gene no longer expressed, a decline in levels of the inhibitor would allow for a resumption in RNase activity, rapidly destroying all of the thermoinhibition associated transcripts and allowing for a resumption in germination.

The sequence homology of the THB-6 sequence with the gene for miraculin is difficult to explain. Miraculin is a taste-modifying protein, causing sour tastes to be perceived as sweet, which was first isolated from berries of the west African shrub *Richadella dulcifera* by THEERASILP and KURIHARA (1988). THEERASILP, HITOTSUYA, NAKAJO, NAKAYA, NAKAMURA and KURIHARA (1989) determined that the protein contains 191 amino acids and has a deduced relative molecular mass of 24.6 kDa. MASUDA, NIRASAWA, NAKAYA and KURIHARA (1995) deduced a similar sequence which differed from the previous sequence by only a single amino acid, but which also appeared to contain a further 29 amino acid sequence which may have encoded a signal sequence. The predicted size for this protein does correlate well

with the size of the second group of thermoinhibition-specific polypeptides identified during this study, in which seven polypeptides ranging in size from 22 kDa to 26.7 kDa were observed. Some insight may be gained from the fact that miraculin falls within the class of Kunitz-type protease inhibitors (BRENNER, LAMBERT, KALOSHIAN and WILLIAMSON, 1998), as well as the fact that some homology was obtained with other protease inhibitors such as one identified from *Theobroma cacao* (SPENCER and HODGE, 1991). Protease inhibitors of several classes are found in a variety of seeds. These have been suggested to have a defensive function as well as a possible storage function (SPENCER and HODGE, 1991). Kunitz protease inhibitors fall within a family of inhibitors which also inhibit α -amylase (SPENCER and HODGE, 1991). This may suggest a possible role for this gene in the prevention of starch hydrolysis under thermoinhibitory conditions, which might correlate with the increase in starch deposition in thermoinhibited *T. minuta* achenes observed by DREWES (1989).

As previously mentioned, the cDNA sequences identified in this study are too short to give definitive sequence data for the identification of these genes by homology searches. Whilst this study has shown that thermoinhibition in *T. minuta* achenes shows a number of characteristics suggestive of an active, genetically-based control over germination at supra-optimal temperatures, much work remains to be done. The full length sequences for these genes need to be isolated from either genomic or cDNA libraries. This should be relatively simple to accomplish as the clones isolated during this study may be used as probes for these sequences. Once the full length sequences have been identified, conclusive identification of the genes should be possible. This would enable a much more definitive hypothesis for the basis of thermoinhibition to be determined. This hypothesis may then be tested by transgenic studies, possibly using these genes in the antisense direction to eliminate the expression of these genes to determine whether or not prevention of their expression would be able to overcome thermoinhibition. A likely system for such testing would be through the transformation of somatic embryos. It would indeed be interesting to determine whether somatic embryos of *T. minuta* behave similarly to zygotic embryos in response to temperature changes. Should this be the case, an embryogenic callus

culture would form an excellent system for transgenic studies of the effects of these genes. A protocol for the production of somatic embryos of *Tagetes erecta* L. has been published (BESPALHOK and HATTORI, 1998) and should provide a sound basis for the establishment of a protocol for somatic embryo formation in *T. minuta*. In addition to the cDNAs described in this Chapter, a number of other cDNAs which are differentially expressed in thermoinhibited achenes have been isolated which have yet to be cloned and analysed. It is also possible that thermoinhibition involves a number of novel genes which would not be identified by homology searches. These may include some of the cDNAs analysed in this study as well as the many which have yet to be cloned and analysed. In such cases, a transgenic system which would allow for an analysis of the effects of these genes may be the only way in which these genes may be functionally correlated with thermoinhibition.

5.4 CONCLUSION

Following differential display, thirteen differentially expressed cDNAs were cloned into the vector pGEM[®]-EASY and transformed into *Escherichia coli*. Three clones for each cDNA were selected for sequencing. All 39 clones were shown to be expressed only in thermoinhibited achenes by reverse northern blotting. Following sequencing, 27 clones produced readable sequences. The sequences of these cDNAs were analysed and similar sequences were combined into alignment groups representing identical sequences. The insert sequences were compared to known gene and protein sequences using the BLAST tools, which resulted in tentative identification of three sequence groups as a possible transcription factor-like or RNA-binding protein-like gene, an RNase L-inhibitor-like gene and the gene encoding the taste modifying protein miraculin. These results should be viewed with caution until such time as full length transcripts of the relevant cDNAs may be obtained and sequenced. However, based on these preliminary results, it may be hypothesized that thermoinhibition in *T. minuta* may be regulated at the transcriptional level, the post-transcriptional level or at both levels. At the transcriptional level, some of the differentially expressed cDNAs may encode proteins which act like transcription

factors and which actively impose thermoinhibition by causing the expression of germination-repressive genes or by directly preventing the expression of specific germination-associated genes. At the post-transcriptional level, these cDNAs may represent hnRNP proteins which modify the processing of the mRNA transcripts of genes required for germination, preventing their normal expression. Another of the cDNA alignment groups may regulate the longevity of thermoinhibition-associated mRNAs, ensuring tight regulation of the expression of the thermoinhibition-associated genes. The third group of cDNAs appears to encode a possible protease-inhibitor.

5.5 GENERAL CONCLUSION

All previous reports on thermoinhibited species in the literature have indicated that for those species, thermoinhibition is a passive process. Germination is prevented by temperature-dependant changes in morphology or metabolism or, in some instances, in both. None of these phenomena can be related to thermoinhibition in *T. minuta*. This study, however, from molecular analyses of both proteins and nucleic acids, has revealed that a number of genes are specifically expressed in thermoinhibited achenes. The behaviour of the thermoinhibition-associated proteins identified in this study by two-dimensional polyacrylamide gel electrophoresis follows a pattern similar to that observed for species in which seed dormancy is genetically-imposed. These polypeptides are only synthesized in imbibed, thermoinhibited achenes and not in dry achenes held at the thermoinhibitory temperatures. Their expression is exceptionally tightly linked to the onset and release of thermoinhibition, and they are completely specific to the thermoinhibited condition, not being found at any stage in germinating achenes.

These observations, as well as the tentative identifications of the cDNAs as described in this Chapter, are consistent with the hypothesis that thermoinhibition in *T. minuta* is an actively-controlled process which is imposed on imbibing achenes when the ambient temperature is not conducive to seedling establishment.

The identification of genes which are critical for germination has long eluded seed biologists. To date, the major tool available for the isolation of these genes was through comparative studies of dormant versus non-dormant seeds. Once those genes which prevent germination under adverse conditions have been identified, a study of their functions and interactions with other genes expressed during normal germination should lead to the identification of the genes which represent the critical steps in germination. Such dormancy studies, however, are made more difficult by the fact that various treatments have to be applied to the dormant seeds before dormancy is broken. Furthermore, dormancy is relatively slowly imposed or released. In *T. minuta*, however, thermoinhibition also appears to be genetically controlled in a manner analogous to certain forms of dormancy. Unlike dormancy systems, however, with thermoinhibited achenes the only treatment needed for germination to continue is a reduction of only a few degrees in the temperature at which the achenes are incubated. This change from the thermoinhibited state to a germinable state is exceptionally rapid, taking no more than a couple of hours. This allows for simple analysis of the behaviour of thermoinhibition-associated mRNAs and proteins, as time courses of their expression and disappearance can easily be plotted relative to the germination of the achenes. This therefore represents a powerful new model system for the study of the molecular biology of seed germination. It is hoped that the findings and ideas presented in this thesis will provide a foundation for further studies which will lead to greater understanding of the molecular biology of not only thermoinhibition in this species, but on a broader scale to the molecular control processes involved in the transition from a quiescent seed to a seedling.

LITERATURE CITED

- AGRAWAL, P.K., K.L. SETHI and R.B. MEHRA. 1980. Effect of temperature on germination of *Lathyrus sativus* seeds: An analysis. *Seed Research* **8**: 25-32
- ALTSCHUL, S.F., W. GISH, W.MILLER, E.W. MYERS and D.J. LIPMAN. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403-410
- ALTSCHUL, S.F., T.L. MADDEN, A.A. SCHÄFFER, J. ZHANG, Z. ZHANG, W. MILLER and D.J. LIPMAN. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389-3402
- AUSUBEL, F.M., R. BRENT, R.E. KINGSTON, D.D. MOORE, J.G. SEIDMAN, J.A. SMITH and K. STRUHL. 1989. *Short Protocols in Molecular Biology*. J. Wiley and Sons, New York. pp 1.4.1-1.4.9, 1.8.1-1.8.8, 6.1.2-6.1.3
- BAILEY, P.C., G.W. LYCETT and J.A. ROBERTS. 1996. A molecular study of dormancy breaking and germination in seeds of *Trollius ledebouri*. *Plant Molecular Biology* **32**:559-564
- BANKIER, A.T. and B.G. BARRELL. 1989. Sequencing single-stranded DNA using the chain termination method. In: *Nucleic Acids Sequencing: A practical approach*. C.J. Howe and E.S. Ward (eds.). IRL Press, Oxford. pp 37-78
- BAUER, D., H. MÜLLER, J. REICH, H. RIEDEL, V. AHRENKIEL, P. WARTHOE and M. STRAUSS. 1993. Identification of differentially expressed mRNA species by an improved display technique (DDRT-PCR). *Nucleic Acids Research* **21**: 4272-4280
- BÄUMLEIN, H., S. MISERA, H. LUERSSSEN, K. KÖLLE, C. HORSTMANN, U. WOBUS and A.J. MULLER. 1994. The *FUS3* gene of *Arabidopsis thaliana* is a regulator of gene expression during late embryogenesis. *The Plant Journal* **6**: 379-387
- BAUW, G., M. VAN MONTAGU and D. INZE. 1992. Microsequence analysis of *Arabidopsis* proteins separated by two-dimensional polyacrylamide gel electrophoresis: A direct linkage of proteins and genes. In: *Methods in Arabidopsis Research*. C. Koncz, N-H. Chua and J. Schell (eds.). World Scientific Publishing Company, Singapore. pp 357-377
- BEIS, A. and A. LAZOU. 1990. Removal of artifactual bands associated with the presence of 2-mercaptoethanol in two-dimensional polyacrylamide gel electrophoresis. *Analytical Biochemistry* **190**: 57-59
- BERTIOLI, D.J., U.H.A. SCHLICHTER, M.J. ADAMS, P.R. BURROWS, H-H. STEINBIß and J.F. ANTONIW. 1995. An analysis of differential display

shows a strong bias towards high copy number mRNAs. *Nucleic Acids Research* **23**: 4520-4523

- BESPALHOK, J.C. and F.K. HATTORI. 1998. Friable embryogenic callus and somatic embryo formation from cotyledon explants of African marigold (*Tagetes erecta* L.). *Plant Cell Reports* **17**: 870-875
- BEWLEY, J.D. 1997. Seed dormancy and germination. *Plant Cell* **9**: 1055-1066
- BEWLEY, J.D. and M. BLACK. 1982. *Physiology and Biochemistry of Seeds in Relation to Germination. Volume 2*. Springer Verlag, Berlin. pp 297, 304
- BEWLEY, J.D. and M. BLACK. 1986. *Seeds: Physiology of development and germination. Volume 2*. Springer-Verlag, Berlin. pp 297, 304
- BEWLEY, J.D. and M. BLACK. 1994. *Seeds: Physiology of development and germination*. Plenum Press, New York.
- BEWLEY, J.D. and A. MARCUS. 1990. Gene expression in seed development and germination. *Progress in Nucleic Acid Research and Molecular Biology* **38**: 165-193
- BIDDINGTON, N.L. and T.H. THOMAS. 1978. Thermodormancy in celery seeds and its removal by cytokinins and gibberellins. *Physiologia Plantarum* **42**: 401-405
- BISBAL, C., C. MARTINAND, M. SILHOL, B. LEBLEU and T. SALEHZADA. 1995. Cloning and characterization of a RNase L inhibitor: A new component of the interferon-regulated 2-5A pathway. *Journal of Biological Chemistry* **270**: 13308-13317
- BOTHA, F.C., G.P. POTGIETER and A.-M. BOTHA. 1992. Respiratory metabolism and gene expression during seed germination. *Plant Growth Regulation* **11**: 211-224
- BRADBEER, J.W. 1988. *Seed Dormancy and Germination*. Blackie, London. pp 1-3, 6
- BRADFORD, K.J. and O.A. SOMASCO. 1994. Water relations of lettuce seed thermoinhibition. I. Priming and endosperm effects on base water potential. *Seed Science Research* **4**: 1-10
- BRADFORD, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**: 248-254
- BRENNER, E.D., K.N. LAMBERT, I. KALOSHIAN and V.M. WILLIAMSON. 1998.

- Characterization of *LeMir*, a root-knot nematode-induced gene in tomato with an encoded product secreted from the root. *Plant Physiology* **118**: 237-247
- BROWN, T.A. 1990. *Gene Cloning: an introduction. Second edition.* Chapman and Hall, London. pp 90-93
- CALLARD, D., B. LESCURE and L. MAZZOLINI. 1994. A method for the elimination of false positives generated by the mRNA differential display technique. *Biotechniques* **16**: 1096-1103
- CHEN, E.Y. 1994. The efficiency of automated sequencing. In: *Automated DNA Sequencing and Analysis.* M.D. Adams, C. Fields and J.C. Venter (eds.). Academic Press, London. pp 3-10
- CHOMCZYNSKI, P. and N. SACCHI. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry* **162**: 156
- COMAI, L. and J.J. HARADA. 1990. Transcriptional activities in dry seed nuclei indicate the timing of the transition from embryogeny to germination. *Proceedings of the National Academy of Sciences of the USA* **87**: 2671-2674
- CÔME, D. and T. TISSAOUI. 1973. Interrelated effects of imbibition, temperature and oxygen on seed germination. In: *Seed Ecology.* W. Heydecker (ed.). Butterworths, London. pp 157-168
- COPELAND, R.A. 1994. *Methods for Protein Analysis: A practical guide to laboratory protocols.* Chapman and Hall, New York. pp 59-62
- CORBINEAU, F., R.M. RUDNICKI and D. CÔME. 1989. ACC conversion to ethylene by sunflower seeds in relation to maturation, germination and thermodormancy. *Plant Growth Regulation* **8**: 105-115
- DATTA, K., L. MARSH and A. MARCUS. 1983. Early growth of wheat embryonic axes and the synthesis of RNA and DNA. *Plant Physiology* **72**: 394-397
- DEBOUCK, C. 1995. Differential display or differential dismay? *Current opinion in Biotechnology* **6**: 597-599
- DE KLERK, G.J. 1981. Degradation of early synthesized proteins in imbibed dormant and after-ripened *Agrostemma githago* L. embryos. *Developmental Biology* **83**: 183-187
- DRENNAN, P.M. and J. VAN STADEN. 1989. Enhancement of emergence in *Tagetes minuta* by light and temperature pre-treatment. *Seed Science and Technology* **17**: 115-124
- DREW, R.L.K. and P.A. BROCKLEHURST. 1984. Investigations on the control of

- lettuce seed germination at high temperatures. *Journal of Experimental Botany* **35**: 986-993
- DREW, R.L.K. and P.A. BROCKLEHURST. 1987. The effect of sodium hypochlorite on thermo-inhibition of lettuce seed. *Acta Horticulturae* **215**: 179-184
- DREWES, F.E. 1989. *An examination of the thermoinhibition of Tagetes minuta L. achenes*. MSc. Thesis, University of Natal (Pietermaritzburg). Unpublished.
- DREWES, F.E. and J. VAN STADEN. 1990. Germination of *Tagetes minuta* L. II. Role of gibberellins. *Plant Growth Regulation* **9**: 285-291
- DUNLAP, J.R., B.T. SCULLY and D.M. REYES. 1990. Seed coat-mediation of lettuce germination responses to heat and sodium chloride. *Journal of the Rio Grande Valley Horticultural Society* **43**: 55-61
- DUTTA, S. and K.J. BRADFORD. 1994. Water relations of lettuce seed thermoinhibition. II. Ethylene and endosperm effects on base water potential. *Seed Science Research* **4**: 11-18
- DUTTA, S. K.J. BRADFORD and D.J. NEVINS. 1994. Cell-wall autohydrolysis in isolated endosperms of lettuce (*Lactuca sativa* L.). *Plant Physiology* **104**: 623-628
- DYER, W.E. 1993. Dormancy-associated embryonic mRNAs and proteins in imbibing *Avena fatua* embryos. *Physiologia Plantarum* **88**: 201-211
- EGLEY, G.H. 1995. Seed germination in soil: Dormancy cycles. In: *Seed Development and Germination*. J. Kigel and G. Galili (eds.). Marcel Dekker, New York. pp 529-544
- ESAU, K. 1977. *Anatomy of seed plants. Second edition*. John Wiley and Sons, Singapore. p 435
- FARRELL, R.E. Jr. 1993. *RNA Methodologies: A laboratory guide for isolation and characterization*. Academic Press, London. pp 57-69
- FELLICIELLO, I and G. CHINALI. 1993. A modified alkaline lysis method for the preparation of highly purified plasmid DNA from *Escherichia coli*. *Analytical Biochemistry* **212**: 394-401
- FOOTITT, S. and M.A. COHN. 2001. Developmental arrest: from sea urchins to seeds. *Seed Science Research* **11**: 3-16
- FORSYTH, C. and J. VAN STADEN. 1983. Germination of *Tagetes minuta* L. I. Temperature effects. *Annals of Botany* **52**: 659-666
- FUKUHARA, T. and H.J. BOHNERT. 2000. The expression of a Vp1-like gene and

seed dormancy in *Mesembryanthemum crystallinum*. *Genes and Genetic Systems* **75**: 203-209

- GALAU, G.A., K.S. JAKOBSEN and D.W. HUGHES. 1991. The controls of late dicot embryogenesis and early germination. *Physiologia Plantarum* **81**: 280-288
- GALLARDO, M., M. DEL MAR DELGARDO, I. SANCHEZ-CALLE and A.J. MATILLA. 1991. Ethylene production and 1-aminocyclopropane-1-carboxylic acid conjugation in thermoinhibited *Cicer arietinum* L. seeds. *Plant Physiology* **97**: 122-127
- GALLARDO, M., I. SANCHEZ-CALLE, P. MUÑOZ DE RUEDA and A.J. MATILLA. 1996. Alleviation of thermoinhibition in chickpea seeds by putrescine involves the ethylene pathway. *Australian Journal of Plant Physiology* **23**: 479-487
- GEGENHEIMER, P. 1990. Preparation of extracts from plants. *Methods in Enzymology* **182**: 174-182
- GOLDMARK, P.J., J. CURRY, C.F. MORRIS and M.K. WALKER-SIMMONS. 1992. Cloning and expression of an embryo-specific mRNA up-regulated in hydrated dormant seeds. *Plant Molecular Biology* **19**: 433-441
- GÓMEZ, J., D. SÁNCHEZ-MARTINEZ, V. STIEFEL, J. RIGAU, P. PUIGDOMÈNECH and M. PAGÈS. 1988. A gene induced by the plant hormone abscisic acid in response to water stress encodes a glycine-rich protein. *Nature* **334**: 262-264
- GRABANDT, K. 1985. *Weeds of Crops and Gardens in Southern Africa*. Seal Publishing, Jeppestown. p69
- GREENLAND, A. 1987. Synthesis and cloning of cDNAs from eukaryotic mRNAs. In: *Gene Cloning and Analysis: A laboratory guide*. G.J. Boulnois (ed.). Blackwell Scientific, Oxford. pp 9-36
- GROOT, S.P.C. and C.M. KARSSSEN. 1987. Gibberellins regulate seed germination in tomato by endosperm weakening: A study with gibberellin-deficient mutants. *Planta* **171**: 525-531
- HANAHAH, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology* **166**: 557-580
- HANCE, B.A. and J.M. BEVINGTON. 1992. Changes in protein synthesis during stratification and dormancy release in embryos of sugar maple (*Acer saccharum*). *Physiologia Plantarum* **86**: 365-371
- HECK, G.R. and D.E. FERNANDEZ. 1997. Differential display of mRNA. In: *Differentially Expressed genes in Plants: A bench manual*. E. Hansen and G.

