Seed dormancy in barley (*Hordeum vulgare* L.): comparative genomics, Quantitative Trait Loci analyses and molecular genetics.

YUMIKO GRACIELA BONNARDEAUX

BSC. Biology (Hons)

This thesis is presented for the degree of
Doctor of Philosophy
The University of Western Australia

The School of Earth and Geographical Sciences (Soil Science Discipline)
Faculty of Natural and Agricultural Sciences
The University of Western Australia

2008

ABSTRACT

Under prolonged wet and damp conditions, barley grain with low dormancy can germinate precociously, a condition known as preharvest sprouting that causes a number of detrimental effects in grain quality. In particular, preharvest sprouting renders the grain unsuitable for malting. The aim of this study was to take a genomics approach to identify and characterise candidate genes that could be linked to the control of seed dormancy in barley. This thesis developed a bioinformatic strategy that exploited the availability of gene sequences with functional evidence in the model species of Arabidopsis and rice. The bioinformatic strategy integrated phenotypic data (QTL data) and comparative genomics for a targeted approach in identifying candidate genes with a high probability of having a conserved function in cereals. This bioinformatic study identified two candidate genes *ERA1* and *ABI2* with strong evidence for a role in seed dormancy based on their function in Arabidopsis in abscisic acid (ABA) signal transduction and their co-location to seed dormancy QTLs in Arabidopsis, rice and wheat.

In order to establish whether the candidate genes mapped to seed dormancy QTLs in barley, QTL analyses were performed on a double haploid population, not previously studied, developed from a cross between Stirling, a major Australian malting cultivar, and Harrington, a major Canadian malting cultivar. This cross was specifically chosen for this study, as elucidation of chromosomal regions associated with seed dormancy in the background of a malting cultivar would make a significant contribution for the malting industry. A genetic linkage map with 128 molecular markers was constructed and QTLs controlling seed dormancy were characterised in the population. A major

QTL controlling seed dormancy and accounting for over half the phenotypic variation (52.17%) was identified on the distal end of the long arm of chromosome 5H. Minor QTL were also detected near the centromeric region of 5H and on chromosomes 1H and 3H. These minor QTL with additive effects accounted for 7.52% of the phenotypic variance measured. Examination of epistatic interactions further detected additional minor QTLs near the centromere of 2H and on the long arm and short arms of 4H. Combinations of parental alleles at the QTL locations in predictive analyses indicated dramatic differences in germination. This study found the combination of Stirling alleles at the two QTL locations on the 5H chromosome and Harrington alleles at the 1H and 3H QTL locations significantly affected dormancy.

Identification of a seed dormancy QTL on the long arm of 3H, in a region syntenic to the wheat chromosome locations of ESTS aligning to the *ERA1* and *ABI2* genes, laid the foundation for physical and genetic mapping of the candidate genes to investigate whether the genes co-located to the QTL on 3H. Physical mapping of the genes in wheat barley addition lines confirmed their positions on the long arm of 3H. Genetic mapping of the *ERA1* gene was performed using a CAPS marker developed in this thesis. The genetic mapping of the *ERA1* gene did not place the gene within either of the minor QTLs on 3HL, although segregation distortion may have influenced the map position of this gene. Further investigation is required to resolve the positioning of the *ERA1* and *ABI2* genes in relation to the 3H seed dormancy QTL.

The main outcomes of this study have been 1) identification of candidate genes for further study; 2) identification of QTLs on the long arm of 3H that were previously unknown; 3) demonstration of the potential differences in dormancy that can be

achieved through the use of specific gene combinations, highlighting the importance of minor genes and the epistatic interactions that occur between them and; 4) the development of a CAPS marker for the *ERA1* gene, which can be used to track the gene in barley breeding programs to observe its association with important agronomic traits. This thesis also pioneered the implementation of several new technologies including multiplex-ready PCR (Hayden et al. 2008) for fluorescence–based SSR genotyping and QTLNetwork (Yang et al. 2008) for statistical analysis of QTLs. Seed dormancy is a complex trait and is likely to involve the interplay of a number of genes that have a role in other developmental and regulatory processes.

ACKNOWLEDGMENTS

Completing this thesis has required the help, support and insights of many people and the acknowledgments could easily be as long as the thesis. To begin, I would like to sincerely thank my supervisors, Professor Rudi Appels, Dr Chengdao Li and Professor Sivasithamparam for their support and guidance. I am especially grateful to Rudi for giving me his valuable time, support and knowledge throughout the project from beginning to the end. Dr Reg Lance and Dr Mike Francki also provided me with mentoring and support.

The MPBCRC funded this project and I would like to acknowledge the education team, Heather, Michael and Amanda, for the great education program they put together that inspired my interest in science communication. The University of Western Australia provided financial support for project operation and conference travel, Kirin Australia kindly provided the plant material and the research conducted using SABC facilities.

I received invaluable support from the CCG group at Murdoch, for the bioinformatics of my project. I would especially like to thank Paula for her incessant time, assistance and delightful company over a wee drink.

In particular, I would like to extend my gratitude to everyone from the DAFWA lab at the SABC, who provided a wealth of knowledge, experience, technical support as well as friendship (Esther, Allison, Dean, Allen, JJ, Karon, Dan, Vera). Frances Brigg painstakingly performed mountains of sequencing for me. Fiona and Sharon had endless patience assisting me with the multiplexing. I thank Reet, my office roomy, for her company and sharing the ‘PhD’ experience. Dora and Meredith were the sources of

advice, guidance and sanity from laughs over the many coffee breaks we shared.

I would like to thank all the technical staff at DAFWA, particularly Kathy and Mike for their help with harvesting and threshing, and the grain lab (Stefan, Marty and Grant) for lending me their facilities to do the seed germination tests. The field trials were also well looked after by the staff at the DAFWA Research Stations. Dean and Katia helped me tremendously with the statistical analyses of the project.

I would like to thank all my friends for supplying me with encouragement throughout my candidature. This thesis would not have been possible without the support from John to whom I am eternally grateful for a list too long to mention here. The last leg of the thesis, was undoubtedly the hardest part for more reasons than one and I am so grateful to Dani for her refuge and friendship. In particular I would like to thank Lu who provided a net of constant and consistent unwavering support, faith and encouragement during this time, and the courage and direction for forging onwards.

Last but in no way the least I want to thank my family. I find I am speechless when trying to find the words to thank them, as there are no words that could do justice to my gratitude. But in trying, I would like to thank my parents for their steadfast support, encouragement, inspiration, love and belief in me, and emergency meals. To my sister Miki, forever my hero and role model, I thank for not only being the greatest sister in the world but my personal support and cheering squad. I thank her and Jason for taking me in and looking after me. I cannot go without including little Tomi, who is totally unaware of all the joy she brings. Finally, in an appreciation of all they have done for me in my life, I would like to humbly dedicate this thesis to my family.

DECLARATION AND LIST OF PUBLISHED PAPERS

I declare that the research presented in this thesis is original and was undertaken and written by myself. The thesis has been completed during the course of enrolment in a PhD degree at The University of Western Australia and has not previously been submitted for a degree or diploma at any other education institution.

A paper that has been published from the research described in Chapter 3 is:

Y. Bonnardeaux, C. Li, R. Lance, X.Q. Zhang, K. Sivasithamparam and R. Appels (2008). Seed dormancy in barley: Identifying superior genotypes through incorporating epistatic interactions. *Australian Journal of Agricultural Research* **59** (6): 517-526.

A section of my Chapter 5 discussion has also been included in:

W. Ma, O. Anderson, H. Kuchel, Y. Bonnardeaux, H. Collins, M. Morell, P. Langridge, R. Appels (2008). Genomics of quality traits. In: (C. Feuillet & G. Muehlbaur, eds) *Triticeae Genomics.* Springer, Heidelberg London New York, Paris, in press.