- Harper (eds.). Taylor and Francis, London. pp 45-62
- HEINER, C. and T. HUNKAPILLER. 1989. Automated DNA sequencing. In: *Nucleic Acids Sequencing: A practical approach*. C.J. Howe and E.S. Ward (eds.). IRL Press, Oxford. pp 221-236
- HEYDECKER, W. 1977. Stress and seed germination: An agronomic view. In: *The Physiology and Biochemistry of Seed Dormancy and Germination*. A.A. Khan (ed.). North Holland, Amsterdam. pp 237-282
- HILLIARD, O.M. 1977. *Compositae in Natal*. University of Natal Press, Pietermaritzburg. pp 328-329
- HOLDSWORTH, M., J. LENTON, J. FLINTHAM, M. GALE, S. KURUP, R. McKIBBIN, P. BAILEY, V. LARNER and L. RUSSELL. 2001. Genetic control mechanisms regulating the initiation of germination. *Journal of Plant Physiology* **158**: 439-445
- HOLM, L., J.V. PANCHO, J.P. HERBERGER and D.L. PLUCKNETT. 1979. A *Geographical Atlas of World Weeds*. John Wiley and Sons, New York. p 358
- HOROWITZ, M. and R.B. TAYLORSON. 1983. Effect of high temperatures on imbibition, germination, and thermal death of velvetleaf (*Abutilon theophrasti*) seeds. *Canadian Journal of Botany* **61**: 2269-2276
- HUANG, X.L. and A.A. KHAN. 1992. Alleviation of thermoinhibition in preconditioned lettuce seeds involves ethylene, not polyamine biosynthesis. *Journal of the American Society of Horticultural Science* **117**: 841-845
- INOUE, H., H. NOJIMA and H. OKAYAMA. 1990. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**: 23-28
- JENDRISAK, J. 1980. The use of α -amanitin to inhibit *in vivo* RNA synthesis and germination in wheat embryos. *Journal of Biological Chemistry* **255**: 8529-8533
- JOHNSON, R.R., H.J. CRANSTON, M.E. CHAVERRA and W.E. DYER. 1995. Characterization of cDNA clones for differentially expressed genes in embryos of dormant and nondormant *Avena fatua* L. caryopses. *Plant Molecular Biology* **28**: 113-122
- KARSSSEN, C.M., A. HAIGH, P. VAN DER TOORN and R. WEGES. 1989. Physiological mechanisms involved in seed priming. In: *Recent Advances in the Development and Germination of Seeds*. R.B. Taylorson (ed.). Plenum Press, New York. pp 269-280
- KEITH, K., M. KRAML, N.G. DENGLER and P. McCOURT. 1994. *Fusca3*: A heterochronic mutation affecting late embryo development in *Arabidopsis*.

- KHAN, A.A. and X.L. HUANG. 1988. Synergistic enhancement of ethylene production and germination with kinetin and 1-aminocyclopropane-1-carboxylic acid in lettuce seeds exposed to salinity stress. *Plant Physiology* **87**: 847-852
- KHAN, A.A. and J. PRUSINSKI. 1989. Kinetin enhanced 1-aminocyclopropane-1-carboxylic acid utilization during alleviation of high temperature stress in lettuce seeds. *Plant Physiology* **91**: 733-737
- KILEDJIAN, M., C.G. BURD, M. GÖRLACH, D.S. PORTMAN and G. DREYFUSS. 1994. Structure and function of hnRNP proteins. In: *RNA-Protein Interactions*. K. Nagai and I.W. Mattaj (eds.). pp127-149
- KLEBER-JANKE, T. and K. KRUPINSKA. 1997. Isolation of cDNA clones for genes showing enhanced expression in barley leaves during dark-induced senescence as well as during senescence under field conditions. *Planta* **203**: 332-340
- KLEIN, S., H. BARENHOLZ and A. BUDNIK. 1971. The initiation of growth in isolated lima bean axes. Physiological and fine structural effects of actinomycin D, cyclohexamide and chloramphenicol. *Plant and Cell Physiology* **12**: 41-60
- KOORNNEEF, M., G. REULING and C.M. KARSSSEN. 1984. The isolation and characterization of abscisic acid insensitive mutants of *Arabidopsis thaliana*. *Physiologia Plantarum* **61**: 377-383
- KRISTIE, D.N. and A. FIELDING. 1994. Influence of temperature on the Pfr level required for germination in lettuce cv. Grand Rapids. *Seed Science Research* **4**: 19-25
- LAEMMLI, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685
- LAM, E. 1995. Plant gene expression regulation. In: *Molecular Biology and Biotechnology: A comprehensive desk reference*. R.A. Meyers (ed.). VCH Publishers, New York. pp 696-699
- LANDSBERGER, M., Z.J. LORKOVIC and R. OELMÜLLER. 2002. Molecular characterization of nucleus localized RNA-binding proteins from higher plants. *Plant Molecular Biology* **48**: 413-421
- LÉON-KLOOSTERZIEL, K.M., G.A. VAN DE BUNT, J.A.D. ZEEVAART and M. KOORNNEEF. 1996. *Arabidopsis* mutants with a reduced seed dormancy. *Plant Physiology* **110**: 233-240
- LI, B. and M.E. FOLEY. 1995. Cloning and characterization of differentially

expressed genes in imbibed dormant and afterripened *Avena fatua* embryos. *Plant Molecular Biology* **29**: 823-831

- LIANG, P., L. AVERBOUKH and A.B. PARDEE. 1993. Distribution and cloning of eukaryotic mRNAs by means of differential display: refinements and optimization. *Nucleic Acids Research* **21**: 3269-3275
- LIANG, P. and A.B. PARDEE. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**: 967-972
- LINSKENS, M.H.K., J. FENG, W.H. ANDREWS, B.E. ENLOW, S.M. SAATI, L.A. TONKIN, W.D. FUNK and B. VILLEPONTEAU. 1995. Cataloging altered gene expression in young and senescent cells using enhanced differential display. *Nucleic Acids Research* **23**: 3241-3251
- MCCARTY, D.R., T. HATTORI, C.B. CARSON, V. VASIL, M. LAZAR and I.K. VASIL. 1991. The *Viviparous-1* developmental gene of maize encodes a novel transcriptional activator. *Cell* **66**: 895-905
- MAHOU, A. and F.G. DENNIS Jr. 1994. Protein changes in peach seeds during chilling are not associated with breaking dormancy. *Journal of the American Society for Horticultural Science* **119**: 131-135
- MALAN, C., J.H. VISSER and N. GROBBELAAR. 1981. Control of problem weeds of maize on the Transvaal Highveld (South Africa). I. *Tagetes minuta* L. *Weed Research* **21**: 235-241
- MANDEL, M. and A. HIGA. 1970. Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology* **53**: 159-162
- MANIATIS, T., E.F. FRITSCH and J. SAMBROOK. 1982. *Molecular Cloning: A laboratory manual*. Cold Spring Harbour Laboratory Press, New York. pp 112-113.
- MASUDA, Y., S. NIRASAWA, K. NAKAYA and Y. KURIHARA. 1995. Cloning and sequencing of a cDNA encoding a taste-modifying protein, miraculin. *Gene* **161**: 175-177
- MAXAM, A.M. and W. GILBERT. 1977. A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the USA* **74**: 560-564
- MAYER, J.E., G. HAHNE, K. PALME and J. SCHELL. 1987. A simple and general tissue extraction procedure for two-dimensional gel electrophoresis. *Plant Cell Reports* **6**: 77-81
- MAYER, A.M. and A. POLJAKOFF-MAYBER. 1975. *The Germination of Seeds. Volume 2*. Pergamon, Oxford. pp 32-35

- MEINKE, D.W., L.H. FRANZMANN, T.C. NICKLE and E.C. YEUNG. 1994. Leafy cotyledon mutants of *Arabidopsis*. *Plant Cell* **6**: 1049-1064
- MEISSNER, R., P.C. NEL and E.A. BEYERS. 1986. Allelopathic influence of *Tagetes*- and *Bidens*-infested soils on seedling growth of certain crop species. *South African Journal of Plant and Soil* **3**: 176-180
- MEZEL, L.M. and D.R. STORTS. 1994. Purification of PCR products. In: *PCR Technology: Current innovations*. H.G. Griffin and A.M. Griffin (eds.). CRC Press, Boca Raton. p 21
- MORRIS, C.F., R.J. ANDERBERG, P.J. GOLDMARK and M.K. WALKER-SIMMONS. 1991. Molecular cloning and expression of abscisic acid-responsive genes in embryos of dormant wheat seeds. *Plant Physiology* **95**: 814-821
- MORRISSEY, J.H. 1981. Silver stain for proteins in polyacrylamide gels: A modified procedure with enhanced uniform sensitivity. *Analytical Biochemistry* **117**: 307-310
- MOU, L., H. MILLER, J. LI, E. WANG and L. CHALIFOUR. 1994. Improvements to the differential display method for gene analysis. *Biochemical and Biophysical Research Communications* **199**: 564-569
- MUÑOZ DE RUEDA, P., M. GALLARDO, I. SANCHEZ-CALLE, P. and A.J. MATILLA. 1994. Germination of chickpea seeds in relation to manipulation of the ethylene pathway and polyamine biosynthesis by inhibitors. *Plant Science* **97**: 31-37
- OBROUCHEVA, N.V. and O.V. ANTIPOVA. 1997. Physiology of the initiation of seed germination. *Russian Journal of Plant Physiology* **44**: 250-264
- O'FARRELL, P.H. 1975. High resolution two-dimensional electrophoresis of proteins. *Journal of Biological Chemistry* **250**: 4007-4021
- OLD, R.W. and S.B. PRIMROSE. 1987. *Principles of Gene Manipulation: An introduction to genetic engineering. Third edition*. Blackwell Scientific Publications, Oxford. pp 10-12
- PAEK, N.C., B.-M. LEE, D.G. BAI and J.D. SMITH. 1998. Inhibition of germination gene expression by *Viviparous-1* and ABA during maize kernel development. *Molecules and Cells* **8**: 336-342
- PAPI, M., S. SABATINI, D. BOUCHEZ, C. CAMILLERI, P. CONSTANTINO and P. VITTORIOSO. 2000. Identification and disruption of an *Arabidopsis* zinc finger gene controlling seed germination. *Genes and Development* **14**: 28-33
- PIROVANO, L., S. MORGUTTI, L. ESPEN and S.M. COCUCCHI. 1997. Differences

- in the reactivation process in thermosensitive seeds of *Phacelia tanacetifolia* with germination inhibited by high temperature (30°C). *Physiologia Plantarum* **99**: 211-220
- POLLARD, J.W. 1994. Two-dimensional polyacrylamide gel electrophoresis of proteins. In: *Methods in Molecular Biology, Vol. 32: Basic protein and peptide protocols*. J.M. Walker (ed.). Humana Press, Totowa. pp 73-85
- PONS, T.L. 1986. Responses of *Plantago major* seeds to the red/far-red ratio as influenced by other environmental factors. *Physiologia Plantarum* **68**: 252-258
- PONS, T.L. 1992. Seed responses to light. In: *Seeds: The ecology of regeneration in plant communities*. M. Fenner (ed.). CAB International, Wallingford. pp 259-284
- PRUSINSKI, J. and A.A. KHAN. 1990. Relationship of ethylene production to stress alleviation in seeds of lettuce cultivars. *Journal of the American Society of Horticultural Science* **115**: 294-298
- ROBICHAUD, C. and I.M. SUSSEX. 1986. The response of *viviparous-1* and wild type embryos of *Zea mays* to culture in the presence of abscisic acid. *Journal of Plant Physiology* **126**: 235-242
- RODRÍGUEZ, E. and T.J. MABRY. 1977. Tageteae-chemical revue. In: *The Biology and Chemistry of the Compositae. Volume II*. V.H. Heywood, J.B. Harborne and B.L. Turner (eds.). Academic Press, London. pp 785-789
- RUSSELL, L., V. LARNER, S. KURUP, S. BOUGOURD and M.J. HOLDSWORTH. 2000. The *Arabidopsis* COMATOSE locus regulates germination potential. *Development* **127**: 3759-3767
- SAINI, H.S., P.K. BASSI, E.D. CONSOLACION and M.S. SPENCER. 1986. Interactions among plant hormones, carbon dioxide, and light in the relief of thermoinhibition of lettuce seed germination: Studies in a flow-through gaseous system. *Canadian Journal of Botany* **64**: 2322-2326
- SAINI, H.S., E.D. CONSOLACION, P.K. BASSI, and M.S. SPENCER. 1989. Control processes in the induction and relief of thermoinhibition in lettuce seed germination. *Plant Physiology* **90**: 311-315
- SAMBROOK, J., E.F. FRITSCH and T. MANIATIS. 1989. *Molecular Cloning: A laboratory manual. Second edition*. Cold Spring Harbour Laboratory Press, New York. p 2.57
- SANGER, F., S. NICKLEN and A.R. COULSON. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the USA* **74**: 5463-5467

- SCHÄGGER, H. and G. VON JAGOW. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry* **166**: 368-379
- SHIRZADEGAN, M., P. CHRISTIE and J.R. SEEMANN. 1991. An efficient method for isolation of RNA from tissue cultured plant cells. *Nucleic Acids Research* **19**: 6055
- SIMMONDS, J.A. and G.M. SIMPSON. 1971. Increased participation of pentose phosphate pathway in response to afterripening and gibberellic acid treatment in caryopses of *Avena fatua*. *Canadian Journal of Botany* **49**: 1833-1840
- SMALL, J.G.C., C SCHULTZ and E. CRONJE. 1993. Relief of thermoinhibition in Grand Rapids lettuce seeds by oxygen plus kinetin and their effects on respiration, content of ethanol and ATP and synthesis of ethylene. *Seed Science Research* **3**: 129-135
- SMITH, B.BJ. 1994. SDS polyacrylamide gel electrophoresis of proteins. In: *Methods in Molecular Biology, Vol. 32: Basic protein and peptide protocols*. J.M. Walker (ed.). Humana Press, Totowa. pp 23-34
- SOCKOLOV, B.P. and D.J. PROCKOP. 1994. A rapid and simple PCR-based method for isolation of cDNAs from differentially expressed genes. *Nucleic Acids Research* **22**: 4009-4015
- SOMPAYRAC, L., S. JANE, T.C. BURN, D.G. TENEN and K.J. DANNA. 1995. Overcoming limitations of the mRNA differential display technique. *Nucleic Acids Research* **23**: 4738-4739
- SOULE, J.A. 1993. *Tagetes minuta*: a potential new herb from South America. In: *New Crops. Proceedings of the Second National Symposium: New Crops, exploration, research and commercialization. Indianapolis, Indiana, October 6-9, 1991*. J. Janick and J.E. Simmonds (eds.). John Wiley and Sons, New York. pp 649-654
- SPENCER, M.E. and R. HODGE. 1991. Cloning and sequencing of the cDNA encoding the major albumin of *Theobroma cacao*. Identification of the protein as a member of the Kunitz protease inhibitor family. *Planta* **183**: 528-535
- SPONSEL, V.M. 1983. The localization, metabolism and biological activity of gibberellins in maturing and germinating seeds of *Pisum sativum*, cv. Progress N9. *Planta* **159**: 454-468
- TAKEBA, G. and S. MATSUBARA. 1976. Analysis of temperature effects on the germination of New York lettuce seeds. *Plant and Cell Physiology* **17**: 91-101
- TASHEVA, B. and G. DESSEV. 1983. Artifacts in sodium dodecyl sulfate-polyacrylamide gel electrophoresis due to 2-mercaptoethanol. *Analytical*

- THEERASILP, S. and Y. KURIHARA. 1988. Complete purification and modification of the taste-modifying protein, miraculin, from miracle fruit. *Journal of Biological Chemistry* **263**: 11536-11539
- THEERASILP, S., H. HITOTSUYA, S. NAKAJO, K. NAKAYA, Y. NAKAMURA and Y. KURIHARA. 1989. Complete amino acid sequence and structure of the taste-modifying protein, miraculin. *Journal of Biological Chemistry* **264**: 6655-6659
- THOMAS, T.L. 1993. Gene expression during plant embryogenesis and germination: An overview. *Plant Cell* **5**: 1401-1410
- THOMPSON, J.D., T.J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN and D.G. HIGGINS. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic acids Research* **24**: 4876-4882
- ULLMAN, A., F. JACOB and J. MANOD. 1967. Characterization by *in vitro* complementation of a peptide corresponding to an operator-proximal segment of the β -galactosidase structural gene of *Escherichia coli*. *Journal of Molecular Biology* **24**: 339-343
- VAN STADEN, J. 1973. Changes in endogenous cytokinins of lettuce seed during germination. *Physiologia Plantarum* **28**: 222-227
- VAN STADEN, J. and P.F. WAREING. 1972. The effect of light on endogenous cytokinin levels in seeds of *Rumex obtusifolius*. *Planta* **104**: 126-133
- VERWOERD, T.C., B.M.M. DEKKER and A. HOEKEMA. 1989. A small-scale procedure for the rapid isolation of plant RNAs. *Nucleic Acids Research* **17**: 2362
- VIDAVER, W. and A.I. HSIAO. 1975. Secondary dormancy in light-sensitive lettuce seeds incubated anaerobically or at elevated temperature. *Canadian Journal of Botany* **53**: 2557-2560
- WALTON, D.C. and G.S. SOOFI. 1969. Germination of *Phaseolus vulgaris* III. The role of nucleic acid and protein synthesis in the initiation of axis elongation. *Plant and Cell Physiology* **10**: 307-315
- WAN, J.S., S.J. SHARP, G.M.-C. POIRIER, P.C. WAGAMAN, J. CHAMBERS, J. PYATI, Y.-L. HOM, J.E. GALINDO, A. HUVAR, P.A. PETERSON, M.R. JACKSON and M.G. ERLANDER. 1996. Cloning differentially expressed mRNAs. *Nature Biotechnology* **14**: 1685-1691
- WANG, C-H. and L.O. VODKIN. 1994. Extraction of RNA from tissues containing

- high levels of procyanidin that bind RNA. *Plant Molecular Reports* **12**: 132-145
- WANG, X. and G.Z. FEUERSTEIN. 1995. Direct sequencing of DNA isolated from mRNA differential display. *BioTechniques* **18**: 498-451
- WATSON, J.D. 1977. *Molecular Biology of the Gene. Third edition.* W.A. Benjamin, Taipei. pp 384-390
- WEST, M.A.L., K.M. YEE, J. DANAQ, J.L. ZIMMERMAN, R.L. FISHER, R.B. GOLDBERG and J.J. HARADA. 1994. *Leafy Cotyledon1* is an essential regulator of late embryogenesis and cotyledon identity in *Arabidopsis*. *Plant Cell* **6**: 1731-1745
- WOBUS, U. and H. WEBER. 1999. Seed maturation: genetic programmes and control signals. *Current Opinion in Plant Biology* **2**: 33-38
- YANG, S.F. 1985. Biosynthesis and action of ethylene. *HortScience* **20**: 41-45
- YOSHIOKA, T., T. ENDO and S. SATOH. 1998. Restoration of seed germination at supraoptimal temperatures by fluoridone, an inhibitor of abscisic acid biosynthesis. *Plant Cell Physiology* **39**: 307-312
- ZHAO, S., S.L. OOI and A.B. PARDEE. 1995. New primer strategy improves precision of differential display. *BioTechniques* **18**: 842-850

APPENDIX A

MEDIA FOR BACTERIAL GROWTH AND STORAGE

All quantities provided are those required for the preparation of 1 l of medium. All media are solidified by the addition of 15 g/l agar.

LB (Luria-Bertani) Medium

10 g tryptone
5 g yeast extract
5 g NaCl

Adjust pH to 7.0 with 1 N NaOH

M9 Minimal Medium (5X Stock)

30 g Na_2HPO_4
15 g KH_2PO_4
5 g NH_4Cl
2.5 g NaCl
15 mg CaCl_2 (optional)

Note: M9 minimal medium stock should only be autoclaved for 15 min. Before use, dilute to 1X using sterile dH_2O and add (per l):

1 ml 1M MgSO
1 ml 1M thiamine-HCl
10 ml 20% sucrose

Do not add supplements before the medium has cooled to below 50°C . For solid medium, dissolve the agar in the dH_2O used to dilute the 5X stock.

MEDIA AND SOLUTIONS FOR TRANSFORMATION OF *ESCHERICHIA COLI*

SOB medium

20 g tryptone
5 g yeast extract
0.584 g NaCl
1.864 g KCl

Dissolve in 980 ml dH₂O. After autoclaving, add 10 ml 1 M MgCl₂ and 10 ml MgSO₄.

TB buffer

10 mM PIPES
15 mM CaCl₂
250 mM KCl

After autoclaving, add filter-sterilized 1 M MnCl₂ to a final concentration of 55 mM.

SOC medium

To 100 ml SOB medium, add 2 ml 1 M glucose.

Glycerol Freezer Store Solution

65% (v/v) glycerol
100 mM MgSO₄
25 mM Tris-Cl (pH 8.0)

SOLUTIONS FOR ALKALINE LYSIS MINIPREP

STE Buffer

0.1 M NaCl
10 mM Tris-Cl (pH 8.0)
1 mM EDTA

Solution I

50 mM glucose
10 mM Tris-Cl (pH 8.0)
1 mM EDTA

Solution II

0.2 M NaOH
1% SDS

4 M potassium acetate-2 M acetic acid

Mix 4 volumes 5 M potassium acetate with 1 volume 10 M (57%) glacial acetic acid.

PREPARATION OF DNASE-FREE RNASE A

DNase-free RNase was prepared as described by FELLICIELLO and CHINALI (1993). RNase A (Roche) was dissolved in water to a concentration of 1.5 mg/ml. To this solution was added 0.9 volumes of 0.2 M HCl. The solution was then transferred into 1.9 ml microfuge tubes in 1 ml aliquots. The tubes were placed in boiling water for 5 min and then immediately chilled on ice. The RNase solution

was neutralised at 4° by adding to each tube, at 2 min intervals and with stirring, 100 µl 0.2 M Tris-Cl (pH 7.6) and 5 aliquots of 50 µl 80 mM NaOH. The final solution thus given contains 1 mg/ml RNase A in 15 mM NaCl-15 mM Tris-Cl (pH 7.6). Samples were stored at -20°C.

BUFFERS FOR AGAROSE GEL ELECTROPHORESIS

All quantities provided are those required to prepare 1 l of buffer or buffer stock.

50X TAE buffer stock

242 g Tris

57 ml glacial acetic acid

100 ml 0.5 M EDTA (pH 8.0)

Dilute to 1X before use.

1X TAE gives a working solution of 0.04 M Tris-acetate/0.002 M EDTA.

10X TBE buffer stock

108 g Tris

55 g boric acid

9.3 g EDTA

Dilute to 0.5X or 1X before use.

1X TBE gives a working solution of 0.089 M Tris-base/0.089 M boric acid/0.002 M EDTA.

10X loading buffer for agarose electrophoresis

50 mM NaOH

1 mM EDTA

2.5% glycerol

0.25% bromophenol blue

SOLUTIONS AND BUFFERS FOR POLYACRYLAMIDE GEL ELECTROPHORESIS

Always wear latex gloves when working with acrylamide powder or solutions. Acrylamide powder should only be worked with in a fume hood while wearing a breathing mask.

PROTEIN GELS

SDS-PAGE (Laemmli-system)

Acrylamide stock (% T = 30 %, % C = 2.7 %)

73 g acrylamide

2 g N'N' methylenebisacrylamide (*bisacrylamide*)

Dissolve in 250 ml deionized distilled water (ddH₂O). Filter through Whatman's No. 1 filter paper to remove insoluble impurities and store in the dark at 4°C (stable for approximately 1 month).

Separating gel buffer stock

1 g SDS

45.5 g Tris

Dissolve in 250 ml ddH₂O, adjust pH to 8.8 with HCl. Store at 4°C (stable for

several months)

Stacking gel buffer stock

1 g SDS

15.1 g Tris

Dissolve in 250 ml ddH₂O, adjust pH to 6.8 with HCl. Store at 4°C (stable for several months). Check the pH of the stacking gel buffer each time before use and adjust with HCl if necessary.

2X Sample Solvent

4.6 % [w/v] SDS

10 % [v/v] 2-mercaptoethanol

0.125 M Tris-Cl (pH 6.8)

0.01 % [w/v] bromophenol blue

20 % [v/v] glycerol

TRICINE-SDS-PAGE (SCHÄGGER and VON JAGOW, 1987)

Acrylamide-*bis*acrylamide stock (% T = 49.5 %, % C = 3 %)

48 % (w/v) acrylamide

1.5 % (w/v) *bis*acrylamide

Anode buffer

0.2 M Tris-Cl, pH 8.9

Cathode buffer

0.1 M Tris-Cl, pH 8.25

0.1 M Tricine

0.1 % (w/v) SDS

Gel buffer

3.0 M Tris-Cl, pH 8.45

0.3 % (w/v) SDS

TWO-DIMENSIONAL POLYACRYLAMIDE GEL ELECTROPHORESIS

IEF equilibration buffer

0.5 M Tris-Cl (pH 6.8)

2.5 % [w/v] SDS

0.1 % [w/v] DTT

10% [v/v] glycerol

SEQUENCING GELS FOR DIFFERENTIAL DISPLAY

40% Acrylamide stock

Dissolve 19.0 g acrylamide powder and 1.0 g *bis*acrylamide in 50 ml ddH₂O and filter through a 0.22 μm filter. Wrap the storage bottle in foil and store at 4°C.

6% Polyacrylamide /8 M urea sequencing gel (150 ml)

72.072 g urea

22.5 ml 40% acrylamide stock

15 ml 10X TBE buffer stock

Add ddH₂O to 150 ml. Degas until solution stops bubbling under vacuum. Add 1.5 ml freshly prepared 10% (w/v) APS and 23 µl TEMED. Pour the gel immediately.

HYBRIDIZATION SOLUTIONS

20X SSC

0.3 M sodium citrate

3 M NaCl

Adjust pH to 7.0

20X SSPE

3.6 M NaCl

0.2 M Na₂HPO₄

0.02M EDTA

5X Denhardt's solution

0.1% BSA fraction V

0.1% FICOL 400 000

0.1% PVP

APPENDIX B

UNEDITED REVERSE SEQUENCES

ATJ4-02/1

GGTGACACTATAGAATCTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTC
GACCTGCAGGCGGCCGCGAATTCAGTAGTGATTCCGAACACGGCCACGTTCCACTAGTCT
TGGTGGTGGATGCACCAAAGTGTCTATAGAAGAGAGTGACTCTGCTTACACAGTTGCTAA
GTGATACCCAAGGGTCATGTAGGGTAACATGAAACTGGAATTGAAAAATTAGTCATCCAG
GTAATAGTAAGATCCAGCAAGTTTCTGCTCCCTGTCCTTTGTTGAATTAAGCTTGTTAAT
CCTTACTATGTCTGGACCTGGTAAGTTTCCCCGTGTTCCGAATCGAATTCCCGCGGCCGC
CATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTCGTATTACA
ATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACCTTA
ATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCG
ATCGCCCTT-CCAACAGTTGCGCAGCCTGAATGGCGAATGGAC-CCCCCTGTA-CGGCGC
ATTA-CGCGG-GGG-GGGG-GGTTACCC-CA-G-GAC-GTTACCT-GCAGGGCCCTA-G-
CCG-TC-TTCC-TTTTTCCCTTCTTT-CCCC-GTT--CGGTTTTCCCG-AAGT-TAA--G
GGGGTCCCTT-AGGTCCCA-TTAGG-TT-CGGCCC-CC-CCC-AAAA-T-TTTAGGGGG-
GGTCCC-AAGGGGCCCCCCCCCT-AAAGGTTTTCCCTTGGGGGG---CC-C-TTTAAA-
-G-CCTT-TTC-AAGGAAAA

ATJ4-02/2

GTGACACTATAGAATACCTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGT
CGACCTGCAGGCGGCCGCGAATTCAGTAGTGATTCCGAACACGGGGAACTTACCAGGTC
CAGACATAGTAAGAATTAACAAGCTTAATTCACAAAGGACAGGGAGCAGAACTTGCTG
TATCTTACTATTACCTGGATGACTAATTTTTCAATTCAGTTTCATGTTACCCTACATGA
CCCTTGGGTATCACTTAGCAACTGTGTAAGCAGAGTCACTCTTCTATAGACACTTTGG
TGCATCCACCACCAAGACTAGTGGAACGTGGCCGTGTTCCGAATCGAATTCCCGCGGCCG
CCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTCGTATTAC
AATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACCTT
AATCGCCTTGCAGCACATCCCC-TTTCGCCAGCTGGCGTAATAGCGAA-AGG-CCGCAC
CGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGAC-C-CCCTGTAA-CGG
CGCATTAAGCGCGG-GGG-G-GG-GGT-ACCC-CA--G-GAC-GTTA--T-GCCAGG-CC

TA-GCCCGCTCCTTTTCCTTTTTTCCCTTC-TTTTC-CCCGTTCC-GGTTTTCC-GTAAG-
-T-AA--GGGGG-T-CCTTTAGGT-C-AATT--GGGTT-CGG-CC----CCCCAAAAATT
T-T-AGGGG-GG-TCCCCA-GGGGCCCCCCCG-AAA-GGTTTTCCCTTTG-G-GG---
C-CC--TT-AAA-GGAC-T-T

ATJ4-02/3

GTGACACTATAGATTCC--TAGCATA-G-A-A-AACGCGTTGGGAGTCCTCCCATATGGT
CGACCTGCAGGCGGCCGCGAATTCCTAGTGATTCCGAACACGGGGAACTTACCAGGTC
CAGACATAGTAAGAATTAACAAGCTTAATTCAACAAAGGACAGGGAGCAGAACTTGCTG
GATCTTACTATTACCTGGATGACTAATTTTTCAATTCCAGTTTCATGTTACCCTACATGA
CCCTTGGGTATCACTTAGCAACTGTGTAAGCAGAGTCACTCTTCTATAGACACTTTGG
TGCATCCACCACCAAGACTAGTGGAACGTGGCCGTGTTTCGGAATCGAATTCCTCGCGGCCG
CCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTCGTATTAC
AATTCCTGCGCCGTCGTTTTACAACGTCGTGGACTGGGAAAACCCCTGGCGTTACCC-
AACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAG-CTGGGGTAATAAGCG-A-AAG
GCCCCGCCACCGATCGGCCCTT-CCCAAAG-TTGCACA-GCCCTGAATGGG-GAAATGGA
ACCCGCCCT-GTAACGGG-GCATTAAAAC---GG-GGG-GGGGGGGGTTACCC-CAAG
GG-GACC-GTTTAA-TTG-CAGGGGGCCCTAA-GGCC-CCCCTTTTC-GTTT--T-CCT
TCCTTTTCCCCCA-ATT-GCC-GG-TTTTTCC-GC-AA-G-TTAAAAGGGGGGGT-CC
CTTTAAGG-TCC-AATTAGGAGTTTT-GGGCCCC---CCCCAAAA-TT-G-TA-GGGG
G--G-TCCCCAAGGGGCC-C-CCCCG-AAAAGGGTTTTCCCCCTT-G-GGGGACCCCC-
-TTTTA-A-GGGCCTT

ATJ6-01/38

TGACACTATAGA-TT---TAGCTATGCATCCAACGCGTTGGGAGTCCTCCCATATGGTCG
ACCTGCAGGCGGCCGCGAATTCCTAGTGATTTCGTTCCGCACAAGGAATGGCTGGTCGC
TTATATCAAGCTCTAAAAGGCTATCGTATTTTCTTAGACACAGAAGCCGATTTTCGATTTA
CACGATTTGAAAGAGATCGTAAAACATACTAAATTGTTTGTTTTTATACTTACGGAAGGC
ATTTTTGATAGTTTTTGGTGCTTACAAGAGCTCAGAGCTGCTGTTTCGTACAAAGTTCTT
GTTGTAATTATTAGAGATATTAATTCCAATATCCAAATCCTTGGCCTGCGGAACGAAAT
CGAATTCCTCGCGGCCCGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTT
ATAGTGAGTCGTATTACAATTCCTGCGCCGTCGTTTTACAACGTCGTGACTGGGAAAACC
CTGGCGTTACCAACTTAATCGCCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAAT
AACGAAAGAGGCC-CACCGATCGCCCTTCCC-ACAGTTG-C-CAACCTGAAATGG-GAA
TGGAAC-CC-CCCTGTAACCGGG--ATTAA----CGGGCGGGGGGGGGGGTTAACCG-

CAA-G-GACCCGTTAAAA-TTGGCA-GGCCCTA-GG-CCCGGTCCTTTT--GTTTTTTTC
CCCTCCCTTTTCCCCAGGTTTCGCCGGGTTTCCCCGC-AAG-TTAAA-ACGGGGG-C-CC
CTT-A-GG-TCCAAATT-GGG-TTTT--GG-CCC---CCCCCAAAAATT--T-AGGGGG
GGGT-C-CAAGGGGCCCCCCCCCT-AAAAAGGTTTTCCCCCTTT-AGGGGGCCCCCCTTT
TTAA

ATJ6-01/48

GCTTATTTAGGTGAACCTATAGAATTCCTCAAGCTATGCATCCAACGCGTTGGGAGCTCTC
CCATATGGTCGACCTGCAGGCGGCCGCGAATTCACTAGTGATTTTCGTTCCGCACAAGGAA
TGGCTGGTCGCTTATATCAAGCTCTAAAAGGCTATCGTATTTTCTTAGACACAGAAGCCG
ATTTTCGATTTACACGATTTGAAAGAGATCGTAAAACATACTAAATGTTTGTTTTTTATAC
TTACGGAAGGCATTTTTGATAGTTTTTGGTGCTTACAAGAGCTCAGAGCTGCTGTTTCGT
ACAAAGTTCCTGTTGTAATTATTAGAGATATTAATTCCAATATCCAATCCTTGGCCTG
CGGAACGAAATCGAATTCGCGCGGCCGATGGCGGCCGGGAGCATGCGACGTCGGGCC
AATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGAC
TGGGAAAACCCTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGC
TGG-GTAATAGCGAA-AGG-CCGCACCGATCGCCCTTCCAACAGTTGCGCAGCCTGAATG
GCGAATGGAC-CCCCCTGTAACGGGGCATTAAAGCGCGGGGGG-G-GGGGG-TACCCCCAA
GGGGACCCCT-AC-T-GCCAGG-CCTA-G-CCGC-CCTT-G-TTTTT-CCTTC-TTTCCC
C--G-TGCC-GGTTTTCC---AG-TT-AA-GGGGG--CCCTTAGGG-CCAATT--GGGTT
TCGG-CCCCCCCCAAAA-T-TT-GGGGGGG--CCC--GGGGCCCCCCC-AAAA-GGTT
T-CCCCTTGGGGGGG-CCCCC-TTT-