Yumiko Graciela Bonnardeaux

-------------------------------	--------------------------------
Prof Sivasithamparam Coordinating Supervisor	Prof Rudi Appels Supervisor

TABLE OF CONTENTS

Abstract……………………………………………………………………..………….i
Acknowledgements…………………………………………………………..……..…iv
Declaration & List of Publications……………………………………..………vi
Table of Contents……………………………………………………………………..vii
List of Figures………………………………………………………..………...xi
List of Tables…………………………………………………………….………xvi

Chapter 1: Introduction and literature review…………...………………………..1

1.1 General introduction……………………………………………………….…………...1

1.2 The barley plant……………………………………………………….…………..2

1.3 Seed and seed development…………………………………………….……………2

1.4 Seed dormancy and germination………………………………………………………3

1.5 Preharvest sprouting………………………………………………………………………6

1.6 Classification of dormancy…………………………………………………………….7

1.6.1 Exogenous dormancy…………………………………………………………8

1.6.2 Endogenous dormancy……………………………………………………….8

1.7 Factors affecting seed dormancy…………………………………………………9

1.7.1 Physical factors affecting sprouting…………………………………………..9

1.7.2 Environmental effect……………………………………………………...10

1.8 Genetic and molecular determinants of seed germination and dormancy…………11

1.8.1 Arabidopsis mutants…………………………………………………………12

1.8.1.1 Abscisic acid……………………………………………………………………13

1.8.1.2 Gibberellin……………………………………………………………………...15

1.8.1.3 Ethylene……………………………………………………………………...15

1.8.1.4 Brassinosteroids………………………………………………………………...16

1.8.1.5 Transcription factors…………………………………………………………16

1.8.1.6 Auxins and cytokinins………………………………………………………18

1.8.1.7 Other seed dormancy mutants…………………………………………………19

1.8.2 Analysis of natural variation………………………………………………...25

1.8.2.1 Quantitative Trait Loci mapping………………………………………………26

1.8.2.2 Quantitative Trait Loci populations...……………………………………………28

1.8.2.3 Phenotyping……………………………………………………………………29

1.9 Seed dormancy Quantitative Trait Loci in model species…………………………31

1.9.1 Arabidopsis…………………………………………………………………..32

1.9.2 Rice..……………………………………………………………………34
1.9.3 Wheat.………………………………………………………………..36
1.9.4 Barley………………………………………………………………..41
1.10 Comparative genomics……………………………………………44
1.11 General thesis direction…………………………………………...46

Chapter 2: Candidate gene identification using comparative genomics…………..47
2.1 Introduction……………………………………………………………...47
2.2 Materials and methods……………………………………………………..50
2.2.1 General outline of bioinformatic strategy………………………………..50
2.2.2 Mapping of Arabidopsis seed dormancy QTLs………..…………………...52
2.2.3 Identification of Arabidopsis seed dormancy genes within QTLs…..…………..52
2.2.4 Identification of rice orthologs…………………………………………...53
2.2.5 Mapping of rice seed dormancy QTLs and identification of co-location of rice orthologs with QTL positions………………………………………………..54
2.2.6 Confirmation of Arabidopsis candidate genes and putative rice orthologs by gene structure similarity……………………………………………………...54
2.2.7 Bin location of wheat orthologs…………………………………………55
2.3 Results……………………………………………………………………...55
2.3.1 Mapping of Arabidopsis seed dormancy QTLs………..…………………...55
2.3.2 Identification of Arabidopsis seed dormancy genes and their location in relation to seed dormancy QTLs…………………………...………………………56
2.3.3 Identification of rice orthologs…………………………………………...58
2.3.4 Identification of rice orthologs within rice seed dormancy QTLs………..……...60
2.3.5 Comparison of Arabidopsis-rice candidate gene structure………………………61
2.3.6 Candidate gene alignment to binned wheat ESTs………………………………..65
2.4 Discussion………………………………………………………………66

Chapter 3: Seed dormancy in a Stirling/Harrington cross: generation of a molecular genetic map and QTL analyses……………..………………………...74
3.1 Introduction……………………………………………………………...74
3.2 Materials and methods……………………………………………………..76
3.2.1 Plant material……………………………………………………………76
3.2.2 DNA extraction…………………………………………………………..76

3.2.3 Molecular marker analysis……………………………………………………………76
3.2.4 Marker mapping and genetic map construction…………………………………78
3.2.5 Phenotyping trials……………………………………………………………………...79
3.2.6 Germination tests……………………………………………………………………80
3.2.7 Data analysis…………………………………………………………………………...81
3.2.8 QTL mapping…………………………………………………………………………82
3.3 Results………………………………………………………………………………...82
3.3.1 Trial climate…………………………...…………………………………………82
3.3.2 Seed germination……………………………………………………………………84
3.3.3 Statistical analysis…………………………………………………………………..85
3.3.4 Genetic map…………………………………………………………………………86
3.3.5 Detection of QTLs for seed dormancy…………………………..……………87
3.3.6 Seed dormancy QTL analysis on after-ripened seed……………………………..91
3.3.7 High seed dormancy genotype……………………………………………………...93
3.4 Discussion……………………...……………………………………………………..94

Chapter 4: Physical mapping of candidate genes *ERA1* and *ABI2* in wheat-barley addition lines and development of a CAPS marker for the *ERA1* Locus………..101
4.1 Introduction…………………………………………………………………………101
4.2 Materials and methods..……………………………………………….…………..103
4.2.1 Plant material………………………………………………………………………103
4.2.2 DNA extraction……………………………………………………………………104
4.2.3 Primer design………………………………………………………………………105
4.2.4 Mapping of candidate genes in wheat barley additional lines – assigning barley chromosome arm location…………………………………………………………107
4.2.5 Candidate gene sequencing and development of CAPS marker………………..108
4.2.5.1 PCR purification and sequencing reactions…………………………………..108
4.2.5.2 Sequencing reaction purification and sequence analysis……………………..108
4.2.5.3 Analysis and development of *ERA1* CAPs marker…………….…………..109
4.2.6 Genetic mapping of *ERA1* candidate gene in Stirling/Harrington population.…110
4.3 Results………………………………………………………………………………110
4.3.1 Alignment of wheat and barley ESTs to candidate rice genes for seed dormancy……………………………………………………………………...……110
4.3.2 Candidate gene sequencing and comparison of gene structure…………………111

4.3.2.1 Comparison of *ERA1* gene structure and sequence between rice and barley……………………………………………………………………………..111
4.3.2.2 Comparison of *ABI2* gene structure and sequence between rice and barley……………………………………………………………………………..112
4.3.3 Chromosome arm location of candidate genes………………………………115
4.3.4 Development of SNP-CAPs marker for the *ERA1* gene……………………..117
4.3.5 Genetic mapping of *ERA1* gene using the SNP-CAPS marker……………....118
4.4 Discussion…………………………………………………………………………119

Chapter 5: General discussion…………………………………………………...125

References……………………………………………………………………………137
Appendix A: Arabidopsis candidate gene list...……………………………………169
Appendix B: Splicing junctions of candidate genes………………………………174
Appendix C: Pilot map of Stirling/Harrington DH population using SSR and DArT markers…………………………………………………………………………176

LIST OF FIGURES

1.1 Schematic diagram of barley grain showing directions of water movement during imbibition using iodine-potassium iodide solution and sodium thiosulfate (kindly provided by Allen Tarr, Department of Agriculture and Food Western Australia), page 4.

1.2 Diagram of dormancy stages a seed may experience from maturity to germination depending on favourable or unfavourable environmental conditions. Genotype and environmental factors during seed development will also have an effect on dormancy levels of a seed throughout its development, page 6.

2.1 Schematic diagram of bioinformatic strategy used to identify candidate genes for seed dormancy (SD) using comparative genomics between Arabidopsis and rice, page 51.