ATJ6-01/50

CACCAG-----A-G-G-----A-ATTTTC---TAG-AAG--GAAAACGG-G--GGGTTTT
CCCCCATA-GGAAAACTGCAGGCGGCCG-GA-T-CACTAATGATT-C-AC-C-TCTCC
CCCGAATGGCTGGTA-CT-ATATCA-GCTCTAAAAGG-TATCCGCAT-TCCCTTA-ACCC
---A-GCCGATTTTC-ATT--C-CG----G-AA-A----G-A-A-CATACTA-ATTGTTTG
TTTTTATACTTACGGAAGGCATTTTTGATAGTTTTTGG-GCTTACAAGAGCTCAGAGCTG
CTGTTTCGTACAAAAGTTCCTGCC-GTA-TTTTT-GAGATA-CAAATTCCTCAAGTATCCA
AATCC-TTGGGCCTGCGG-AA-GAAAATCGAATTTCCCGCGGCCCG-CATGGC-GGCCGG
GGAGCATGC-GA-G-TCGG-G-CCA-TT-GCCCCTATAAGCGA-GGC-GCTATTTA-AAA
TTCCACGGGCACG--CGGTTCTTACCAA-GGT--C-GG-CCTGGGGAAAA-ACCCCTGGA
GTGT-GACCCCC--CCTTTC-A-T--G-C-TT-G-GTAG-G--AT-CCCCCTTTCG-C-T
CA-GCTGGGCGG-TAA-----CAA-AAGGTG-CC--T--C--GAT-C--CCC-TTTCC

--A-C-G-TGGG-AT-A-TCT--AATGGG--GAA-GGGAC--C-CACCC--T--AGGGG-
C-G-C-TGTAAG-CC--GGGGCGG-CGGGGGGAGG---T-CAC-CCCA-G--GGGG-TCG
-CT-CAAA-TTG--CAAGGGAG-CCCTAAA-GGGG-ACGCGGA-C-C-G-TTGCGTTT--
TG-C-CCTCCCTCGTC-A-TCGC-CT--TG--CT-TCG-GA-TTTTCTCCCCGA-CA--G
-TTCTCATAATA--G-GGGGGC-TCTCCCCT-TCGG

CTJ4-24/5

GTGTTCC-CTCTAGCAA-A-AA--CAACGCGTTGGG-T-CCTCCCATATGGATCGACCTG
CAGGCGGCCGCGAATTCAGTAGTATTGTTAACGACGGCCAGTCCATGATACACAGCAGAA
GGAACCTCCCTCCCCTACTATGTAACGATTGTTGTGTCTCACATTATCTTCTAGATCTTCT
AACTTCTTCC-GTAATACTCCACGAGATCAACTTTGCGCCCAAAAATCCCAAGAAATCCA
ACACGCCCAAACCTTTGAGAAGTCTTTTGTGTTGACTTTAGATGAGCAAGCCTTCTATTC
AGCTTATCAGCATC-CTAATTAGCTTCTGAATCTTGGTAGTGTGATGCACCATGTAATGT
GACAGATATAATGAAGGATAATATTCTGAGAAAAAATTCTCAACACTTGCACTTAAGCTG
GCCCGTGTTCGGAATCGAATTCCCGCGGCCCGCCATGG-GGCCGGGAGCATG-GACGTCGG
GCCCCATTCGCCCCCTATAG-GAG-CG-ATTACAATTCAGTGGCCG-CGTTTTACAACGTC
GTGACTGGG-AAAACCCTGGCGT-ACCCAAACTTAA-CG-CTTG-A-CACATCCCCCTTT
TCGGC-AG-TGG-G-AA-TAGCG-A-A--G-CCGTACCG-ATCGCCCCCTCCCA--CA-G
TG-CG-A-TCCTGAATTGG-GAATG-ACC-C-CCCAGTGTAAACGGCC-CATT-ACC-GC
-TG-GGGCTG--G-T-G-TTT-A-G-CC-A-CGTGACCGG-TA-A--TT-CC--C-GTCC
TAACGG-CC----C-TTA--G-TTTTCT-CCTT-CCGT-CTCCA--AC--TTCGCCGG-
TTTAT

CTJ4-24/15

A-AGTTC--CTCAAG-AAA-AA-ACAACGCG-TGG-TCTCCTCCCATATGGTTCGACCTGC
AGGCGGCCGCGAATTCAGTAGTATTCCGAACACGGCCACGTCCACCTCCACCAGAAGTG
TAGTTTCTTTACCAGCAGCAGGTTTCAAAAACCTCATTGATGCTGAGTACTTCTTGGCT
CTCTCTTCTTTCTCAGCAAGCTCTTTACGCTTATCCTTGTGAGTCCCCAGTTTAGCAAAG
ATTTTCATCGTTGCTCTTCTTGTGAGAGTGTGTCATAGATGCAAGCTCTTGTCAACT
TCCACTTTTCTTTCTTCAAGTCTTGTGAGTGTCTCTAAAGCTTTCTCTTTTCTCAAGAACC
TTTTGATATTCTTCAAGAGTCATCTCCTTATCCTCAGGCTCCTTCTCCTCAGTTTCATTT
ACCGTGTTCGGAATCGAATCCCGCGGACGCCATGGCGGCCGGGAGCATGCGACGTCGGG
CCCAATTCGCCCTATAGTGAG-CGTATTACAAATTCAGTGGCCGTCGA-TTTACAACGTC
GTGACTGGGAAAACCCTGGGCGTTACCCAA-CT-AATCG-CTTGCAGCACATCCCCCTT
TTCGCCAGCTTGGCGGTAATAGCGGAAGAGGCCCG-ACCGATCGCCCTC-CAACAGTTG

CCGCAAACCTG-AATGG-GAA-GGA-GCGCCCTGTAGCGGGCGCATTAA-CGC-GCGGGG-
GTGG-TGGTTAA-CGCAACGTGGA-CGC-TAAACT-G-CAGTGGCCCTAAGGCCGC-TCC
TTTTGG-TTTC-TCCCCTCCTT-AT-GCCAG-TGGC-GGGAT

CTJ4-24/43

GA-GGCGGCCGCGAATTCAGTACTGATTCCGAACACGGCCACGTCCACCTCCACCAGAAG
TGTAGTTTCTTTACCCAGCAGCAGGTTTCAAAAACCTCATTGATGCTGAGTGACTTCTTGG
CTCTCTCTTCTTTCTCAGCAAGCTCTTTACGCTTATCCTTGTGAGTCCCCAGTTTAGCAA
AGATTTTCATCGTTGCTCTTCTTGTGGAGAGTTGTTGCATAGATGCAAGCTCTTTGTCAA
CTTCCACTTTTCTTTCTTCAGTCTGAGTGTCTCTAAAGCTTTCTCTTTTCTCAAGAA
CCTTTTGATATTCTTCAAGAGTCATCTCCTTATCCTCAGGCTCCTTCTCCTCAGTTTCAT
TTACCGTGTTCGGAATCGAATTCCC GCGGCCCGCCATGGCGGCCGGGAGCATGCGACGTCC
GGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTACAACGTC
GTGACTGGGAAAACCCTGGCGTTACCCAACCTAATCG-CTTGCAGCACATCCCCCTTTCC
CCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCA--C
CTGAATGGCGAATGGACGCGCCCTGTA-CGGCGCCATTAAGCGCGGA-GGGGGT-GGGGT
TAC-C-CAGGC-GACCGCTACACTTGCCA-GGCCCTAAGGCCCGGCTCCTTTTCGCTTTCC
TTCCCTTCCTTTC-TCGCCACGTT-GCGGG-T

CTJ4-26/1

-A-GATTCCACTTA-A-----AAAAG-G-G--GTTC--CCTCC-AAAGGAAAAAA-CTG-
GGGCGGTGCGCAA-TCTCTAGTGA-TACGATTATTTCCCCCATGCTGGGG-AAA-GGCAG
GAATGA-CTTT-C-T--TTGGAGACACATGG-AAGATTTG--GG---G-C-CG-TTA-CT
GGCC-GCG-GGCC-AGACGATTTTCGTAAAGCAATAG-AGAAACC-AG-AAC-GCTACGC
-AA-T-GC---G--A-GACGA-AG--GA-TCGTGCC---G-GA-TATGC-ACAAGGGGA
AGAA-G-TTG-TTTTTTTTTTTGGCGA-GGGGAAGAAGA---TCCTGTT---G-TCTCG-
-CG----GAG--CCCCG---C--G--CG-GAC--GAGTA-G--AAAC-CCCC-CCGGC-C
GC-CA-GG--CGGC-AGCAG-AATGCTAAC-TGT---GT-CCC-AATA-CGG-TG-A--G
GCGGAGG-----C-CAAAA---AAATGG-C--GTGT--C---TA-AA--TG-GATGG-T-
-GGGAAAAA--GAGCGA-T-GT-T-CG---GAGAC-TA-TGGC-CGGG-A-C-GGACCGT
TGC-C-ACG--TGTA-CCGC-GGGCG-G-AATAGCG-G--AACGG-GCAC-TGC-CTGAC
--GTGCC-ATACC-GC-CTGGGGGG--GAAG-CCC-G-A-TTGGAAGCAAAGGA-AAG-A
A-AT--GG--AAGGGAGGCAA-TAA--AGT----G-GGT-GGG-C-G--CCA--A-C-TA
-AATGG-C-GACCA--CC--CC-CTG-A-GTCCCGAGAGGGG-C-C-G-TGCGG-AT--A
T--AC-GCA--C--AAC-GAACCAACCTA-GGAAA-CCC-A-G-GGGG-G-C---AGGGG

GCTATTTA-T-G---A----GGCCC--C

CTJ4-26/2

G--TTCT-C-TC-CGAG--G-G-AAAA-G-CGGCG--CGGTTTC-TCTCTCTCTATAA--
AAAAAACCATGCAGGCGG-TC-CGAATGCCCTC-CTGCAATCCGCA-TTTTTTT-CCCGC
ACACCGACT-AGTAATCG-TGAACAGGAA-ACC-----TTC-TCCTCCC--ACC-GGCA-
AATCACTAC-A-AGGT-ATC-AATAT-C-C--TCCTA-AGATCCC-CAAAT--AT-GGTT
AT---T---TCC-TC-A-GAA-TC-CGC-A-C--TG-CGCT--C-TG-GTC-ATGGC--T
-CT-CG-AA-TGG-C--AT--AGA-TAATA-A-TC-CGGA-G-AGCA----CA-TAG---
--AGGAG-CCTGGA-TTTTT-TTTTT-TTTC-----CATGATCT-CG--TCAA-GAG-TA
-T-T-TCCTC-C-CTG-G-CG-GAACAC-CCGTC-TC-TCCCCTCCGG--AC-GTCA--C
CAAC-AACTCT--ACGG-T--C-CA-TGGG--CGTCCGTGG-A-CT----T-AAC-CTC-
TGGG-C-A-ACTT-GGGGCC---T-AGTGGGG-C---TTT-C-A-T-AAAC-GGG-A--G
TCGTT--T-GAATGCT---AG--GTG-GGAT-A-T-TC---G-----A-C-CATA-CGT
-A-TA--CCC-AGC-CTT-AG---C-TCC--TAT-T-G-T-TGCGG--G-CA-ATA-G-G
T-C-TCG---CC-G---T-G-AGCTA-AGGTTT--TCA-A-A-G--AC---T-GT-C--G
T--GGGGA-G-G--CGG-GAG-C---AACCTTGT--A-GG-TGC--GT--T-C----CC-
CACAGGTG--T-A--G--T---TA-CGCAAGAACTACCTTTACTC-G--T-G--T-ATGT
-C-CT-CTAC-G-G-C-G-A--TTATC-T-A--T-TCTTCT-----CTT-CTAC-CG-C
-CCGCGCC-GACT-CT-G-ATA--GGCTGGT---G--AC--CGT-ACGAGCGG--GTGCG
-A----TC-GAG-T-T-GACG-GTTGC-G--CGT-GT--A-G--G-GCTCT-G

CTJ4-26/3

T-TTTCCTTCTCTAAGAAA-GG-AAACAACGCGTTGGTTC-TCCTCCCATATGGTCGACC
TGCAGGCGGCCGCGAATTCAGTAGTGATTCCGAACACGGCAGAATAATCAACAGAAATGT
CTGGGATGATGTAAAATGAAATCAAATATTTATATTCATACATTTTGACACATGTAACT
CTTCTTAAAATCTCCAGGTTGACCTTTGACTTTTCAGTTAACTTGACTGATGCCTGGTTT
TCTTTGAAAGAGAAGTGACCCAATAAGAATGGCAATGCCAGTTATTGCTATGGCGAAAAG
AAGCCGCCTCGTAGAGATGTCTTGTCAATCCTTACTATGTCTGGACCTGG-AAGTTTCCC
CGTGTTCCGAATCGAATTCCC CGCGCCGCCATGGCGGCCGGGAGCATGCGACGTCCGGCC
CAATTCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCG-TTTACAACGTCG-GA
CTGGGGAAAACCCTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCA
GCTGGCGGTAATAGCGAA-AAGCCCCG-A-CGGATTCGCCCTTCCCAACAG-TTGCGCC
AGCCCTGAAAGG-GAATGGAACGCGCCCCTGTTAGCGGCGCATTAAACGC-GGCGGG-TG
TGGTGGG--TAC-CGCAACG--GACCCCTACACTTGCCC--CGCCCCTA-CGCCCCGCTCC

CTTCCGCT-TCCTTCCCCTTC-TTTC-C-CC-CCGATCGCCCGGCTTTCCCCGGCAA---
C-TAAATCGGGGG-CTCCCTTTAAGGGGTCCCAATTTAG-GC-TTAA-GCACCTCC

CTJ4-27/16

GGGTGACACTATAGATTT-TTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATG
GTCGACCTGCAGGCGGCCGCGAATTCAGTAGTGATTTTTTTTTTTTTTTGCCAATTGAAAA
GAGATAGCAAGGAGAAAACATTGTTCTAGTCAACATCCAAAGGTCTTGAAAGCAAATGG
GTCGGGTACAAATCATGCAAACAATATCCATACTCAAACCCTAACCACAATGACCAAGT
CAAACCACCACACATGTTGACTAAATAATCTATGTGCTGCTAGCTGCTGCTGCTGCAACT
TCTGCTTTCTCCTCAGAACCTCCTTCTGCTGTGTATCATGACTGGCCCGTGTTCGGAATC
GAATTCCCGCGGCCGCCATGGCGGCCGGGAGCATGCGACGTCCGGGCCAATTCGCCCTAT
AGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCT
GGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTCGCCAGCTGGCGTAATAGC
GAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACG
CGCCCTGTAGCGGCGCATTAAAGCGCGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTA
CACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCTTCTCGCCACGT
TCGCCGGCTTT-CCCGTCAAGCTCTAAATCGGGGGCTCC-TTA-GGGTCCGATT-AGTGC
TTTACGGCACCTCGACCCCAAAAACCTTGATTAGGG

CTJ4-27/30

GTGACACTATA-TT--CTTCAAGCATAAAAATCCAACGCGTTGGG-TCTCTCCCATATGG
TCGACCTGCAGGCGGCCGCGAATTCAGTAGTGATTCGGAACACGGCCACGTTCCACTAGT
CTTGGTGGTGGATGCACCAAAGTGTCTATAGACGAGAGTGACTCTGCTTACACAGTTGCT
AAGTGATACCCAAGGGTCATGTAGGGTAACATGAAACTGGAATTGAAAAATTAGTCATCC
AGGTAATAGTAAGATCCAGCAAGTTTCTGCTCCCTGTCCTTTGTTGAATTAAGCTTGTTA
ATCCTTACTATGTCTGGACCTGGTAAGTTTCCCCGTGTTCCGGAATCGAATTCCCGCGGCC
GCCATGGCGGCCGGGAGCATGCGACGTCCGGGCCAATTCGCCCTATAGTGAGTCGTATTA
CAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAAC
TAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCAC
CGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGC
GCATTAAGCGCGGGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCC
CTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCCCTTCTCGCCACGTTCCGCCGGCTTTCCC
CGTCAAGCTCTAATCGGGGGCTCCTTTA-GGTTC-GATTTAGTGCTTTACGGCACCTCGA
CCCAAAAACCTTGATTAGGG

CTJ4-27/40

GGGGTTTCCTCTT-A-CCAAAGGAAAAAA-CG-G-CGGTTC-TCCTCCCATAA-GAAAAC
CTG-AGG-GG-CGCGAAT-CACTAGTGATTCGGTTTTTTTTTCCC-TACACAG--GAAGGA
ACTGT-ACTCGA-CCAATCA-TC-A-GATCTACTCTTGATCATGTGAT-CTGGAGCAATA
AC--CCCC-T-ATGAT-A-CAACAGCTCGATCAAATCC-CC-AATTCAGTCCGGCAGAGA
AGTTG-TAGAATCACCGGACCAAGATG-ACTTGCCG-CGTGACCTC-T-TTTTTCGG-GA
TCCGTTTTTTTTTGT-TGGC-GATTT-CTATCCTTGCGTGTTC-TGC-CGTGTTTCGGAA-
CCA-TTCC---GGCCGC-ATGGGGG--GGGAGCATGCGAAAG-CGGGCCAAT-CGCCCT
ATAGTGAG-TCGTATTA--ATTTCACTGGCCGTAT-A-ATTA-AA---CG-GGACTGGG-
AAAAC-CTGG-CGTTACCC-AA-CTT-AATCGA-CTTG-AGCA-CATGCCCCCT-TCGGC
CAGACTGAG-G-TAATGAGCGAA-GA-GCC--ACA--GGAAG--CACTTTC--AAAT-AG
TT-G-CGGCGGCC--TAGATGGCCGAT-TGGGACAGCGCCCGTGG--AG-CCGGCCC-A-
-TA--C-CCG-T--G--ATCGTGGT-G-GTTACATCTC--CCGATGAACCCG-TA---C-
TTGTACTA-ACGCT-CTGATA-G-CCCG-TCCAT-TTCG-C--T--T--ACCGTGCGAT-
TT-T-GT-CAC--TA-A-CC-GTCTAT-T-ACAG-CCA-GTTCT-AAA-CAGGG-GGCC-
TCCTTCC-TGG-C-CC-AA-T-T-GGGC-T---GGG-G--CTTCTAA-AC-GAAAAC-