2.2 Arabidopsis genetic (left) and physical (right) maps showing published seed dormancy genes and QTLs and common reference markers between the maps (indicated by grey lines between the maps). Published seed dormancy QTLs are defined by green vertical lines (DOG QTLs from Alonso-Blanco et al, 2003) and blue vertical lines (van der Schaar 1997). Seed dormancy genes are highlighted in red, khaki and brown. Genes in red are confirmed to locate within a seed dormancy QTL on the Arabidopsis physical map, khaki genes are those identified to locate within a seed dormancy QTL based on distances between common markers of the two maps. These QTLs are also marked with an asterix. Genes highlighted in brown are additional genes incorporated as candidate genes due to their close

proximity, (within 500,000bp) to a seed dormancy QTL defined by only one marker, page 57.

2.3 Arabidopsis physical map with published seed dormancy genes and QTLs with correspondences to rice physical map (Gramene Tigr Assembly map, pseudomolecule version 3) generated in CMAP (http://ccg.murdoch.edu.au). Rice physical map shows location of published rice seed dormancy QTLs. Figure shows the location of the genes *ERA1*, *ABI2*, *HYL1* and *ETR1* within an Arabidopsis QTL as well as the location of their rice orthologs. These rice orthologs also locate to a rice seed dormancy QTL, page 62.

2.4 Intron-exon structure of Arabidopsis candidate genes (red) and putative rice orthologs (blue). Translated regions are represented by boxes, and introns by horizontal lines with the lengths of exons and introns displayed as number of base pairs. Grey shading indicates regions between the orthologs that are significantly aligned. GenBank accession numbers are given for the cDNA and BAC/PAC clones, pages 63-64.

3.1 View of marker screening performed in GeneMapper v3.7 demonstrating clear polymorphism in marker fragment size between DH lines of the Stirling/Harrington population, page 78.

3.2 Daily rainfall and average daily temperature for the Wongan Hills 2006 site. Sowing date (S) and harvest date (H) are shown, page 83.

3.3 Distribution of seed germination for germination tests carried out immediately after harvest in the Stirling/Harrington DH population from the two different sampling years and sites, page 85.

3.4 Genetic map of Stirling/Harrington DH population showing marker order and distances between them. Putative centromeric regions (based on published consensus maps, see text) are indicated by a black triangle. Chromosome regions that harbour QTL are displayed by coloured vertical lines. QTL defined by a blue line represent seed dormancy QTL identified in the first germination test. QTL marked by an asterix were only detected in epistatic interactions. QTL defined by a green line represent seed dormancy QTL detected using after-ripened seed 7 weeks post harvest. Main QTLs identified for seed weight are also shown (defined in red). The inset summarises QTL interactions and the interaction type, page 89.

3.5 QTL-likelihood curves (in red) generated by QTLNetwork, showing the locations of QTLs on 1H, 3H and two QTLs on 5H for the first germination tests carried out immediately after harvest, taking into account all three environments, Wongan Hills 2005, Wongan Hills 2006, Mt Barker 2006. The horizontal blue dash line indicates the QTL threshold equivalent to a critical F value of 6 (P<0.01), page 90.

4.1 Diagram showing the exon (boxes) and intron (horizontal lines between boxes) structure of the candidate genes *ERA1* and *ABI2*. Gene structure is based on the rice *ERA1* and *ABI2* gene homologs. Diagram shows the primer locations used to amplify regions of the candidate genes and are represented by arrows. The number of the primers refers to the numbers used in Table 3.1, describing the primer

sequence and name, page 106.

4.2 Intron-exon structure of candidate genes *ERA1* and *ABI2* in rice (blue) and sequenced barley (green). Exons are represented by boxes, and introns by horizontal lines with the lengths of exons and introns displayed as number of base pairs. Base pairs displayed above the exons for the barley *ERA1* and *ABI2* genes represent the number of base pairs sequenced. Dotted exons and introns indicate regions without sequence data. The size of non-sequenced exon regions are based on aligned wheat and barley ESTs. Grey shading indicates regions between the orthologs that are significantly aligned. Diagram shows location and sequence of SNPs detected between Stirling and Harrington parental lines for the *ERA1* gene. The SnaBI restriction site for the *ERA1* CAPS marker is outlined, pages 113-114.

4.3 Physical mapping of candidate genes *ERA1* (A, B) and *ABI2* (C). Gel pictures of 1% agarose gels (1 x TBE) demonstrating the fragment sizes amplified in the wheat barley addition lines. Primer pairs used to amplify the fragments are indicated. The allocation of the *ERA1* gene and the *ABI2* gene to the long arm of 3H is based on the presence of a band in the Betzes (B) cultivar and the ditelosomic line for 3HL that is absent in Chinese Spring (CS) and the other wheat barley addition lines, page 116.

4.4 Genetic of mapping of the *ERA1* gene in the Stirling/Harrington double haploid population using the developed SNP CAPS marker. Mapping of the *ERA1* gene positions the locus at the top of the long arm of 3H. Location of seed dormancy QTLs identified in Chapter 3 is shown (SD QTL). CAPS loci were produced by

amplification of genomic DNA using primers ERAi1F and ERAi1R followed by digestion of the PCR product by the restriction enzyme SnaBI and analysed on 1% agarose gel (shown in inset). Marker bands for Stirling (428 bp) and Harrington (521 bp) cultivars are indicated. Location of the centromere is approximate, page 118.

5.1 Hypothetical graph demonstrating germination behaviour of a fast germination phenotype with no dormancy such as the Harrington cultivar, a slow release phenotype with strong levels of seed dormancy such as Stirling and what may be the ideal seed dormancy phenotype displaying strong dormancy with a sudden timely release of dormancy such as Baudin, page 127.

5.2 Aspects of seed germination pathway in barley. Examples are given, of environmental conditions conducive to germination or dormancy and highlighting seed dormancy genes (most of which are described in the literature review) functionally characterised in Arabidopsis that either promote or repress germination. Candidate genes *ERA1* and *ABI2* are in bold. Environmental conditions would be influential throughout the process from seed development to germination. Certain phases would be more sensitive to the environment, particularly after the loss of primary dormancy. The figure makes suggestions for where QTLs may come into effect, page 131.

[1]*GA 20-oxidase* in cereals is orthologous to *GA5* in Arabidopsis.

[2]See Figure 5.1 for a description of what may be an ideal seed dormancy phenotype in barley.

[3]*Vp1* in cereals is orthologous to *ABA3* in Arabidopsis.

LIST OF TABLES

1.1 Table of selected Arabidopsis mutants related to seed germination/dormancy identifying genes with various roles in hormone biosynthesis, signalling, response and action page 20-24.

1.2 Summary table of Quantitative Trait Loci (QTL) studies in rice, wheat and barley, page 33.

2.1 Summary of the BLASTP analysis in Arabidopsis (At) and rice (Os) identifying putative rice orthologs of the Arabidopsis candidate genes for seed dormancy. Table shows the rice BAC/PACs for the rice protein accessions and the e values for the Arabidopsis-rice protein sequence alignments. Percent identity (pid) and high scoring pair length (HSP length) are calculated from combined high scoring pairs of a single hit. Rice accessions that align to the same BAC all align to the same position on the BAC. BLAST results for the controls are presented at the top of the table (highlighted with an asterix) – *ABA3* and *GAI* Arabidopsis and rice orthologs are well established in the literature and serve to ensure the BLAST stringencies were discriminating, page 59.

2.2 Results of BLASTN and TBLASTX alignments between rice candidate gene cDNA or BAC clones and bin mapped wheat ESTs. Table shows the chromosome location of the wheat EST hits and the wheat library that they belong to. Details of BLAST search are given, including e value, percent identity (pid) and high scoring pair length (HSP). The pid and hsp length are calculated from combined high scoring pairs of a single hit, page 66.

2.3 Wheat deletion bins containing wheat ESTs that align to rice orthologs of candidate genes *ABI2* and *ERA1*. Table also shows wheat microsatellite markers that have been deletion bin mapped (Sourdille et al. 2004) and are associated with seed dormancy QTLs in wheat (Kulwal et al. 2004, Groos et al. 2002), page 72.

3.1 Mean, variance, standard deviation and coefficient of variation given for seed weight and seed germination under the three environments tested, page 86.

3.2 Number of mapped markers, total genetic distance and maximum distance (genetic) between marker intervals for each chromosome of the genetic map constructed from the DH population derived from the Stirling/Harrington cross (see Figure 3.3 for map diagram), page 87.

3.3 Description of significant (P<0.05) seed dormancy QTLs detected in the Stirling/Harrington DH population in three environments (two locations and two seasons). QTL heritability, additive main effects and epistatic and environmental interactions are shown, page 92.

3.4 Description of seed dormancy QTLs detected in the individual three environments tested using after-ripened seed (seed stored for 7 weeks after harvest). Chromosome assignment, marker interval, additive effect, heritability, standard error and P value are given for each QTL. Parental allele for dormancy or germination is given for each locus, page 92.

3.5 Table of means and standard variation for genotypic combinations at minor QTL locations on chromosomes 1H and 3H for germination percentage in the three environments tested. R^2 values are shown for P<0.0001 (***). Allele 'A' is the Stirling allele and allele 'B' is the Harrington allele. The first allele of the genotypes is representative of the allele at the 1H QTL and the second allele is representative of the allele at the 3H QTL e.g. AB = Stirling allele at the 1H QTL and Harrington allele at the 3H QTL. MB = Mount Barker, WH = Wongan Hills, page 93.

4.1 Primers used to successfully amplify *ERA1* and *ABI2* candidate gene regions. The primer number corresponds to the numbers indicated in Figure 4.1. Accession numbers of wheat and barley ESTs from which the primers were designed from are provided, page 106.

4.2 Summary of BLAST alignment of wheat and barley ESTs to exon regions of rice candidate genes *ERA1* and *ABI2*. Table demonstrates BLAST e values, percentage coverage of the rice cDNA query, maximum identity of blast hit to wheat and barley EST accessions and the region of alignment on the *ERA1* and *ABI2* genes, page 111.

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

1.1 General Introduction

Cultivated barley (*Hordeum vulgare* L.) is the world's fourth most important cereal crop (Li & Foley 1997) and one of the first crop species to be domesticated (Badr et al. 2000). The crop is now distributed virtually worldwide from temperate areas to even the subtropics (von Bothmer et al. 2003). Barley was initially cultivated primarily for human consumption and animal feed. Although its uses for human food are limited, barley is still used for bread, porridge and other food products today in regions of Asia and northern Africa (von Bothmer et al. 2003). In temperate regions, it is still a major crop used as animal feed for cattle, pigs and poultry. The use of barley in malting for beer brewing and in fermentation and distillation to produce whisky was developed early in history and is currently is one of the crop's most significant high-value end use (Fox et al. 2003). The effort by plant breeders and geneticists to improve the yield, nutritional quality and commercial value of the crop in order to satisfy the diverse end uses is global and ongoing.

Preharvest sprouting is a grain condition that reduces the starch quality of barley grains (Ringlund 1980). Grains displaying more than 5% germination are rendered unsuitable for malting (Brewing and Malting Barley Research Institute 2005) causing substantial economic losses for farmers when they are forced to sell their grain in the less profitable animal feed market. Seed dormancy is the most significant factor affecting grain levels of preharvest sprouting. Preharvest sprouting susceptibility is a consequence of low grain dormancy prior to harvest. This thesis contributes an understanding of seed

dormancy by examining the chromosomal regions that contain genes controlling seed dormancy and investigating possible candidate genes contributing to this trait.

1.2 The barley plant

The Grass family of Poaceae is a large family of monocots comprising many economically important cereal crops such as rice (*Oryza sativa* L.), wheat (*Triticum aestivum* L.), maize (*Zea mays*) and barley. Within this family, barley belongs to the tribe Triticeae, a highly evolved diverse group containing members demonstrating various degrees of polyploidy and interspecific and intergeneric hybridisations (von Bothmer et al. 2003). Apart from being an important agricultural crop, barley is additionally popular for biological research due to its low chromosome number and large chromosomes. The barley genus *Hordeum* contains approximately 32 species most of which are diploids with seven pairs of chromosomes (2n=2x=14) (von Bothmer et al. 2003) made up of approximately 5 x 10^9 bp of DNA (Bennett & Smith, 1976). Several tetraploids (2n=4x=28) and (2n=6x=42) also exist in this genus (von Bothmer et al. 2003).

1.3 Seed and seed development

The structure of cereal grains is well characterised. Figure 1.1 shows a schematic diagram of a barley grain and movement of water into the grain during imbibition. The greater part of a cereal grain consists of the starchy endosperm, which functions as the seed food reserve containing both carbohydrate and protein resources (King 1989). The other tissues of the grain are made up of the scutellum, aleurone and embryo. The endosperm is surrounded by the aleurone tissue layer, typically 2-3 cells thick, and is separated from the embryo by the scutellum (King 1989, Duffus & Cochrane 1992).

The embryo is made up of the coleoptile, three to four embryonic leaves and rootlets. Enveloping both the endosperm and the embryo are the maternal tissues of the testa (inner seed coat) and the pericarp (outmost tissue) (King 1989, Duffus & Cochrane 1992). Seed development consists of two major phases, namely embryo development and seed maturation. Embryogenesis commences with the formation of a single cell zygote and is completed with the formation of all the embryo structures (Mayer et al. 1991). A growth stage follows in which the embryo enlarges and fills the seed sac (Goldberg et al. 1994). Once the embryo reaches its full size, cell division in the embryo is arrested and the seed is said to have reached physiological maturity (Duffus & Cochrane 1992). However, to reach full maturity, the seed must undergo the dehydration phase of the maturation stage. The seed maturation stage involves the accumulation of food reserves and the development of dormancy and defence against desiccation (Goldberg et al. 1994). In the course of normal seed development and under suitable environmental conditions, the dry seed will imbibe water and germinate during which embryo arrest and dormancy processes are reversed (Bentsink & Koornneef 2002). The rate at which these events take place depends on both the genotype and the environmental conditions during seed development (Duffus & Cochrane 1992).

1.4 Seed dormancy and germination

Germination and dormancy processes overlap and both are determined by the balance between the degree of dormancy and the capacity of the embryo to overcome dormancy (Bentsink & Koornneef 2002). Germination can be defined by the uptake of water and the subsequent elongation of the embryonic axis (Bewley & Black 1994). The initial uptake of water by the mature seed comprises a rapid preliminary uptake, followed by a

plateau phase. Post germination, when the embryonic axis has elongated, there is a further increase in water uptake (Bewley 1997). Morphologically, germination is complete when the radicle penetrates the surrounding embryo structures (Bentsink & Koornneef 2002).