TTJ6-52/28

GGGGGTTTC----TTA-AGAGAAA-AAAAA--GGG-GCG-TTTC--CCTG-GGG--AG-
-AAAAGGT-GG-GCGGCGGC-C-CGCAGTG-G--G-AGG-GAAA-AAA-TTTTTTTT-C
CCC-GGGGGA-TGGGGAA-AGAA--G-AGG-C--GAAG--GAGGGCACGG-A-T-TAT-A
AAAGG-GAGCGGGAA--A--A-G---TG-GA--GTC-ACA-G-AA---G-TGC-A-GCG-
GAGC-CCGA-G-AGG---ACGAGGGGGGGG-G--CG-G-GGAAAGTA-A-G---GAAA--
AG-A-AGCG-GG-GGGAG----A--AAATGCTGGG--GCGC-GT-GAGGG--GGTGTAGA
---T-TGCA-AACAG--AGA-AG--GAAG-GAA-GTG-TTTTTTTTTT--G-G-AGA-A-
CG-G-GAG--C-AA-GGC-TATG---TGGC--G-GCAA-T-GT--TA-GGG-GA-G-T-C
-C--TTGAC-GCAA-ACGGGGGGG-GGG--GAG--T-AGA---A-GAGT----TAGGG--A
--A--G--GGGGCGGAGGAG-GAG-T-AGA-GTG--GG-CA--A-GGTGGGAG-CAAGAC
ATA-CT-A-GTAAAG-G--AAC-GA-TG--A-AG--AA--G-G-G--GGAG-AGG-A-C-
GGAGG-CGT-GGA--AGGAG-GA-G-A-CAC--TATTGA-TATTGTGAGTAGCAAGTAAG
A--A--AAG-G-GCA---AGGGGGGAGAAGC-G-CGATAG-ATG-A-AG-GA-GATAT-A
-AG-A-GGAG-CG-GGAAGAAA-A-GAAGAGGGG--TGGTG-ACT-AAGAC-A---GGGT
CGGGGC-AAGATA-TGA-T-AAGAGAAG---GGTAA-A-AG-G-CTGAG--CGA---GAG
GGAACGT-GAG--ACGAGGGGCA-GT-A-G-AA-A---ATCGA-A-GT-GACGATG---G
CG-TAG--AC--TG-C-ATAA-AGGGATGGGCAT-CGGCAAT-ATT-ACAGA-AT-GGAC

GTAGGA-C-GCA-GA-C--AG--AC-G-C-A-ACAG--AT-GA-GG--AG-TAGG-A--A
CAACAGGCCGGAGG-GA--TA-T

TTJ6-52/31

G-GGATTTCTC-TTTA-AA--AAAAA--GCGG-G-TTC-C-TCC-AAAAAG-AAAACA-
GG-GGCGGCC-C-AA-A-CC-TA--TGTTT-TT-TTTTTTTT-CCC-ATG--GCTGAATAA
-AG-A-CACCGC-CC--A--CTTT-C---CATCC--TTTT---AAATAT-TA--A--T--
CAG--CCCTTCCCCC-G--G--G-CGGAT--GTCG--GCTCGG-CCT--TAC-T-TG--T
CACT--T----C---C--C-G-TCT--C-C--CA--C-TC--GC--AC-T-A-CCCTTGG
CCA--C-CACGCCTTTTTTTTTTTTTTTG-C-CACATCAT-CT---A-C-C-TT-TTC-GG-
A--ACG-G--T---A-CA--AG---TTC--A-TA-CCCC-CTCCT-TT-C-GC-CTTCT-
GTCGC--T--CCA---AA-TT-AACATCC--TCACA-T-AGAC----G-GCAGTA-TTAG
TCT---TT-CCTTTT--GATTTT-A-TTT---T-TGTT-T-A-TT-G---T-C-CCTCG-
-CC-CT--A-T--ACGTG-CTC-AT-ACC--G-ACGC--AG--TT--C--TCCCCT-T--
AA---C-CGAACAAC-TA-----TGGGA----GGT---CA----A---CAC-T--
----TTCTATC--CC---CTCG--A-G-GATCTCGCAT--CTCG-AA-CA---GA-TCTG
TAAA--CAA--T--T-C-TGTT-TGTCG---TC---G-A-ATCTAC---TGG--CT-T-T
TG-C-T-TG-G-T--A-A-TT---AG--GA----CC-TCAAATT-AT-ACAT-AT-TAA
T--AA----T-TT-TTCTC-TA-TTA--T-TG-GGA

TTJ6-52/50

GGGTTC-T----AG----G-AAAGGGGGGGTTC--CTCCCAA-G--AAAA-GGGGGGCC
CCCGCC-G---A-G-AA-A-T-TTTTTTTT-CCCC-GGG-G-AGAAAAAAGGGCT-TC
A--AAG-TT-G-GCGC-GG-A-GAG-AGGGATAT-GG--G-AC-C-CC-TA---GG-C-G
C-G---CCAG-C-A-GGG-GGC-TA-A-G----GAC-GTA-A--T---GTGTACCG-CGA
CCGGTAAC-GGCAG-CG-AGGC-CAAA-G-A-GC--G-G-GCGG-ATA--G-AC--GA-A
GG-GCTTTTTTTTTTTA-G-GACC-G-GG-A-G-AAC-TTTTT-ATGGG-A-TG-TA-GG
-A---G-G---CATGAAG-GG--G-GACCCA-GA-CG-GACGGGG-G-GC-TA-ATACCT
GC-GAAGA-A-AG-T--G-G--G--C--GGGGG-GG--G-AGAAAAG-CGGG--GGGGA-
CGA-GATCCGG-T-C-AAG----G--AAAG---AGGGC-GG---A--AG---AG-GA-AG
-AG-CA-AAGA---G-AGA-GAAT----ACAG-CCTC-A-A--A--G-----C-CTG-A-
T-CACGAGAGG-G-C-GG-AAGGAG---G--T--A-AAG-GTG-C-CGCAC-A-GAC---
GAA-AG--GAGC-T-ATA--GAGA--C--GG-TG-ATG-GG-GGCGG----GG--GACAG
GGGC--AATG-G-AGCG--G-T-T--CAAG-CG-AA--GTGG--AGA-A-GAT-CAGC--
GG---TGT-TC-AAG-GGGG-GAGGA--ATGGG-G-AGCTA-CG--A-AG-GA-AGCA-C

A--GACGGA-CAG--AAAG--G----AGG-GACGGGG

TTJ6-53/2

A-AGTT----CAAG-AAAG-AAA-CAACGCGTTGG-T---CCTCCCATAAGGACGACCTG
CAGGCGGCCGC-AAT-CACTAGTGATTTTCGTTCCGCAGGCCAAGGATTTGGATATTGGAA
TTTAATATC-CTAATAATTACAACAGGAACTTTGTACGAAACAGCAGCTCTGAGCTCTTG
TAAGCACCAAAAACCTATCAAAAATGCCTTCCGGAAGCATAAAAACAAACAATTTAGTATG
TTTTACGATCTCT-TCAAATCGTGTAATCGAAATCGGCTTCTG-GTC-AAGAAAATACG
ATAGCCTT--AGAGCTTGATATAAGCGACCAGCCATTCTTG-GCGGAACGAAATCGAAT
TCCCGCGGACGCCATGGCGGACGGGA-CATGCGACGCC-GGCC-AATTCGGCCCTATAG
TGAGTCG-A--ACAA-TC-C-GGCCGCCGT--TACAACGTCG-GAC-GGGGAAAACCC-G
-CGA-T-CCCAACTTAA-CG-CT-GCAGCACA-CCCC-TTTCGCCAGC--GG-GTAATA
GCGAAAGAGGCCCGCCACC-GAC-GCCCCTTCCCAACAGA-AGCGCA-GCCTGAA--GGC
-GAATGGACGCGGCC-G--AGAGGGC-CATTAA-CC-CGG-CGGCGG-GGGGGG-TA-C
GC--CAGCGGGACCCGC-ACCACT-GGCCACGGCCC-AACG--CCCGC-CCTT-TC--T-
-C-CAC--ACAATCT-GC-ACGATCC-CCGATTAC-CCGGCAGCTC-AAAACGGGGG-C
-C-CC-CAGGG-AACC-ATCAGAAGCC-AAGG

TTJ6-53/3

G--TTTCTCTCT-TAG--A-G-AAA-CAACGCG-CTGG-TTCCTCT-CCCATAATGGTCG
ACCTGCAGGCGGCCGCGAATTCAGTAGTGATTTTCGTTCCGCACAAGGAATGGCTGGTCGC
TTATATCAAGCTCTAAAAGGCTATCGTATTTTCTTAGACACAGAAG-CGATTTTCGATTTA
CACGATTTGA--GAGATCGTAAAACATACTAAATTTGTTGTTTTTATACTTACGGAAGGC
ATTTTTGATAGTTTTTGGTGCT-ACAAGAGCTCAGAGCTGCTGTTTCGTACAAAGTTCCT
GTTGTAATTATTAGAGATATTAAATTCGAATATCCAAATCCTTGGCCTGCGGAACGAAAT
CGAATTCGCCGCGGCCCATGG-GGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTA
TAG-GAGTCGT-TTACAATTCAGTGG-CGG-CG-TTTACAACGTC-CGACTGGGGAAAAC
C-CTGGCGTTAC-CAACTTAATCGCCTTGCAGCACATCCCCCTT-TCGCCAGCTGGCGTA
ATAGC-GAAGAGGC-CG-ACCGATCGCCCTTCCCAACAGATGCGGCAGA-CTGAATGG-C
-GAA-G-A-CCCTCCCTGGAAGC-G-GCATTAA-C-GCGGTCGGT-GTGGCGGGC--CCC
-CAG-GGTGACCGCTA-AATTAGCCA-C-GCCCTAGCG-CCCGCTACTTTT--CTTTCCT
TCCCTT-CT-T-C-CG-C-T--GTT-GCCGGGTTTTCC--GG-CAA-GCCTA-ATC-GAG
GGCCTCCCCT--AG-GG--CCAGAAT--A-AGCG--G-

TTJ6-53/50

GTGACACTATAGA--C-TTCAAGCTA-GCATCCAACGCGTTGGGAGCTCTCCCATATGGT
CGACCTGCAGGCGGCCGCGAATTCAGTAGTGATTTTCGTTCCGCACAAGGAATGGCTGGTC
GCTTATATCAAGCTCTAAAAGGCTATCGTATTTTCTTAGACACAGAAGCCGATTTTCGATT
TACACGATTTGAAAGAGATCGTAAAACATACTAAATTGTTTGTTTTATACTTACGGAAG
GCATTTTTGATAGTTTTTGGTGCTTACAAGAGCTCAGAGCTGCTGTTTCGTACAAAGTTC
CTGTTGTAATTATTAGAGATATTAAATTCCAATATCCAAATCCTTGGCCTGCGGAACGAA
ATCGAATTTCCCGCGGCCGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCC
TATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC
CCTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAAT
AGCGAAGAGGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGG
ACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGCGGGTGTGGTGGTTACGCGCAGCGTGACCG
CTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTCTTCCCTTCCTTTCTCGCCA
CG-TCGCCGGCTTTCCCGTCAAGCTCTAATCGGGGGCTCCTTTTAGGGTTCCGA-TTAG
TGCTTTACG

GTJ1-12/5

GGATTTCTC-TCTAAAA-AG-G-AAAAA-G--GG-GTTGTTTTT-TCCTCCAA-GAA-A
A-AAAAAC-TGGC-GGGGGGT-CCCT-AATTCACTAATGTGA-T-T-GGTTTT-TTTTTC
C-CGC---AA-GGT-AAAAAC-ATGAT-CAA-T-A-ATC-C--TC-TCTTGACG-G-ACC
-TA-TGT-TA--A-G-AATGCGCC-AAAAT--TCCT--GCCCCCTTTAT-AAACAT-AAA
AGG--A--CACTGCATTATG--CC-GAATCA-ATTC--GC-G-CGACATGGCGG-CGGGA
GCAT-C-AC--CAGGGCCAA-TC--ACCTACATATTTTTT---TTAACAA-AT-A-GTG-
C-C--CGCCTTTACAAG-GGGTG-TAA--GAGG-CAAACCC--GCGCGA-TTGT-GA--T
-GA-T-TG-T-TTG-C-CAATAATCT-TA-CTTTAAGG-GC-GA-TGGA-GA-A-TGAGA
CGGT--AAGGAGTAAGC--A-GGAT-TG-GC-C--C--T-A-CTA-GATG--GAC--GGT
C-GGAA---GGT--G--TGCGAGCCCCCG-CGCTTG--A-GA-GGGGGCATA-A-GGAA-
-CG-C--CG-A-T-AGGAAA--TATAAT-GCGTCCGCAG-GAGCCCGCCC--CA-TTTGG
GT---A-G--CA-GGA--A-TG-CGGA-TACG--TT-A--AAAC-GTAAA---GACT--A
TC-CAC--AC-G--T-CGATGAA-GC-GTGT-CTC-AGACATGGT-G---AG-A---ATG
G-GGAATA-TAG-ATTG-GGTTGTT---AGAATTATGAT-GCT-TAAG-G-CATCCC-AT
CTGC-AT-AT-ATTC-A-T----GGGTA---G--GCACT-C--TA--AGGGCCAAA-AGT
---T-CTGAT-A-CGGTG-AT-TT--CAA--GA-GCGTC-G-CTAA--ACC-C-TTGTGA
ACG--AAAT---G---GATAAGAA-

GTJ1-12/17

A-ATT---CTCAAG---AAAAACAACGCG-TGG-TCTCCTCCCATATGGTCGACCTGCAG
GCGGCCGCGAATTCAGTAGTGATTCCCGCGGCCGCCATGGCGGCCGGGAGCATGCGACGT
CGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTGTTTTACAACG
TCGTGACTGGGAAAACCCTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTT
CGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAG
CCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGGGTGTGGTGGTT
ACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTCTTC
CCTTCCTTCTCGCCACGTTTCGCGGCTTTCCCG-CAAGCTCTAAATCGGGGGCTCCCT
TTA-GGTTCCGATTTAGTGCTTTACGGCACCTCGA-CCCAAAAACCTTGATTAGGGTGAT
GGTTCACGTAGTGGGCCATCCCCCTGATAAGAC-GTTTTTTCGCCCTTTG-ACG-TGGA
GT-CACGTTCTTT-ATAGTGGA-TCTTGTTCCCAACTGGGAACAACACTCAACCCCTATC
TC-GTCTA-TCTTATGA-TTTATAA-GGATTTTTGCGGAATTCGGCCCTTTGG-CTAAAT
AA-GAGC-GATTTACCAAAA-TTTA-CC--AATTTAACAAATATTACACGCCTTACAA
TTT-CTGGAGGGGAA

GTJ1-12/34

--TATTC--CTCAAGAAA--AAA-CAACGCGTTGG-TCCTCCTCCCATATGGTCGACCTG
CAGGCGGCCGCGAATTCAGTAGTGATTCCCGGCATAAATAATGTAGATGGTGAAAACCTAG
AAAGTCAAAGTTTTATTTATAATTTAAAAGTCAAAGCTTTACATGAGAAATGTAGGATAT
TTTACGTTATAATTGTATTTATAAATCTTTCTGCCCAAAAAAAAAAAAAAAAAATCGAATCCC
GCGGCCGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTC
GTATTACAATTCAGTGGCCGTGTTTTACAACGTGACTGGGAAAACCCTGGCGTTAC
CCAACCTAATCGCCTTGCAGCACATCCCCCTTTCCGAGCTGGCGTAATAGCGAAGAGGC
CCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGT
AGCGGCGCATTAAAGCGCGGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCC
AGCGCCCTAGCGCCCGCTCCTTTTCGCTTCTTCCCTCCTTTCTCGCCACGTTTCGCGGC
TTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTA-GGTTCCGATTTAGTGCTTTACGG
CACCTCGACCAAAAACCTTGATTAGGGTGATGGTTC-CGTAATGGGGCCATCGCCCTGA
TAGA-G-TTTTTCGCCC-TT-GAA-GTTGGGAGTCCA-GTTC-TTTAATAGGGGGA--T-
TTG-TTCAA-TGG-AA

GTJ3-28/5

GGTGACACTATAGA-TTCTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGT
CGACCTGCAGGCGGCCGCGAATTCAGTAGTGATTTTTTTTTTTTTTTGGCACCTCAACCCT

GTAATAGTTTTTTTTTTTAAACGTAAGCCATATATGATATTTAAAATGTTTCAAACACTACTAA
CCACACCGTCACCTTGAAACAAACATGGATAACAACACTAACAACATTCGTTTTAAGTAACAT
GAAGTAGATAAAGAGGAGGGATAAATATGACATGAATTTGCTCATAAAGACATGTTCCACC
TTGCTCGTTTTTGGGCCAAGCGGAGAAATCGAATTCGCGCGGCCGCCATGGCGGCCGGGAG
CATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCG
TTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACCTAATCGCCTTGCAGCAC
ATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAAC
AGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGGCGG
TGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTT
CGCTTTCTTCCCTTCCCTTCTCGCCACGTTGCGCCGGCTTTCCCGTCAGGCTCTAAATCG
GGGGGCTCCCTTTTAGGGGTTCCGATTTAG-GCTTTACGGCACCTCGACCCC-AAAAACT
-GATTAGGGGGGATGGTTCACGTAGGGGGGCCATC-CCCTGAAAAACGG-TTTTC-CCCT