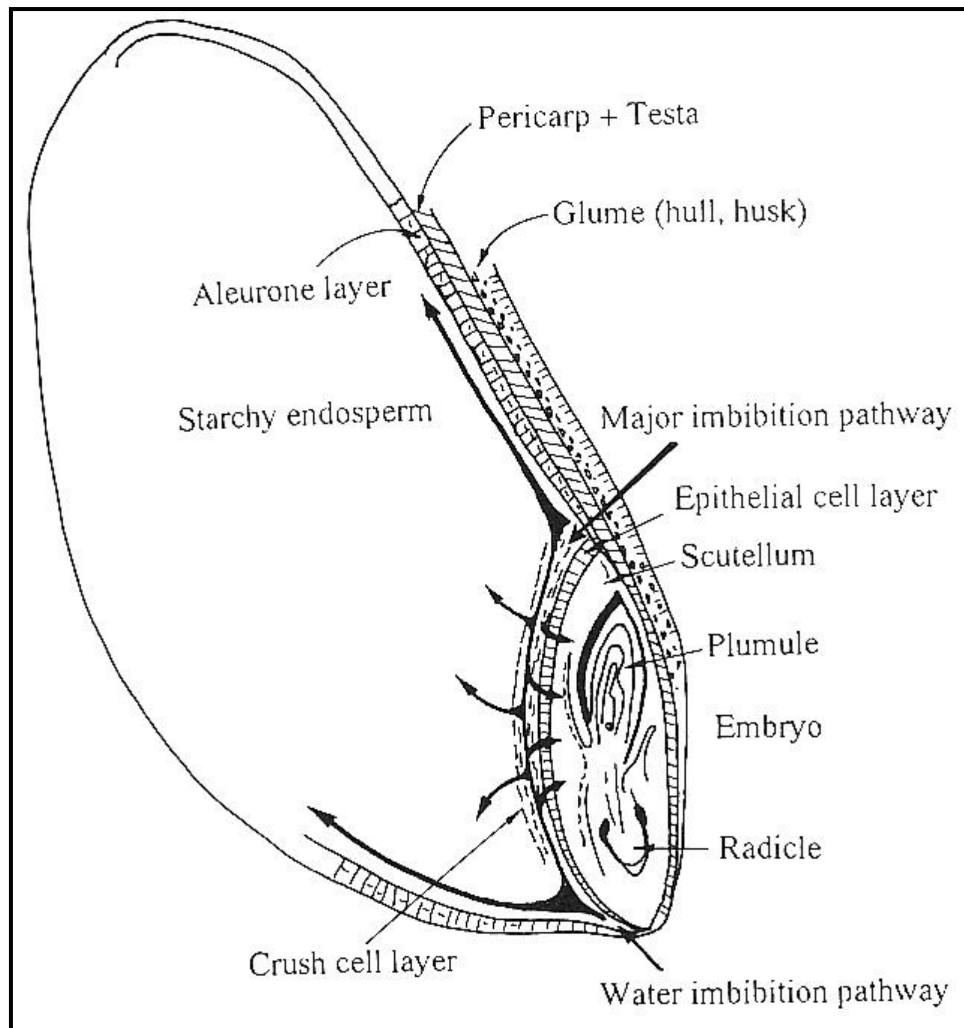


Figure 1.1: Schematic diagram of barley grain showing directions of water movement during imbibition using iodine-potassium iodide solution and sodium thiosulfate (kindly provided by Allen Tarr, Department of Agriculture and Food Western Australia).

Dormancy is widely adapted trait that prevents a seed from germinating in climatic and environmental conditions that are not conducive to seedling survival. The terms quiescent, after-ripening, primary and secondary dormancy are terms used by researchers to describe and categorise the seed state with reference to its germinability and/or dormancy. As viable seeds mature they may be dormant or non-dormant. Generally non-dormant seeds are those that do not require specific environmental cues to germinate (Finch-Savage & Leubner-Metzger 2006). Quiescent seeds are non-dormant seeds that will germinate when given specific environmental conditions of temperature, water and oxygen (Baskin & Baskin 1998). Primary and secondary dormancy refer to the timing of the occurrence of dormancy. Primary dormancy is induced when the seed is attached to the plant and is characterized by the failure of an intact viable seed to germinate under favourable conditions that would normally induce germination and development (Li & Foley 1997, Foley 2001). The transition of the seed from the primary dormant state to a more readily germinable state is termed after-ripening. The duration of after-ripening depends on the degree of dormancy the seed has and is affected by specific environmental conditions (Li & Foley 1997, Foley 2001). Secondary dormancy can develop from primary dormancy in dispersed, mature seeds under certain unfavourable environmental conditions that induce a quiescent state (Bewley & Black 1994, Baskin & Baskin 1998). Secondary dormancy can be lost and induced repeatedly in parallel to seasonal changes in temperature to avoid germination in small windows of favourable environmental conditions that occur within unfavourable seasons (Hilhorst 2007). Seeds with secondary dormancy will generally respond to the same after-ripening conditions and treatments that induce germination in primary dormancy (Karssen 1982). In summary, dormancy is regulated at different stages of seed development and maturity and is continuously changing with time and

the environmental conditions that the seed is exposed to. Figure 1.2 shows the stages of seed dormancy with favourable and unfavourable conditions.

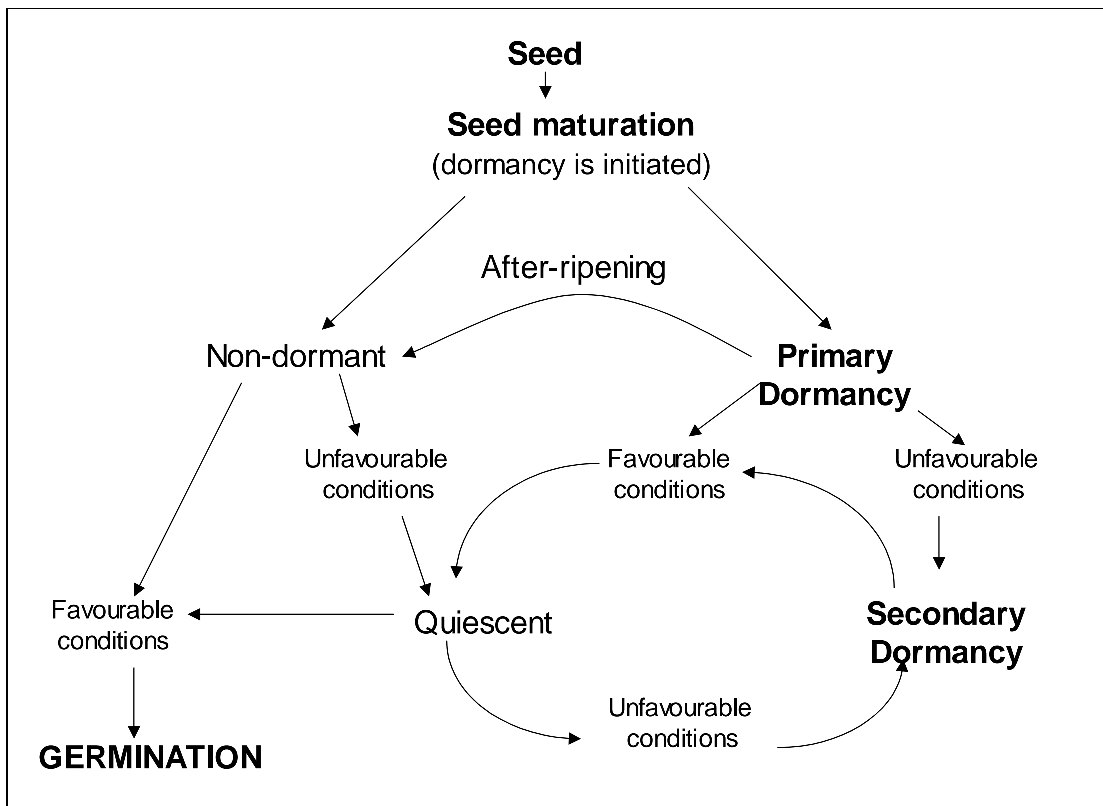


Figure 1.2 Diagram of dormancy stages a seed may experience from maturity to germination depending on favourable or unfavourable environmental conditions. Genotype and environmental factors during seed development will also have an effect on dormancy levels of a seed throughout its development.

1.5 Preharvest sprouting

Seed dormancy is an important trait for many cereal species because it affects resistance to preharvest sprouting. The selection and breeding for rapid germination and growth during the domestication of economically important cereal crops has lead to the removal of most seed dormancy mechanisms (Adkins et al. 2002, Finkelstein 2008). In some cases this has resulted in plants with seed that are fully germinable prior to harvest. In the final maturation stages, under wet or even damp and humid conditions, these grains will germinate while still on the mother plant prior to harvesting. This phenomenon is known as pre-harvest sprouting and the ensuing consequences are detrimental, causing substantial losses in crop yield and grain quality. Hydrolysis of

starch in the endosperm is induced and this leads to a decrease in grain weight in addition to promoting saprophytic fungal growth (Castor & Frederiken 1977). The grain storage life of preharvest sprouted grain is also lowered (Bason et al. 1991). Preharvest sprouting severely reduces the quality of malting barley in particular (Hoyle 1999). Germination results in the production of endoproteolytic enzymes that hydrolyse storage proteins into soluble proteins, peptides and amino acids, the ratio of these soluble products being an important factor in determining brewing quality (Hoyle 1999). As the degree of embryo dormancy at grain harvest essentially determines the genetic variation found in preharvest sprouting tolerance, the terms preharvest sprouting tolerance and grain dormancy are often used interchangeably (Kulwal et al. 2004).

1.6 Classification of dormancy

Dormancy can be loosely classified into two types of dormancy, exogenous dormancy and endogenous dormancy, based on the morphological and physiological properties of the seed respectively (Nikolaeva 1977, 2004). Baskin & Baskin (2004) devised a scheme that further classified seed dormancy into five dormancy classes including physical (PY), physiological (PD), morphological (MD), morphophysiological (MPD) and combined dormancy (PY + PD). While this classification scheme is useful for phylogenetic, biogeographic and evolutionary studies, it is difficult to categorise the underlying mechanisms that induce these subjective classes of dormancy in genetic and molecular studies (Hilhorst 2007). Therefore, exogenous and endogenous dormancy are described here for assigning dormancy.

1.6.1 Exogenous dormancy

In exogenous dormancy, the structures surrounding the embryo (i.e. maternal tissues) prevent seed germination by acting as permeability barriers that limit imbibition and gaseous exchange, and mechanical barriers that restrict embryo enlargement. In exogenous dormancy, mechanical or physical scarification of the seed coat or its removal, will result in the promotion of a germinating embryo unless the embryo itself is dormant (endogenous dormancy). Seed coat imposed dormancy is maternally inherited and used as a survival strategy in many diverse species, protecting the seed against detrimental environmental conditions (Debeaujon et al. 2000). While exogenous seed dormancy is important in many plant species, the focus of this thesis is primarily on endogenous dormancy. Seed coat function in seed dormancy has been reviewed in Ballard (1973), Rolston (1978), Finch-Savage & Leubner-Metzger (2006) and Debeaujon et al. (2007).

1.6.2 Endogenous dormancy

Endogenous dormancy involves some characteristic of the embryo inhibiting germination (Baskin & Baskin 1998). The embryo will only germinate if exposed to specific abiotic or biotic conditions. Seeds with endogenous dormancy may require environmental (temperature, light, smoke, stratification) or physiological (plant growth regulators such as gibberellic acid and ethylene, nitrogen compounds) treatments to trigger germination. Cold and moisture treatments can also release embryo dormancy in cool season species by mobilising carbohydrate and lipid reserves and allowing the embryo to access food reserves necessary for germination to occur (Adkins et al. 2002). In addition, the growth potential of the embryo must be great enough to overcome the

restrictions imposed by the surrounding tissues for germination to occur (Bentsink & Koornneef 2002).

1.7 Factors affecting seed dormancy

Dormancy is regulated at different stages of seed development and maturity and is continuously changing with time and the environmental conditions that the seed is exposed to. While seed dormancy is largely a function of genotype, environmental and physical factors play an important role in influencing the expression of the trait.

1.7.1 Physical factors affecting sprouting

Morphological characteristics of cereals such as the size, shape and structure of the ear and glume, ear nodding angle and the presence of awns, affect exposure of the grains to water and thus influences their capacity to absorb and retain moisture (King & Richards 1984, King & Wettstein-Knowles 2000). The exposure of the grain to moisture would then in turn affect preharvest sprouting of the grain. Early studies also related seed dormancy in barley to grain positions on the ear. Grains in positions closer to the centre of the ear were thought to attain a favourable degree of assimilation (Scott et al. 1983). Although differences in grain dormancy in different ear positions have not been significant, several studies have noted smaller grains with low viability and greater dormancy at the extremities of the ear compared to those situated toward the centre of the ear (Chaussat et al. 1983, Carreck & Christian 1997). These minor factors may contribute to variation in susceptibility to preharvest sprouting, however, these are not elaborated on in this review, as their roles relative to genetic and environmental effects are small in barley.

1.7.2 Environmental effect

The large effect of the environment on seed dormancy is evident from the variation in dormancy observed between seeds of the same genotype. Dormancy is a common trait throughout the higher plants and has evolved differently across species existing in various climatic regions (Finch-Savage & Leubner-Metzger 2006). Seeds generally need water, oxygen and an appropriate temperature to germinate, however, plants have adapted divergent responses to the range of environmental conditions they are subjected to, in order to time germination when climatic conditions are most favourable for plant establishment and growth (Finch-Savage & Leubner-Metzger 2006). Generally, plants use dormancy to time their emergence by maintaining dormancy prior to the period with environmental conditions that are adverse for seedling survival. This dormancy is lost in the period preceding the season in which the conditions are conducive for seedling establishment (Allen et al. 2007). Thus, the specific conditions that a seed needs to overcome dormancy, varies with species, its habitat and the mechanism of dormancy it utilizes (Foley 2001, Finch-Savage & Leubner-Metzger 2006).

Environmental factors such as light, temperature, moisture and dissolved gases, can act as cues and initiate a cascade of plant responses that will either encourage germination or enhance dormancy. For practical reasons, factors influencing germination and dormancy are characterised on the basis of whether they inhibit or promote germination. However the continuum between dormancy and germination makes it difficult to distinguish which of these two processes certain factors are affecting (Cohn 1996). For example, when a seed germinates in the presence of light, it is unclear whether the light is causing the seed to release dormancy (e.g. Benech-Arnold et al. 2000, Batlla et al. 2005) or if it is promoting germination (eg. Vleeshouwers et al. 1995). This has lead to

confusion in categorising factors based on whether they control the induction or release of dormancy depending on how one separates the two processes (Finch-Savage & Leubner-Metzger 2006). Until germination and dormancy pathways are completely understood, establishing the exact effect certain conditions have on dormancy will remain indistinct.

Adding to the complexities in defining environmental conditions that affect dormancy and germination is the fact that there are windows in which the seed is more susceptible to specific environmental conditions depending its maturity and development stage. For example, temperature and rainfall during grain filling, has been shown to have a large effect on the expression of grain dormancy in wheat (Nielsen et al. 1984, Mares 1993b, Nyachiro et al. 2002, Biddulph et al. 2007). Rainfall can affect dormancy through grain wetting or through the effect of water supply to the plant itself (Biddulph et al. 2007). The effect that certain environmental conditions have in releasing dormancy and inducing germination can change throughout the seed's development.

The specific environmental conditions that influence seed dormancy are also dependent on the genetic background of the plant (Finch-Savage & Leubner-Metzger 2006). The genotype by environment interactions that occur make dormancy an even more complex trait to dissect.

1.8 Genetic and molecular determinants of seed germination and dormancy

Seed dormancy and germination is controlled and regulated by a complex network of cellular and metabolic pathways that involve a wide range of hormones. The pivotal role of abscisic acid and gibberellin in inducing and releasing dormancy, respectively, is

well known and has been covered in several publications (Koornneef & Karssen 1994, Bewley 1997, Finkelstein 2004, Finch-Savage & Leubner-Metzger 2006, Feurtado & Kermode 2007, Finkelstein et al. 2008). The hormones brassinosteroid, ethylene and auxin have also been associated with seed dormancy and germination, as well as sugar, light sensing factors and seed coat metabolites. Greater understanding of the genetic aspects of seed dormancy has been gained through genetic approaches using mutantagenesis and analysis of variation. Emphasis is placed on Arabidopsis in this review as mutant approaches in the last decade have primarily been conducted using this model species.