GTJ3-28/11

AGGAGTTC-CTTTGA-AGG--GAAA--GG-GG-GTTC--CCTCCAAAGG-AAAAAAA-TG
C-GG-GGCCGCCGAAT-CACTAGTGATC-TTTTTTTTTTCCC-GC--GGGATG--AA-AA
ACCTGG-AACG-CAC-T-CAACCGACGAACAGAAAAAAT-GTTA-A-AACAACCTGAATAC
AAAGCAAGAAGAT-GCT-TGATAAGAAAAAC-CCATAG-AATTTAAAGAGT-C-AATATA
TTCAATCAA-ACTTAGGAGAAGAAATA-AA-GG-CCGGTGAT-GAAAA-CATC-CAA-GG
AT-TTTTTTTTTT-TA--TCCG-GA-GATATCAGAATCAT-TGCATC-ATG-TACTCAGG-
C-TGGC-CACGGAAGGACTT--CACGA-G-GGAGA-A-CGAAT-CCCCGC-GG--CGGCA
TGGATGGC--GGG-A-C--G-G-C-GTCGGGC-CAA-TCAG-CC-ATAGATGAGTCGG--
-TACAATT--ACTGG-GCGC-AGCGT-TTAC-CACG-C-G-GGAC-GGG-G-AAC-CCGG
GGGTTA-CCCAA-GT--AGCG-C-TGG-AGGGACACT----CCGTT-C-GC-G-TTGGGG
GT-AA--C-GACACAGG-CC--C-GCGGA-TG--C--TTG-CCGAA-AGT-GCCTCAGAC
-TGAAATGGGG-GAAAGGAATGCT--CCC-G--AGAG-GCGT-CT--A-GCACATGG-G
GGGTGG-GG-G-GG----CGCC-CCA-GC-G-GGAACCGCG---CGT-TGGGCC-AG-G-
GC--TAAGCAGC-AGGGCTCAC-GT-G-----CGACC-GACT-GT-TTGAGTAC-CC-
CGTAG-GG-GGGG---ATC-CGCGT-AA-G-CGC----A-T-AGAGGGGGACT

GTJ3-28/17

G-TTTCTC-TC-CG--AG-G--AAA-ATACGC-CTTGGTTTC-TCCCCCA-AG--AAAG
A-CCTGCAGGCGGCACGC-A--T--CCTAG-TGATTTT-TTTTTTTT-C-GCCAAAC-A-A
ATAAAAATTAAT-CTCGAAC-TACCAATAAAA--CTAA-TTGCAGTGT--TGA CTC-AC
ACC-CCGC-ACTGTG-TGA-CAGG-GAGATA-AAG--AAACCCATTGTACT---AT-C-C

-GCG--GGTC-AAA-C-GGATA-TCC--TA-TATTGAA---G---AATTACACTAAACTG
---ATG---ATTTCTTATTTTT-T-C-TC--GATTATGTA-A-GTATGTT-TGC-TTT-T
T-CATACAAC-A--ATA-ATT-TTTC---TATGGC-ACCCGG--GAA--TCG---TCCC
-GC-GGCGCTCC-GGGGGCC-G--AGCCATGCC-AA-TT-CGGT-CGCAA-CTTCACC-C
-ATT-GTGAG--G--GCT--ACAATTT-C-TGGG--CG-TCGCA-TCAGAAAC--C-T-C
ACTGGGGCAA-AA-CCCC-GGG-G-TTTA---TC-ACC-TCGAT--CG--CCC---TT--
C--CC-TCCCCCT-GT--ATCCA--C--G---TTAT-CAGAGCAA-AAGG-CA-CAT--
-AG-CTCC-C--TTCCCAA-ACTC--G-GT-A-CCC-GGAAATGGCT-AAC-GG-A-TGT
CT-C-CT-GCAAA-CGG-GCTCTTTT--C-G-TGCG-G--GCGGG--TT-A-CC---CA-
CTG-T--GC-AAACTCTGGCT--CG--GCC-----CTCCGC-AT-TT-TA--T-TT-TC
-CCATCCAC-TCTGT---GAACTAGAC-CGC---T-CCCTGTACAC-CATCTAACA--CG
GCGAG-CTCTTTA-G-GGTCG--CAT-ATATCGG-TATTAA-AGT-GA

GTJ4-43/2

GTGACACTATAGA--C-TTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGT
CGACCTGCAGGCGGCCGCGAATTCAGTAGTGATTCCGAACACGGGAAGATATGGTGGGAA
AGATTTTATGTGACAACAGTGATGGAATAACTGTGGTTGGTAATGGTATGTATTAGGTTT
TAATTTGGTTCCATATTGAGGATGATTCACGTTATGTCTGGTCTTGCTTTTTTTGGGGCTT
TTTTTTTTTTTTGGGACCACTTGAAAAGAAATTATAAAGTGTCTTTTTATAGGAATTTGAA
TTTTAGATGTTATATACTCATATAGCACTGTAAATTTACAACCTTGCAGCAGCAGCAGCA
GCATATGTAAAAACCTCCCAGCTTAAGTAGGCCGTGTTCCGAATCGAATTCCCAGCGCC
GCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTCGTATTA
CAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAAC
TAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCAC
CGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGC
GCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCC
CTAGCGCCCGCTCCTTTCGCTTTTCTTCCCTTCCCTTCTCGCCACGTTTCGCC-GGTTTTC
CCCGTCAAGCTCTAAATCGGGGGCTCCCTTT-AGGG-TCC-

GTJ4-43/39

GTGACACTATA-TTCC-CTCAAG---A--AAA--CAACGCGTTGGGAGCCCTCCCATATG
GTCGACCTGCAGGCGGCCGCGAATTCAGTAGTGATTTTTTTTTTTTTTTGGCATGTTCCGGA
TACTCCGCATTGAGTGATTACGCATAACAGCAAGGCCTGCCAGTGCCCTGGTAGTACTT
TTTCTTCTTGGGGGCGATGACCATGATGAGTTGGCGTCCTTCCATCTTAGGATGCTGCTC
AACGGCGCCGTGTTCCGAATCGAATTCCCAGCGGCCCATGGCGGCCGGGAGCATGCGAC

GTCGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAA
CGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGC
AGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGCGGGTGTGGTGG
TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCT
TCCCTTCCCTTCTCGCCACGTTTCGCCGGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCC
CTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACCTTGATTA-GGG
-GATGGTTC-CGTA-TGGGGCCATCCCCCTGATAGACGGTTTTTC-CCCTTTGA-CGTTG
GAA-TCCACGGTCTTTAA-AGTGGACTCTTGTTCCCAA--GG-A-A

GTJ4-43/47

TTTAGGTGACACTATAG-T-CTTCAAGCATA--AATCCAACGCGTTGGGAGCTCTCCCAT
ATGGTGCACCTGCAGGCGGCCGCGAATTCACTAGTGATTTTTTTTTTTTTTTGGCATGTTC
GGATACTCCGCATTGAGTGATTACGCATAACAGCAAGGCCTGCCAGTGCCCTGGTAGTTA
CTTTTTCTTCTTGGGGGCGATGACCATGATGAGTTGGCGTCCTTCTATCTTAGGATGCTG
CTCAACGGCGCCGTGTTTCGGAATCGAATTCCCGCGGCCGCCATGGCGGCCGGGAGCATGC
GACGTCGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTA
CAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCC
CCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTG
CGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGCGGGTGTGG
TGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTT
TCTTCCCTTCCCTTCTCGCCACGTTTCGCCGGCTTTCCCGTCAAGCTCTAAATCGGGGGC
TCCCTTTA-GG-TCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACCTTGATTA-GG
TGATGG-TCACGTA-TGGGGCCATCGCCCTGATA-ACGG-TTTC-CCCTTTGACGTTGGG
AGTCCACGTTCTTT-ATAAG-GGACTCTT

GTJ4-44/16

A-TATATT-C-TCAAGAA-A-AAA-ACAACGCGTTGG-TTTCCTCCCATATGGATCGACC
TGCAGGCGGC-GCGAATTCACTAGTGATTTTTTTTTTTTTTTGGGGTGATAGTGGAAAAGT
TGACAGATAAGTCACCCTTACTGCCACTCTACAGAACCGTACATGAGATTTTCAAATCGA
ATTCCCGCGG-CGCCATGGCGGCCGGGAGCATGCGACGTCGGGGCCCAATTCGCCCTATAG
TGAGTCGTATTACAATT-ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGG
CGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGA
AGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCG
CCCTGTAGCGGCGCATTAAAGCGCGGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA

CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTC-TCCCTTCCTTCTCGCCACGT-T
CGCCGGGCTTTCCTCCG--CAAG-TCTAAAATCGGGGGGGCTCCCTTTTTTA-GGGTTCCCA
ATTT-GATGGCTTTTAC-GGGA-C-TT--AACCCCCAAAAACTTGGATT-AGGG--GAA
GGGTTTC----GAAAAGGGGGGCCA--CG-CCC--AAAAAA-GGGGTTTTT-CCGCCCTT
TG-AAG-T-GGGGAGAT-C-AC--TT-TTTT-AA-AAGAGGGGA---TCTT--G-C-ACA
ACTGG-GGAAAAGAAAC-TTTA-ACC

GTJ4-44/21

GGGGGTTTC----TTTGAA-----AAAA-----GG-TTTTTTC--CC-CGGGGGAG-AAA
AAAAAT-CCCGCG-CGG---CCC--AAATGGAG---AAGG-AGAAA-AGATGTTTTTTT-C
CCCCGGG-GGGGGGAAGA-AAAA-AAACG-GT----G-ACGG--ATGTAGGAC-GCGGTG
GAGGAG-G-GAGAAAGA-AAGGGGCGAGGTGGAAG-AAC-A--C-GAGG---A-GA--GA
GCGAGAAG-AAAGAGA-CGGGGAAGA-CGG-G-GAA--AAG-G-GGA--A-G--GGCA-G
-CAG-----GGG-TGAG-G--AC-GGAAAGG--GAGC--GGTG-CGGAAAG--A--GGT
-GG-G-GGGG-GTGGGGC---GGC-CAG-GCA-G--G---CACGG-G-CATTTTTTTTTT
TTTTTG-GAGGGA--AGG--GTAA-CGGGG---AA-AGT---AG-AA-AGGCA-GACA-G
AGGTGT-CGCA-CGGGC-CG--CC-GG-GCGG-GGGCGGGGAGG--GA--CG--AGA-GG
A--GGGAGA--A-AACCGGGGAACGCC-GA-CGCGC-GC-G-A-CAAGAGC--G-G-GC-
-CGCGGC--AAG--ATC-G-A-CGAGGAAG-GAAGAGT--GGGGGGGAGAA-AGAGGTG-
GCGGAGG-GG-GAGGG---GA---G-G-GG--GGCAGCGGGC-TGTAA-C-ACC-G-GCG
GAGAGGGGAGGG-AG---G--G-A-GA-G--A--G-GGAGAA--G-GG-G---C-GAGGC
-AA-G--AG-GGGC----GGG-G-AAAAGCG-G-GAGC-A--A-G-G-CGT--CC--A--
GG-CGA-AGGG-AAGA-GA--G-T-GAAA-AGA-AGAAAAGGGAGGGG-AAC-A---G-A
GA--A-AGGT-T-G---GGAGAGAG--GCCA-GGG-CG-GG-GGA--CG-ACGA-AGAG-
-GGG-AGGG-G-AA-GGGGTG---AG-AGGTGGC-AGGAG-CA--CAAGCG-G-GG-AGA
GAGAG-AAG-GAAA-CAGGTAAAGAAGC-C-A-G--A-GAA-A--GAGAAC---AGAC-C
GAGCGG-AGGAGAGGGA--GAGG-

GTJ4-44/25

TTTTAGGTGACACTATAG-TTTTTCAAGCTATGCATCCAACGCGTTGGGATCTCTCCCAT
ATGGTTCGACCTGCAGGCGGCCGCGAATTCAGTAGTGATTCCGAACACGGCAAAAAAGCAA
GACCAGACATAACGTGAATCATCCTCAATATGGAACCAAATTAACCTAATACATGCCA
TTACCAACCACAGTTATTCCATCACTGTTGTACATAAAAATCTTTCCACCATATCTTCC
CGTGTTCGGAATCGAATTCCC GCGGCCCGCCATGGCGGCCGGGAGCATGCGACGTCGGGCC
CAATTCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGA

CTGGGAAAACCCTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCAG
CTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAA
TGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGTGTGGTGGTTACGCGC
AGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCC
TTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGG
TTCCGATTTAGTGCTTTACGGCACCTCGACCCCCAAAAAAGTTGATTA-GGTGATGGTTC
ACGTAGTGGGCCATCGCCCTG-ATAAACGGTTTT-CGCCCTTTGACGTTGGG-GTCCACG
TTCTTT-ATAGTGGACTCTTG-T-CAAAGTGGAA-AACA

GTJ10-52/17

GTGACACTATA--T-C-TTCAAGCATAGA-AATCCAACGCGTTGGGAGCTCTCCCATATG
GTCGACCTGCAGGCGGCCGCGAATTCAGTAGTGATTAAGCCCGAGGGACAATACACCAAC
TTATAGTCATCCTCATACTTCTCAATCTTGAACCAGTTATCTATGGTTCCTTGCCCCGGGC
CTTCCCTAATATTCCGCGGGTATACAAATAACGTTGCCCGTTAAACTGGTTGATCTGCCAA
ACCGTGTGGGTGCACATGTGACTTGCTCATTGAACTTAATGTTAAGATCAACATTTTCG
CGAATGATGCCATCTTTGTTGGCGGATATAAAATTCAAGGGTGACCCGTTTCTCCCTTCA
CTAACTTCTTGTACCACTTCACGTGGGCATAGTTCGCCTGGTCCCTCGGGCTTAATCGAAT
TCCCGCGGCCGCCATGGCGGCCGGGAGCATGCGACGTCCGGGCCCAATTCGCCCTATAGTG
AGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCG
TTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAG
AGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCC
CTGTAGCGGCGCATTAA-CGCGGCGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACT
TGGCCAGCG-CCCTA-CG-CCCGCTCCTTTCGCTTTCTTCCCTTCTTTCTCGCCACGTT
CGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGG

GTJ10-52/25

GTGACACTATA-TTCC-CTCAAGCAAAAAA-CCAACGCGTTGGG-TCTCTCCCATATGG
TCGACCTGCAGGCGGCCGCGAATTCAGTAGTGATTAAGCCCGAGGGACAATACACCAACT
TATAGTCATCCTCATACTTCTCAATCTTGAACCAGTTATCTATGGTTCCTTGCCCCGGCC
TTCCTAATATTCCGCGGGTATACAAATAACGTTGCCCGTTAAACTGGTTGATCTGCCAAA
CCGTGTGGGTGCACATGTGACTTGCTCATTGAACTTAATGTTAAGATCAACATTTTCGC
GAATGATGCCATCTTTGTTGGCGGATATAAAATTCAAGGGTGACCCGTTTCTCCCTTCA
TAACTTCTTGCACCACTTCACGTGGGCATAGTTCGCCTGGTCCCTCGGGCTTAATCGAATT
CCCGCGGCCGCCATGGCGGCCGGGAGCATGCGACGTCCGGGCCAATTCGCCCTATAGTGA
GTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGT

TACCCA ACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGA
GGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCC
CTGTAGCGGGCGCATTAA-CGCGGGCGGGTGTGGTGGTTAC-CGCAGCGTGACCGCTACACT
TGGCCAG-GCC-TA-CGCCCCGCTCCTTTTCGCTTCTTCCCTTCCTTTT-TCGCCACGG
TTCG-CCGGCTTTCCCCGGTCAAAGT

GTJ10-52/33

A-A-TC-CTC-AGAAA-GAAACAACGCG--GGTTC-TCTCCCATATGGTCGACCTGCAGG
CGGCCGCGAATTCAGTAGTGATTAAGCCCGAGGGACAATACACCAACTTATAGTCATCCT
CATACTTCTCAATCTTGAACCAGTTATCTATGGTTCCTTGCCCCGGGCCTTCCTAATATTC
CGCGGGTATACAAATAACGTTGCCCGTTAAACTGGTTGATCTGCCAAGCCGTGTTGGGTG
CACATGTGACTTGCTCATTGAACTTAATGTTAAGATCAACATTTTCGCGAATGATGCCAT
CTCTGTTGGCGGATATAAAATTCAAGGGTGACCCGTTTCTCCCTTCACTAACTTCTTGCA
CCACTTCACGTGGGCATAGTTTCGCCTGGTCCTCGGGCTTAATCGAATTCCC GCGGCCGCC
ATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAA
TTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCA ACTTAA
TCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATA-CGAAGAGGCCCGCACCGA
TCGCCCTTCCCCAACAGTTG-CGCAGCCTGGAATGGCGAATGGGACGCGCCCTGTAGCGG
CGCATTAAAGCGCGGGCGGGG-GTGGTGGTTACGCGCCAGCGTGGACCGCTAC-CTTGCCAG
CG-CCCTAGCGGCCCGCTCCTTTTCGC-TTCTTCCCTTCCTTTCTCGCC-CGG-TCGC
CGGCTTT-CCCCGTCA

GTJ10-54/1

GTGACACCTATA-TT-CCTTCAAGCAA-AAAATCCAACGCGTTGGGAGC-CTCCCATAT
GGTCGACCTGCAGGCGGCCGCGAATTCAGTAGTGATTAAGCCCGAGGGTAAGTAATTCAT
CAATTAGCTTCGTAATTTAAATTGATATGCTTGCAACTATTTTATTTTGTGTGATCACCG
CTTGATTCTCACTGGATTAGCAGCCTTCCGTTAGCTACCACACAAGCAACTGTTGTTTCGC
CTCGGGCTTAATCGAATTCCC GCGGCCCATGGCGGCCGGGAGCATGCGACGTCGGGCC
CAATTCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGA
CTGGGAAAACCCTGGCGTTACCCA ACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAG
CTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAA
TGGCGAATGGACGCGCCCTGTAGCGGGCGCATTAAAGCGCGGGCGGGTGTGGTGGTTACGCGC
AGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTCTTCCCTTCC
TTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTA-GG
TTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAA ACTTGATTAGGGTGATGGGTCA

CGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCT
TTTAATAGTGGACTCTT

GTJ10-54/2

AGATTC-ACTCTAA-AG-GAAAAACGCG--G-TC-TCTCCCATAAG-AAAA-CTGCAGG
CGGCCGCGAATTCAGTAGTGATTAAGCCCGAGGGTAAGTAATTCATCAATTAGCTTCGTA
ATTTAAATTGATATGCTTGCAACTATTTTATTTTGTGTGATCACCCTTGATTCTCACTG
GATTAGCAGCCTTCCGTTAGCTACCACACAAGCAACTGTTGTTTCGCCTCGGGCTTAATCG
AATTCGCGGGCCGATGGCGGGCCGAGCATGCGACGTCGGGCCCAATTCGCCCTATA
GTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTG
GCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCCGAGCTGGCGTAATAGCG
AAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGC
GCCCTGTAGCGGCGCATTAAAG-CGCGGCGGGAGTGG-GGTTACGC-CAGCGTGACCGCTA
CACTTGC-AGCGCCCCTA-CGCCCCTCCTTTTCGCTTTCTTCCCTTCTTTCTCGCCA-
GTTTCG-CCGGCTTTTCCCC--TCA-GCTCTAAATCCGGGGGC-TCCCCTTTAGGGTTCCG
GATTTAGTGCTTTTACGGGCACCCTCGAACCCC-AAAAACTT-GATT-A-GGGTGATGG
G-C-A--TA-C-GGGCCATCGCCC-TGAATA---CGGGTTTT-CCGCCCT-TGGAC-GT
TGG-AGA-CC-C--TTCCT-TTAA--AA--GGGAC--CTTGTTTCC

GTJ10-54/3

GTGACACTATAGA-T-TTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTC
GACCTGCAGGCGGCCGCGAATTCAGTAGTGATTAAGCCCGAGGGTAAGTAATTCATCAAT
TAGCTTCGTAATTTAAATTGATATGCTTGCAACTATTTTATTTTGTGTGATCGCCGCTTG
ATTCTCACTGGATTAGCAGCCTTCCGTTAGCTACCACACAAGCAACTGTTGTTTCGCCTCG
GGCTTAATCGAATTCGCGGGCCGATGGCGGGCCGAGCATGCGACGTCGGGCCCAAT
TCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGG
GAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCCGAGCTGG
CGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGC
GAATGGACGCGCCCTGTAGCGGCGCATTAAAGCGGCGGGTGTGGTGGTTACGCGCAGCG
TGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCTTTCT
TCGCCACGTTTCGCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCC
GATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGG-TCACGTA
GTGGGCCATCGCCCCTGATAAACGGTTTT-CGCCCTTTGACGTTGGAGTCCACGTTCTTT
AATAGTGGACTCTTGTTCCAA-CTGGAA-AACACTCAACCCTATCTCGGTCTAT-T

UNEDITED FORWARD SEQUENCES

ATJ4-02/1

TCACGTATAGGGCGAATTGGGCCCCGACGTCGCATGCTCCCGGCCGCCATGGCGGCCGCGG
GAATTCGATTCCGAACACGGGGAAACTTACCAGGTCCAGACATAGTAAGGATTAACAAGC
TTAATTCAACAAAGGACAGGGAGCAGAACTTGCTGGATCTTACTATTACCTGGATGACT
AATTTTTCAATTCCAGTTTCATGTTACCCTACATGACCCTTGGGTATCACTTAGCAACTG
TGTAAGCAGAGTCACTCTTCTATAGACACTTTGGTGCATCCACCACCAAGACTAGTGG
AACGTGGCCGTGTTCCGAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGG
GAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCT
TGGCGTAATCATGGTCATAGCTGTTTCCCTGTGTGAAATTGTTATCCGCTCACAATTCCAC
ACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAAC
TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGC
TGCATTAATGAATCGGCCAACGCGCGGGGAAAAGGCGGTTTGCGTATTGGGCGCTCTTCC
GCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTTCG-CTGCGGCGAGCGGTATCA-CT
C-CTCAAAGGGGGTAATACGGGTATCCC-AAAATCAGGGGGATACC-CAGG-AAAAA-
CATGGGGACCAAAGGGCCACCAAAGG-CC-GGAACCGTAAAAA-GGCCCCGGTTG-CT
GGGGGTTTTT-CCATA-GG-TCC-CCCCC

ATJ4-02/2

AG-A-GG-AAGG-AGGAGGTTCCCTTCG---AGGGAAAAA-GGG-GCGGCGTTTT-CCCC
CCCGGG-AAG-AAAAC-AGGAACCG-G-CC-CTT-AAGGG-ACGT-TCCAG-CAA--CA-
TTTTTCCC---GATG--GTA-AC-G--C-GTAGT---GT-T-T--GC-AC-ATC--CAGA
C-CGAC-A-C--CA-AG-GCAC-CG-CTC-GGG----CAG--AA-GGT--GT-AGT--AA
-AAC--CCGG-G---GCGCCGC-C-C-AG--CC-C--GACC--T---CT-AACGT-T---
GGGC--AAC-GA-AC--GAGA-----TGT--A-GGC-GA-G-G--GATGAAAG-GA-GG-
A-GC---G-G-AG---C-CAACCC--G-TAAA---AA-GT-AGTTTT-TATT-A-TAT-T
G-T-A-T-C-G--C-TCG--G-G-G--A-AGGAG----AG--G-CT-T--A-GG-A--A-
GGT--C-A-----GCT-GTA-G--A-GC--ACCT---CAA-GA-A-TC--G-GAA-TGG
A-G-AG-CC-G-A--GT-TAT-A-G--A---A-CGTGACTCG--GC-AGTTAG--C--AT
AC-----GCGT-GTA--TGT-GACGG---GAT-G---AG--AT-AT--G--A--TGT-AT
G-G---A--AC-GAG-T--GATC---T-ATG---T-AGA-AGG-T-A-C-----T---AC
GGCAG-AGA-ACT-----GA-GCAG--AC--AA-AA--A-GGT-AGC-A--C-----G
TAACA-T-AGA---T---C--G--TCCG-A-AG-GC--G--C--C-AAG--T-GG---GT
ATGCC-AT-AG-TAAGGAG--TGGGGA--GA-GGA---C-GCCGGGTCG--GA-T-C---

----C--A-CTCAAGAGT----A---GA-A----TT-TG-GTG--AAG-AGA---GTACT
-C-T--GT-GGGC--A--AGAGC--T--AG-TA--CAC--GTA--TATG-G-AC-GA--C
--G---A-GCT-GG-T-G-G-CA--C--A-C---AG---CGT-G----T-AT----G-AT
GC-CGC-CA-A-GT--T-GGTCATGA-GAC--GG-C--T--G-GCT-CG-AAGG---GT-
CA-TAG

ATJ4-02/3

TCACTATAGGGCGAATTGG-C-CGACGTCGCATGCTCCCGGCCGCCATGGCGGCCGCGGG
AATTTCGATTCCGAACACGGCCACGTTCCACTAGTCTTGGTGGTGGATGCACCAAAGTGTC
TATAGAAGAGAGTGA CTCTGCTTACACAGTTGCTAAGTGATACCCAAGGGTCATGTAGGG
TAACATGAAACTGGAATTGAAAAATTAGTCATCCAGGTAATAGTAAGATCCAGCAAGTTT
CTGCTCCCTGTCCTTTGTTGAATTAAGCTTGTTAATTCTTACTATGTCTGGACCTGGTAA
GTTTCCCCGTGTTTCGGAATCACTAGTGAATTCGCGGCCCGCTGCAGGTCGACCATATGGG
AGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTACCTAAATAGCTT
GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA
CAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTG-CCTAATGAGTGAGCTAAC
TCACATTAATTGCGTTGCGCTCACTGCCCCGCTTTCAGTCGGGAA-CTGTCG-GCA-CTG
CAT-ATGAATCGGCCAAC-CCCGGGGAAAG-GGGTTTGC-TAT-GGGGCCTTT-CGGTTT
CC-CC--C-CTGA--CCCTGGCC-CGGGCGTTTCGGTTGGGGGAA-GGGATTTTCGC-CCC
C-CAAGGGGGG-AA-A-GG-TT-CCC--AAAT-CGGGGGT-ACCCCGGGAAAAATTTT-
GGCC-AAAGGCCCCAAAAGGGC--GAACCCC-AAAGGGCC-GGGGTTGGGGG-TTTTCT
AAAGGG-CCCCCCC--GGGGGAATTCAAAAAAAACCCCCCTTC-AA-GGGGGG-AACC
CCCCGGC-TTTTA-ACC

ATJ6-01/38

GGGGGGGGGGGG-TTC---TTCGGGAAAG--AAACG-GGCCCGCGTTTTCCC-CGGGAAA
GAAAAAAA--C-ATTC-G-TGAGA-A-ATGA--G-A-A---AA-AAGTTTGTTCCCCCG
--A--GGG-GGA--A-GAT--C-C-CAG--T--G---AA-----C-TC--CG-GA-TA-
-TG-A--A-GTG----AGACAATG-AAATT-GG-CCA-AG--AG-A----GA--GG-TCG
--ACT--T-TG--T-A--TGA-GTAA-G-AG---T--T---CG---TCA--G-AGCAAGA
-GC-G--CA-GCTAG-G-AG-G-A-ATG-A-TGTGAG-C--GGCA--C-GC---AGC--C
G-AAT--GAGGGA--AGT--GG-T-TTTTTT-T---G----AAGATG-A--AGC--AG-A
ACT---TTGCCG-GTG-TGTG-T-AGAC-TTAT--G-AGAGTGGA ACTAT-TA--G--G-
G-TGAG--C--G---ACTAGTCTC--T-----CG-C---CCAGC-TCG-AGA--CCCCGA
CG-T---TTGGAG-A-CGA--AGG-AGACG---GCAGGG-AAC--GTACAT--A----AA

-CGGGAC-----GG-AG-----C--CTCAG-CTC--A--T--AA---GCAC-CGGTC--C
T-A-ATGCGGA----A-TG---TC-TAAG-G-C-AAAA---A-CCG--G--AC--TA-A-
-GCA---ATG-AAG-GACT-ATGATGCC-----AACGA-TG-GT--GGGGCAAG--G-A-G
AAAATGAC-A--AGCAAC---ATGCAACA-AG---G--GCCAG-A--GATGGAT----GA
-GGAAT---GG--CTG--C-GG-GG--AG-AGGGAA-GGAAAA---GC-GC--GGC-GTA
-AAGA-A-GCGTA-TGCG-AGTG-AAG-AGCA-G--T-T-T--AG-AAGG---AG--T-T
AA-CAGT-TAC-T----TGGC-T-AG-TACTCC---CATCT----AAGAGCCACT--TAC
-TA--GTCGTAAG-A--GG-CAG-GGATT-T-TATG-AC--CAGCGA--GC-ATA-GTT-
A-AGATGA--C--G-GAGTGT-CAA-G-C--A-CG-GA-GGCG--G-CATA-GC-AGAT-
--GCGA-G--CG--A-GG--GACT-T--CAGAGGG-G-G--C-GGTCATGGCG--CGA--
----AG----GTA-T-A-GT-CT-GA-CA-CA-AGAT-GAATCGG----T--TC-CACC-
GAT-TGATCACAT-AG

ATJ6-01/48

-GGGGGGGGGTTTC--TC---AAAA-AAAA-GGG-----TTTTCC-CCGGA--A-AAAAA
TATGTT--AGA-TG--AA-GGAT-TGGAT-TATCTTATTTTT-CCCCCTA-TA-TT--AA
-A-G-TCT---CAC-GTTGT-----TTAA-G-G-TATTGAA--T-A--A-TAT-TAAT--
-ATATGTC-T-TGT--ATGTG-GTA-T-T-TAT-TT-TTA-GT-TGAA-T-G-GTTCT--
A--TTTAAG--AA-CTC-----GAG--GGAT---AC-----
-G-T-GGTGAA---T-TTTTTTT-ATG--A--GGG-T--AACG-AA-TTGT-TTT-ATGA
---GT-TTG-GGTAG-AA-GA-TT-TTATAGTGATT----GT--AGGT--AC--CG-TC-
TTC-G-GTTGAT-TT-A-TG-GGA---TGT-TCAT-ATGGTA-CC---TTTCA---ACT-
-G--GGTAGGC--TGG-A----AGATA--TGT-TATA-GA-GTT--GTAAA-CGATG--A
AATA--G-TT--TA-A---AGATAAT-ATAT-T---T-TTGC-GA--TGGAGGTG--AAT
AGT-GGT-GA-T-TAAA-GA-----GGGAT-T-TGG-AACTATAA--AGA--GAA-GGAT
GGTATACTT--T-T-TT--T-A-C---ATTGT--G-AT-CG-G-GACTCA-TCAG-GGTA
-ATT-A--GGTGG--GGA-CTG-AT--T-GAG-GA-AATGAA--TTT-AACG-CGA-G-G
-TG-----TATG-CATGC-GGC-ACT-TTA--T-T-C-TAT-T--C-TTCA-GATACTT
-T-A-A-TGG-CAATATC--C-T--GGT--AT-T-TG-ATTAAGTA-TG--GTTG---TT
A--AT-ATGCT-C-G-TGTTATGAT-ATTC---T-G-ATGT-GCTC-TTGA-AAGAGA--
--T-TCC-A-TG-C--CCAT-TG---TGTAT-ACTGTA-GGACT--A--T-GTA-CGCAT
AA-AGCTTT-TATT-ATGACCAGTCCGGCGAT-CAGCA-GGTG-TAA--TG-GTATATAG
-TAGCT-TCTCTA-TT---CAC---CTTATG-ATAGAGTC-TCTTAG-A-CGAT--AGCT
A-ACAATA-ACG

CTJ4-24/5

CGCTCC-ATAGGGCG-C-TG-GCCCGACGAC-GATGC-CCCG-ACGT-ATG----AC--G
--CAATTCTACTC---GCAC-GCTCCACTTAGATGCACTTGGCGCC-AT-A-TCCCTCAG
AATA-CATCACTGTCCTATCC-TGG---ATT---TGGTGC-TCC-TCTACCGAGATT-C
C-TCTATTCAGCTGATGCTGAT-ACT-G---AC--CGG--GCTCGTCTAAAGTGA-GC-C
CA-GACTTGT-ACAGGT-TG--G-AGGAG-ATATTCTTGAGAATA-ATG-GCG-CA--TG
-A--TCAGG-GG---TGC--C-GAATC-AAT-CC--CT----C---G-G-GACG-CA-GC
ATGCTCACGTACTTACCCG-GT-CG---GTAT-CTGCTGCTGTA-TA-CATG-ACTGGCC
GTCGTT--AC--CCTAGTGACTGCGCAAACCCC-GCGT-ACC-C--TATG--A---T-C
-CCAC-T-CC--AT-C-T-A--TGA-GAT---ATAGA-GA-GCC---A--GATCGGCCT-
C-CATG-GCA--GC-GCCT-CTGGG-GAATG-GC--TCCCTGCACA-TCC-TCAACAT-C
GAGGC-GTAGCGT-AAC-GT-AA--CTGCCGC-ACTATTG--AGAGCCAACCGCCCGGAT
CTGTGC-GTTGCTTAC-GTCC--TT-TCCACACGGTAAACC-G-TTTGCCC-GCTGGCTT
TA-AAT-GGGCTCCCTTCT-GGGAACGA--T-TTGCTTT-GGGGACCTCTACCCCTTA-A
-CTTAAT-A-GTG-AGGGTTCCGGT-AGTGGG-CT-CCGCCAGAAGAAA-CGGTTT-CC
CCCC-TGGGG-G-ATTGG-GGCCA-CCCCCTT-AAAAGGGGA-AC-CGGGGGCAAAA-T-
--GGAACCCCAA-GC

CTJ4-24/15

TACTATAGGGGCGAATTT-C-----GACG-TA---GTGATCCCGGCCGCCATGG-GGCCGC
GGGAATTTCGATTCCGAACACGGTAAATGAAACTGAGGAGAAGGAGCCTGAGGGTAAGGAG
ATGACTCTTGAAGAATATCAAAGGTTCTTGAGGAAAAGAGGAAAGATTTAGAGACACTC
AAGACTGAAGAAAAGAAAAG-GGAAG--GA-A-AGAGGTTGCATCTATGCAAC-A-TCTCC
AA-AAGAAGAGCAACGATGAAAT-TTTG-TAAA-TGGGGACTG-A-AA-GATAAG-TGT-
-AGAGCTTGCTG-TTTTTTT-AG-GAG-CAAGAAG-C--TC--T-TCGT-TG--T--T-G
-AAAGGTGCTGCTGGC-GAAAGAAA--TACAC--CTGGA-GCG-GGG-GG---GCG-GA-
AG-GTATAGGAA-A--AA-TAG-GTAAT--AACAG---CG---GTGAAA-GTT-TA-GGA
TC-CCGG-AG----GT-CCC--G-GAGT-TGGGC-GATAT-AAT--AGAAG--GATGCTG
AGA-T-G---A-CA-AAA-GTG-TAT-G---GT-AA-TC-TT-GCGG-AT-A-CTA-G--
GTG-CGGGA-GGT-GA--CTGGG-AT-A-TC-T-T-TA-G--TA-GC--GT-AGG--A--
----C-CATGGG-GG-----CAGTAA--GTGGACG-ACCC--TGG-G-G---GTGC-T
TA-ATAGA-GT---AAAT-AAAGGGAA-CCTTG-AG-TAG-A--TT-T--GG-TTAT--G
---AG-GT--TTGTAG-G-T-GGAGAACAGT-GT-AG-GAG-TGA--GA--T-GA-TA-G
TAA-TGATG-TAATCG-GTGTAGGA-G-ATAG-TGG---TT-C-AAA-GA--AGA-ACAC
TG-ATC--T-GATG--AG--AGAGGA-CAC---AGGGCGG-CGTGA--A-AT-G--A-A-

C-TGGACA-GGGAGC-GATA---ATGC-CTA-TGAG-GA-C-A-AGA-AAGAG--GTCTC
T-CTTAGC-AAGCG-GA-C-CGAG--G--A-G--ACATT--A---TG

CTJ4-24/43

GSTATCGACTCCTATAGGGCGAATTGGGC-CGACGTCGCATGCTCCCGGCCGCCATGGCGG
CCGCGGGAATTCGATTCCGAACACGGTAAATGAAACTGAGGAGAAGGAGCCTGAGGATAA
GGAGATGACTCTTGAAGAATATCAAAAGGTTCTTGAGGAAAAGAGGAAAGCTTTAGAGAC
ACTCAGGACTGAAGAAAGAAAAGTGAAGTTGACAAAGAGCTTGCATCTATGCAACAACCT
CTCCAACAAGAAGAGCAACGATGAAATCTTTGCTAAACTGGGGACTGACAAGGATAAGCG
TAAAGAGCTTGCTGAGAAAGAAGAGAGAGCCAAGAAGTCACTCAGCATCAATGAGTTTTT
GAAACCTGCTGCTGGTGAAGAAACTACACTTCTGGTGGAGGTGGACGTGGCCGTGTTCG
GAATCACTA-TGAATTC-CGGCCGCCTGCA-GTCGACCATATGGGA-AGCTCCC-ACGCG
TTGGATGCATA-CTTGAGTATTCTATA-TGTCCCCTAA-TACCTTGGC-GTAATC-TGG
-TC-TA-CTGTTTCCTGT--GA-AT-GTTATCCG-TCAC-TTCCCCCCCACC-T-CCAC
CC-GAA-C-TAAAG-TG--AAA-CTGG--GTGCCCTCTTG-CTGGACTAA--TC-C--TA
C-TTT-C-GTC-CCC-C-C--CCCCCTTTCC--TCG--AA-CC-TG--C--CCC-CT-CT
TTAATGA-TC-CCC-CCC-CC---C-A-TGG-T-TT-TCC--TTTG--C-CTTTTTCCC-
TT-C-CC-TT-CTT---CCTCT-CC-CCC-C----T---CCTTCCCT-C-CT-G-T-T-T
-CTCCCTC--A---C-G---CTAC---T-TCCCC-C--TT---G-G-AT--CCCC---
ACCCTCTC-C-C-C--CG-C-