1.8.1 Arabidopsis mutants

Arabidopsis thaliana, referred to as Arabidopsis hereafter, has contributed greatly as a model plant for functional studies in genetic and molecular research. The extensive functional analyses performed using gene disruptions to determine gene function have provided a valuable resource to study higher plants (Koornneef et al. 2004). Gene disruption, through the use of T-DNA, transposon insertions or RNAi technology to inhibit or overexpress gene-expression, have enabled gene function investigations by observing the phenotypes of the genotypes in which the genes have been disrupted (Koornneef et al. 2004). These studies in Arabidopsis have identified an ever increasing number of genes involved in the control of seed dormancy and germination down to a molecular level (Bentsink & Koornneef 2002, Bentsink et al. 2007). Not surprisingly, as these hormones have multiple functions in plant development and growth processes other than seed dormancy and germination, many of the Arabidopsis hormone mutants demonstrate pleiotropic phenotypes.

Arabidopsis mutants involved in the regulation of seed dormancy and germination have been reviewed in Bensink & Koornneef (2002) and Koornneef et al. (2002). Arabidopsis provides a model for plants in general, however, it is recognised that certain mechanisms controlling seed dormancy in Arabidopsis may be specific to the plant, especially when comparing the dicotyledon plant to monocot crop species. For example, Arabidopsis requires after-ripening, light and responds to stratification (Foley 2001, Baumbusch et al. 2004) for germination, however, cereal grain such as white wheat and barley germinate in darkness and seed coat dormancy appears to play a minimal role. There are various reviews on Arabidopsis genes involved in the regulation of germination and seed dormancy (McCarty 1995, Foley et al. 2001, Bentsink & Koornneef 2002, Koornneeff et al. 2002, Kucera et al. 2005, Bentsink et al. 2007, Finkelstein et al. 2008) that include descriptions of mutants affected in seed coat characteristics or light mediated germination. This review focuses on the Arabidopsis hormone mutants that may represent fundamental genes that are involved in endogenous seed dormancy control in broad plant groups including cereals. Table 1.1 describes selected Arabidopsis hormone mutants, most of which are briefly summarised below.

1.8.1.1 Abscisic acid

The Arabidopsis genotypes Landsberg erecta (*Ler*) and Columbia (*Col*) are frequently used in Arabidopsis research and both accessions have low dormancy levels. This can be a problem in screening for variability in the trait or false positives may be detected due to misclassification of individual seeds in the mutant screens (Bentsink et al. 2007). Some of these problems have been overcome by using the more dormant accession Cape Verde Islands (Cvi) (Koornneef et al. 2004).

Abscisic acid (ABA) and gibberellic acid (GA) play crucial roles in inducing dormancy and enhancing germination, respectively. The level of dormancy is based on the ABA:GA ratio resulting from ABA biosynthesis and GA degradation (Finch-Savage & Leubner-Metzger 2006). Therefore genes controlling cellular perception, sensitivity, response, signalling and regulation of these phytohormones will be fundamentally important for germination and dormancy (Kucera et al. 2005).

Mutants that are altered in ABA biosynthesis or its mode of action have clearly shown that the phytohormone is involved in the onset and maintenance of dormancy. ABA biosynthesis mutants (ABA mutants) demonstrate reduced seed dormancy and germinate in the presence of GA biosynthesis inhibitors like paclobutrazol and tetcyclacis (Leon-Kloosterziel et al. 1996a). Crosses between these mutants and the wild type indicated that dormancy was controlled by the genotype of the embryo rather than the maternal genotype. ABA produced by the maternally inherited seed coat or exogenously applied ABA during seed development was not sufficient to produce long lasting dormancy (Karssen et al. 1983, Koornneef & Karssen 1994, Nambara & Marion-Poll 2003). Major effects in seed dormancy are also observed in mutants that are defective in ABA signalling. Mutants in ABA signalling have been identified by selecting seeds that are able to germinate despite the presence of ABA concentrations that are normally inhibitory for the wild type (Koornneef & Karssen 1994, Leung & Giraudat 1998, Brocard-Gifford et al. 2004). Their response to ABA has termed them ABA insensitive mutants. These mutants can have a marked reduction in seed dormancy and display increased levels of endogenous ABA during seed development. *ABH1* and *ERA1* mutants are also involved in ABA signalling, however, they are ABA hypersensitive, demonstrating a response to low levels of ABA that are not normally inhibitory (Cutler et al. 1996, Hugouvieux et al. 2001).

1.8.1.2 Gibberellin

The pivotal role that gibberellins (GA) play in inducing seed germination has been shown by the germination of dormant seeds that require after-ripening, light and/or stratification after application of exogenous GA (Karssen et al. 1989). Also, the addition of GA inhibitors to seeds prevents their germination (Karssen et al. 1989). The majority of GA-deficient mutants require GA treatment following the impairment of GA biosynthesis, in order to germinate (Koornneef & van der Veen 1980, Sun et al. 1994, Yamaguchi et al. 1998a, 1998b, Helliwell et al. 1999, Hisamatsu et al. 2005), indicating the primary role GA plays in germination.

Transcription factors encoding the DELLA motif make up a prominent family involved in GA signalling that includes *GAI* (*GA insensitive*), *RGA* (*repressor of ga1-3*), *RGL1* (*RGA-LIKE1*), *RGL2* and *RGL3*. These genes regulate response to GA as well as responses to ethylene and auxin stimulus (Achard et al. 2006). Another locus *SPINDLY* (*SPY*) is a negative regulatory of GA signal transduction and is believed to act epistatically downstream from *GAI*. The *GCR1*(*G protein-coupled receptor* gene) whose overexpression results in reduced seed dormancy is also suggested to play a role in GA signalling (Colucci et al. 2002).

1.8.1.3 Ethylene

Ethylene has a role in growth and developmental processes varying from germination to scenescence. Ethylene generally enhances germination efficiency and ethylene insensitive mutants for the genes *ETR* and *EIN2* demonstrate greater dormancy (Bleeker et al. 1988). Recent studies indicate that ethylene and abscisic acid signalling and response pathways may also interact to influence seed dormancy and germination

(Beaudoin et al. 2000, Ghassemian et al. 2000). Hypersensitivity to abscisic acid has been observed in ethylene response mutants and ethylene mutants also have a similar phenotype to mutants defective in ABA signalling, suggesting that ethylene may act as an inhibitor of ABA action during germination (Beaudoin et al. 2000, Ghassemian et al. 2000). Ghassemian et al. (2000) further suggested that *EIN2* may influence *ABA* biosynthesis by suppressing the ABA-insensitive *ABI1* gene. Cross-talk between ethylene and sugar signalling has also been indicated by the sugar insensitive phenotype of *CTR1* (Gibson et al. 2001) and hypersensitivity of *ETR1* to sugar (Zhou et al. 1998).

1.8.1.4 Brassinosteroids

Brassinosteroids (BRs) are plant steroids commonly found in a wide variety of plant species (Clouse & Sasse 1998, Schumacher & Chory 2000). The plant steroid is known to be involved in controlling cell elongation, cell division and skotomorphogenesis and has been additionally associated with the promotion of germination in Arabidopsis. The Arabidopsis mutants *de-etiolated2* (*DET2*) and *BR insensitive1* (*BR1*) demonstrated reduced germination. Steber & McCourt (2001) suggested ABA-induced dormancy was overcome by BR signalling. Indeed, the BR insensitive mutants *BIN2*, *BRI1* and BR deficient mutants exhibit an ABA hypersensitive phenotype (Li et al. 2001). However, there is no evidence that BRs are essential for germination to occur as no germination defects have been observed in mutants of BR biosynthesis or response (Steber & McCourt 2001).