CTJ4-26/3

GGCGAATTGT---CGACGTCGAA-GCTCCCGGCCGCCATGGCGGCCGCGGGAATTCGATT
CCGAACACGGGGAAACTTACCAGGTCCAGACATAGTAAGGATTGACAAGACATCTCTACG
AGGCGGCTTCTTTTCGCCATAGCAATAACTGGCATTGCCATTCTTATTGGGTCACTTCTC
TTTCAAAGAAAACCAGGCATCAGTCAAGTTAACTGAAAAGTCAAAGGTCAACCTGGAGAT
TTTAAGAAGAGTTTACATGTGTCAAATGTATGAATATAAATATTTGATTTTCAATTTACA
TCATCCCAGACATTTCTGTTGATTATTCTGCCGTGTTCCGAATCACTAGTGAATTCGCGG
CCGCTGCAGGTGCACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTAT
TCTATAGTGTACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAA
TTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTG
GGGTGCCTAATGAGTGAGCTAACTC-CATTAATTGCGTTGCGCTCACTGCCCGCTTTCCC
GTCGGGAAACCTGTCGTGCCAGCTGCATTAATG-ATCGGCCAACGCGCGGGGGAGAGGCG
G-TTTGCGTATTGGGCGCTTTCCTCGCTCACCTGA-CTCGCCTGCGCTCCGG
TC-GTTC-GGCTGCGGG-CGAGCGGGTATCAGCTCCCTC-AAGGGCGG--ATAC-GGTT-

CCCC-GAAATCC-GGGGA-AACCGC-GGAAA--AA-TG-GAGCCAAAAGGCC-GCCAAA
-GGCCAGGAACCCG-TAAAAAGGCCCC-TTGCCTGGG-GTTTTTCCATAGGCTCCCCCCC
CC

CTJ4-27/16

GGAA-GGCGAATGTTCCCTGAG-AA-AAAAAT-C-GGCAGC-ATTTCCGCCGCG-GAAAA
AAATT-CGAAC--GGGGCCAGAGGTTGATACACAGC-ATAAGGA-GCCCT-A-GAGAA-G
TATGAAAGTTGC-GACAGCGTG-AGACTCC-AA-CCCCCT--CTTATTT-GGC-AA-TTG
TGTGG-GGGTT-ACT-GG---T-GTGGG----GG-TATGA-TATGGATA--GTT-GCA-G
AAA-TGAA-G-GAT-T--A-AG-TTCAAGAC-TTTGGATGA-GAC-AGAA-A-GG-TG-
C---TT--AAT-T-TTTTTTTTTG-GAG--AAA-GAAA-GATC-GAAT-GAATTA-T-GG
-AGG-G-TGCGAAGG---AAGT-T--GGGAG-G-A-AA--GA---G--GGG---GGA-T-
AGT-G-GTA-TC--AAGG--GCT-ACA----TTAGAT-TGGGAGTAAGTCAT-GCC-AGT
AG-TGTG-T-CAA-GG--GT-AATCGT--TT---GCTA-AT-A-TTA--A---AG-T-GA
-A-A-GG--GG---GC--TGA-GA-T-A-A--A--CT-GGGGG--C-GTTG-GA-A--AA
---TAC--A-AT-TA--AGGA-TG-CGG--A-GAATGA-TGT-T-AC-TT-TCAAGA--G
-GGCGA---GAAAG-GG--A-A-GA--CTT-T--CG--AAA--GATGAGAAATGTGGGGA
GGG-G-ACGG-ATTG----A-TGGCGA--GAT-TAGA-G-G----ACG-C-C--GG-A--
-GAGT-GA-G-ACACG--GA-TGTGA-G--AGG---GGAG-GCAT-G-AACG-ATCTAAT
-G--GAGGGAGCGAAAC---GAG-TA-A-T-TAG-A-AA-A-GAGT---GA-TCAT-C--
A--AGAA-A--AGG-CT-GAGAGGG-CGAGAGTTA-CGGTA-TA-AG-CG-A-T-C-G-G
GTTA--GAGAGGAGTAGG---TTTGA-TTGATA-CTA-GGTT-TG----GT-C--GAGTA
-GC--GGAG--TTA--GTGC--G

CTJ4-27/30

CTATAGGGCGAATTG-TC-CGACGTCGCATGCTCCCGGCCGCCATGGCGGCCGCGGGAAT
TCGATTCGGAACACGGGGAAACTTACCAGGTCCAGACATAGTAAGGATTAACAAGCTTAA
TTCAACAAAGGACAGGGAGCAGAACTTGCTGGATCTTACTATTACCTGGATGACTAATT
TTTCAATTCCAGTTTCATGTTACCCTACATGACCCTTGGGTATCACTTAGCAACTGTGTA
AGCAGAGTCACTCTCTTCTATAGACACTTTGGTGCATCCACCACCAAGACTAGTGGAACG
TGGCCGTGTTTCGGAATCACTAGTGAATTCGCGGCCCGCCTGCAGGTTCGACCATATGGGAGA
GCTCCCAACCGGTTGGATGCATAGCTTGAGTATTCTATAGTGTACCTAAATAGCTTGGC
GTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATCCACACAA
CATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCCTAATGAGTGAGCTAACTCAC
ATTAATTGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAAACCTGTCTGCGCCAGCTGCA

TTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCATATTGGGCGCTCTTCCGCTTC
CTCGCTCACTGACTCGCTGCGCTCGGTTCGGCTGCGGCGAGCGGTATCAGCCTC-CT
CAA-G--G-AATACGGTTATCC-CAAA-TCAGGGGATAACCCAGGAAA-A-C-TG-GACC
AAAAGGCCA-CAAAGGCCAGGA-C-TAAAAA-GGCCGGG-TGGT-GGGGTTTT-CCA-A
GGTCC-CCCCCCTG-AA-AAT-CCAAAAT-C-GCTC-AAGTC-AAGGGGG--AACC

TTJ6-53/2

TCCTATAGGGCGAATTG---C-CGACGTCGCATGCTCCCGGCCGCCATGGCGGCCGCGGG
AATTTCGATTTTCGTTCCGCACAAGGAATGGCTGGTTCGCTTATATCAAGCTCTAAAAGGCTA
TCGTATTTTCTTAGACACAGAAGCCGATTTTCGATTTACACGATTTGAAAGAGATCGTAAA
ACATACTAAATTGTTTGTTTTATACTTACGGAAGGCATTTTGTAGTTTTTGGTGCTT
ACAAGAGCTCAGAGCTGCTGTTTCGTACAAAGTTCCTGTTGTAATTATTAGAGATATTAA
ATTCCAATATCCAAATCCTTGGCCTGCGGAACGAAATCACTAGTGAATTCGCGGCCGCCT
GCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATA
GTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTA
TCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGC
CTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTCCAGTCGGG
AAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTG
CGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTTCGGCTG
CGGCGAGCGGTATCAGCTC-CTCAA-GG-GGAAAACGG-TATCCC-GAATCAGGGGATAA
C-CC-GGAAA-A-C-TGGGAGC-AAAGGCC--CAAA-GGCC-GGAACC--TAAAAAGGGC
CGGGTGCTGGG-GTTTTTCC-TA-GG-TCC-CCCCC-GGAA-AAATCC-AAAAATC

TTJ6-53/3

CTATAGGGCGAATTG-TCTCGACGTCGCATGCTCCCGGCCGCCATGGCGGCCGCGGGGAAT
TCGATTTTCGTTCCGCAGGCCAAGGATTTGGATATTGGAATTTAATATCTCTAATAATTAC
AACAGGAACCTTGTACGAAACAGCAGCTCTGAGCTCTTGTAGCACCAAAAATATCAAA
AATGCCTTCCGTAAGTATAAAAACAAACAATTTAGTATGTTTTACGATCTCTTTCAAATC
GTGTAAATCGAAATCGGCTTCTGTGTCTAAGAAAATACGATAGCCTTTTAGAGCTTGATA
TAAGCGACCAGCCATTCCTTGTGCGGAACGAAATCACTAGTGAATTCGCGGCCGCCTGCA
GGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTG
TCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCC
CCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTA
ATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAAA
CCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCAT

TGGGGCGGCTCTTCCGCTTCCTC-CTCACTGACTCGCTGCGCTCGGTTCGTT-CGGCTGCG
G-CGAGGGGGTATCAGCTCCCTCCAAAGGGGG--AATCCGGTTATCCCCGGAATCC-GG
GGATCACCCCAGGGAAA-A-CATGGGGG--CAAAGGCC-GCAAAAAGGCCCGGGAA-C
CGTTAAAA-GGCC--GTTTGGTTGGG-GTTTTTTCCA-AAGGGT-C--CCCCCCTTGA-
AA-AAAATCC-AAAAAAT-G

TTJ6-53/50

GG-GAGGG-AG-GG-ATTTTC----TTTTAAA--A----AAAAA-----C----CCTTT
TTTC-CCCA-CGAGAAGG-AGAAAAAAA--GG--G--G-AGG--CTTT-AAAGAGGTG-
---GA-C-AC-AA-T-ATT--TTTTCCCC-GGGGGG-GGT-G-CA--CAACCCGACCGC
TAC-G--CTCG-CGC---TGGAAGAGGAGAGGG--AGGA-GAAAG-GGGAAA-AA-C-G-
CA-A-CTGACA---AG-A-GGACCGGA--G-T-G---G-A--G-CGG-TG---GA-CG--
ATG-TACGAGCAAGGG-GAGAGAGACACGG-GG--GG---AGAG--G-G-AGAGA-GGA-
-C---A-GATTG-G---GA-A-GA-G-G---TGA-AAG-AAACGAGCGGG--GTG-AGC-
--AA--AAAAA-GAGAG-AA-A---AA--ATTTTTTTTT-TA-GACC---CG--A-CCA
GACAC--CAAAGT-CT-GG----GCGAAA-GA-G-CTGTGG-AGG-GAC-A-GAG-G-AA
CGAC-TG-GG-G-T-GC-G-GGA--AGT-GA-G-GGGCT--G--AA--GGAAGGCGAGG-
GGAG-TAAGGAG-TGGA-TTAAA-ACAGGTGAGGAGGGG-GC-GGA-GGT-G-GG-AC--
AGGAAGA---CG--C-A-AAGA-CG--G--AAAAGAA-A-GG---ACAG-AGA-A--GCA
---G-A-GCAG-AAGGA-GAGAG-A---AG--GAG--GACAAAGCGG-AGAAA-GGGA--
A--A--G-AGAG--G-GA-AGA-AGG-GAAG-A--AAA--TAAGAGAC-ACGG-G--AGG
GA-G-C-CACGC--AAAGATAA-CCGC--AG-A---GGCT-ATG-GG-TATA--TAGA--
G-A-CGA-AAAGAC-AG-A-ACATGAT--A--GGACGT-AGGGC-GGAGCACCAA--G-A
TA-GAGAG-A-AGTGG-GGC-AGT-AAT-AGGAGGAA--GTATCA--TG-AAG--AGGG-
ATGAAAAGACAGAA---GG-A-AT-AGGAGGATGAA-GAGAAGGA--GAGAGAG--A-A
G-GAGAT-TCGTAGA-A-GGAGC-GTG-CTA-G--TGG-GGACA-CTGC-A-GATGAGAG
CG-CG--GGT-G-AT-GCG-T-GGA-GATA-GG-T-ATGT--A-CGAGA--TG-GT-T--
AGT-GC-GA-TG-GGAGAACGAGATGA-T-AGA-TAGCGT-GGT-C-C-T-AGGT-CGAT
G--AGA-GTGC-TGCAGTCACAG-C--G-GCGTG-GCAGACAGTG---GGGCG

GTJ10-52/17

GTATAGGGCGAATTTTC-CTGACG-AG-AA-A-CCCGGCCGCCATTTCCGCCCGGGAAA
-AGATTAAGCCCAGGACCAGGCCAACTATGCCACGTGAAGTGGCCCCAAGAAGTTAGT
GAAGGGAGAAACGGGTCACCCTTGAATTTTATATCCGCCAACAAAGATGGCATCATTCGC
GAAAATGTTGATCTTAACATTAAGTTCAATGAGCAAGTCACATGTGCACCCAACACGGTT

TGGCAGATCAACCAGTTTAACGGGCAACGTTATTTGTATACCCGCGGAATATTAGG-AGG
CCCGGGCTT-GAACCATAGATAACTGGTTCAAGATTGAGAAGTATGAGGATGACTATAAG
TTGGTGTATTGTCCCTCGGGCTTAATCACTAG-GAATTCG-GGTCGCCCTGCCAGGTCGA
CCATATGGGAGAGCTCCCAACGC--TTGGATGCATAGCTTGA-ATAT-TCTATAG-GTCA
CCTAAAATA-CTTGGC--TAATC-TGGCTCATAGCCTG-TTTCC-TGTGGTGAAAAATTG
G-TAATCCCGCCTCCACAATTTCCACACA-AC-TTAC-AG-CCGGG-AA-CAATAAA-T-
--AAAGCC--GGGGGGG-GG-A-TAA-TGAACCG-A-TCTTA-CT-CAACATTTAA-TTG
GC-AC-GGCGGTTCA--TGG-CCCGGCTCTTCCAAA---CTGGGAAAA-TCCTGC-ACCT
GCCAACCTTG--ATT-AAT-GAAAA--CGCC--ACGCAC-C-GGGGCATAAGA-----T
CT-C-C--ATTTGGGGGGAG-T-CT-C--GACTTCCCC-CGCTCA--TGG-AACC-GC-T
G-C---ACG-GA---GT-CGGGC--A--GCGCTA---CG---TG-TAAG--TCCACCTCC
AAA-GAGTGGC-AAT-TGCTGG--TTTT-CCCCCCA-AAAAC-CG-GGGGT-TAAC-C
CTC--CG-A-AA-A-AC--CG

GTJ10-52/25

TATCGACTCCTATAGGGCGAATTG-TCCCCGACGTCGCATGCTCCCGGCCGCCATGGCGG
CCGCGGGAATTGATTAAGCCCAGGACCAGGCGAACTATGCCACGTAAGTGGTGCAA
GAAGTTAGTGAAGGGAGAAACGGGTCACCCTTGAATTTTATATCCGCCAACAAAGATGGC
ATCATTGCGGAAAATGTTGATCTTAACATTAAGTTCAATGAGCAAGTCACATGTGCACCC
AACACGGTTTGGCAGATCAACCAGTTTAACGGGCAACGTTATTTGTATACCCGCGGAATA
TTAGGAAGGCCCGGGC-AGGAACCATAGATAACTGGTTCAAGATTGAGAAGTATGAGGAT
GACTATAAGTTGGTGTATTGTCCCTCGGGCCTAATCACTA-TGAATTCGCGGCCGCCTGC
AGGTCGACC-TATGGGAGAGCTCCCAACCGC-TTG-ATGCCAT--CTTTG-AC--TTTCT
CTT---TGTT-CCTAA---TT-CCTTGT-TC-TA-CC-TGGC-C-T-CCT-G-TCTCCCC
-TCTC-CTAATTC-TTATCCCGG-TCTT-CA-T-T-CTC--TTTCT-T-CCCACTCC-G-
CAACT-A-CC-CG-CGCTA--G-TCTTGT--CTTCC-CTA--TC---T--T-GC-CTCT-
C-CCGCCCTACTTT-TC-TCC--AC--T---TT---CTG-T-T-TTC-CTC-CC-CTCT-
-CC---GT--GCGGCTCCA-CTT-C--TC-CATCG--TT-CC-CCCTC-TCCTTC-TC--
TCTG--CTCTTC--CT-CCG-CT-C-TC-CAC----T-CT-CTCTCCACT-C-CCT-TAT
TTC---CCTC-CCC-CTCT-T---CC-ATCCCACTGCTTTTCCACTTCTC-CATCCATT-
-A--T-CA-AGCC--GC--TCTT-CC--CTTTCTCT--A-T-CTCTCTCTCT-TCTG--C
GCTAT-TCTACTCCTCC-CTC--AATAC-C--TC-TTAT--GG-CT-C--CC-TCT-A-A
CTCTG--CGTCCCC-CA-

GTJ10-52/33

GCGAATT-T-C-TGACGTCG-A-G-ATCCCGGCCGCCCTGTC-GCCGCGGGAATTCGATT
AACGCCCAGGACCAGGCCGAACCTATGCCACGTGAAGTGGTGCAAGAAGTTAGTGAAGGG
AGAAACGGGTCACCCTTGAATTTTATATCCGCCAACAGAGATGGCATCATTCGC-AAAAT
GTTGATCTTAACATTAAGTTCAATGAGCAAGTCACATGTGCACCCAACACGCCTTGGCAG
ATCAACCAGTTTAACGGGCAACGTTATTTGTATACCCGCGGAATATTAGGAAGGCCCGGG
CAAGGAACCATAGATAACTGGTTCAAGATTGAGAAGTATGAGGATGACTATAAGTTGGTG
TATTGTCCCTCGGGCTTAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGG
GAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTACCTAAATAGCT
TGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCAC
ACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGACGCTAA
CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAAACCTGTCGT-CCAG
CTGCATTAATGAATCGGCCAACGCGGGGGAGAGGCGGTTTGCATATTGGGCGCTCTTC
CGCTTCCTCGCTTACTGACTCGCTGGGCT-CGGGTCGTTCCGGCTGGCGG-GAGCGGTAT
-AGCTTCCTCCAAAGG-GG-AAA-CGGGTTTTCCCC--A-TCAGGGGGA-AAC--A-GA
AAAAAA-ATGGGG-CCAAAAGGCCCC-A-AAGGCC