1.8.1.5 Transcription factors

Arabidopsis mutant studies have highlighted several transcription factors with a role in seed dormancy. These include *FUS3, LEC1, LEC2* and *ABI3*, genes with active roles in

the later stages of seed maturation that affect seed germination (Meinke et al. 1994, Nambara et al. 2000). Mutation of these genes affects seed developmental processes such as storage protein accumulation, desiccation tolerance and seed dormancy. *LEC1*, *LEC2* and *FUS3* have similar phenotypes and are likely to control developmental arrest as immature embryos of Arabidopsis continue to grow following mutation of these genes (Bentsink & Koornneeff 2002). These Arabidopsis mutants are additionally sensitive to ABA. While the functions of these four genes is said to be partially overlapping (Parcy et al. 1997), they induce premature germination at different times and have different requiremnents for germination (Bentsink & Koornneeff 2002). *FUS3* mutants require GA for germination to occur while *LEC1* and *LEC2* do not, although germination is poor. The fatty acid transport *CTS* mutant is thought to be involved in some aspect of seed dormancy based on its 'forever' dormant phenotype and genetic interactions with the seed maturation genes *ABI3, LEC1*, and *FUS3* which rescues the germination phenotype (Russell et al. 2000, Footitt et al. 2002).

The *DAG1* and *DAG2* genes, which are maternally inherited, encode DOF transcription factors which when knocked out have opposing effects in seed dormancy although they are related genes. *DAG1* represses germination while the *DAG2* promotes germination (Papi et al. 2000, Gualberti et al. 2002). As the genes are maternally inherited, Papi et al (2000) suggested they may play a role in transporting compounds from the mother plant into the seed. *GCR1* is also a knockout mutant that demonstrates altered seed germination phenotypes and reduced sensitivity to germination promoting hormones GA and BR.

Mutants of transcription factors *IMB1* and *PKL* also have seed dormancy phenotypes. *PKL* is involved in the transition phase between seed development and the induction of dormancy or germination. The gene interacts with the seed maturation mutants *LEC1* and *LEC2* by repressing their expression during germination as well as reducing the levels of endogenous GA (Ogas et al. 1997, Li et al. 2005). *PKL* and GA-response mutants have similar phenotypes. *IMBI* acts during imbibition to adjust the ABA response factor *ABI5* (Duque & Chua 2003). Connections between ABA signalling and phytochrome response have been made based on the impairment of the PhyA-mediated germination in *IMB1* mutants (Duque & Chua 2003).

1.8.1.6 Auxin and cytokinins

The effect of auxin on seed germination is not very well known, particularly at the molecular level (Kucera et al. 2005). Although there is no clear evidence to suggest that auxin is directly involved in the induction of seed dormancy or its release, it is well established that the hormone plays an important role in embryo development (Feurtado & Kermode 2007). Changes in the levels of indole-3-acetic acid (IAA) have also been observed during various seed development stages (Bewley & Black 1994, Ficher-Iglesias & Neuhaus 2001) and in germination (Chiwocha et al. 2005). The germination of excised embryos of dormant wheat caryopses has been shown to be inhibited by IAA. These embryos also became less sensitive to the hormone during after-ripening (Ramaih et al. 2003). Variations in auxin levels and transport has also been linked to GA in the germination of Arabidopsis seeds (Ogawa et al. 2003, Kucera et al. 2005).

Similarly to auxin, whether cytokinin plays a role in seed germination or dormancy is unclear. However, the hormone has been used to release dormancy in numerous species

including maple (*Acer)*, apple (*Malus*), beech (*Fagus*) and peanut (*Arachis*) (Cohn & Butera 1982). In Arabidopsis, hormonal interactions between cytokinin, ethylene, auxin and GA have been reported (Su & Howell 1992, Greenboim-Wainberg et al. 2005, Pickett et al. 1990). While no auxin or cytokinin-related mutants have been observed to demonstrate significant germination or dormancy phenotypes, the involvement of these hormones cannot be discounted in light of the evidence described above. This thesis includes some auxin and cytokinin mutants with roles in auxin and cytokinin signalling described in Table 1.1.

1.8.1.7 Other seed dormancy mutants

ATPER1 is a stress related gene encoding peroxiredoxin that is expressed only in dormant seeds (Aalen 1999) There is some debate as to whether peroxiredoxin functions in dormancy (Goldmark et al. 1992, Haslekas et al. 1998) as not much is known about inter-related pathways between stress and dormancy signalling.

Leon-Kloosterziel (1996b) isolated two genes *RDO1* and *RDO2* with reduced dormancy phenotypes. These mutants do not display a sensitivity to the hormones ABA, ethylene, auxin and cytokinin that differs from the wild type, however, *RDO2* has a lower level of sensitivity to tetcyclacis, a gibberellin biosynthesis inhibitor (Leon-Kloosterziel 1996b) It was suggested that the mutants have independent pathways ending in dormancy and that the role of *RDO2* is induced by ABA and expressed as a greater need for GA in germination.

Table 1.1: Table of selected Arabidopsis mutants related to seed germination/dormancy identifying genes with various roles in hormone biosynthesis, signalling, response and action.

Hormone	Gene(s)	Mutant class	Function	Mutant Phenotype	Reference
ABA	*ABA1* *ABA2* *ABA3*	ABA deficient	ABA biosynthesis.	Reduced level of endogenous ABA during plant and seed development, reduced seed dormancy, decreased sensitivity to the presence of GA biosynthesis inhibitors	Leon-Kloosterziel et al. (1996a), Schwartz et al. (1997), Xiong et al. (2002), Seo et al. (2006)
	ABI1 *ABI2* *ABI3* *ABI4* *ABI5* *RPK1*	ABA insensitive	ABA signalling.	Reduced sensitivity to ABA inhibition of germination, increased level of endogenous ABA during seed development, reduced seed dormancy.	Finkelstein (1993), Nambara et al. (1992), Nambara et al. (1994), Nambara et al. (1995), Hong et al. (1997), Finkelstein et al. (1998), Arenas-Huertero et al. (2000), Lopez-Molina et al. (2001), Merlot et al. (2001), Osakabe et al. (2005), Verslues & Bray (2006)
	ABH1 *ERA1*	ABA hyper-sensitive	ABA signalling	ABA hypersensitive, decreased germination efficiency.	Cutler et al. (1996), Hugouvieux et al. (2001)
	MARD1	Mediator of ABA-regulated dormancy 1	zinc-finger protein	Rapid germination, reduced dormancy and germination in total darkness, germination resistant to exogenous ABA at radicle protrusion.	He & Gan (2004)
	ATRAB28	*Arabidopsis thaliana* response to abscisic acid 28	ABA signaling, ion cell balance in germination	Reduced germination rates under standard, salt or osmotic stress conditions. Reduced cation toxicity tolerance in germination.	Borrell et al. (2002)
GA	*GA1* *GA2* *GA3* *GA4* *GA5*	GA requiring	GA biosynthesis	Dwarf, responsive to gibberellin stimulus. Some of these mutants germinate without GA treatment.	Sun & Kamiya. (1994), Yamaguchi, et al. (1998ab), Helliwell et al. (1999), Hisamatsu T, et al. (2005), Kim et al. (2005)

	SPY	Spindly	GA signalling	Altered GA signal transduction, resemble wild-type plants that have been repeatedly treated with GA, resistant to the GA biosiynbthesis inhibitor paclobutrazol. Resistant to exogenously applied cytokinin.	Jacobsen et al. (1996), Jacobsen, et al. (1998), Greenboim-Wainberg et al. (2005)
	GCR1	GA signalling	protein similar to G-coupled receptor	Reduced seed sensitivity to GA and brassinosteroid, altered seed germination.	Colucci et al. (2002)
	GAI	GA insensitive	Putative transcription factor. *DELLA*-type *GRAS* transcription factor	Gibberellin-insensitive dwarf. Increased endogenous levels of GA, reduced seed germination	Peng et al. (2002),
	RGL1	*RGA*-Like (Repressor of GA1-3).	*GRAS* transcription factor	GA-independent activation of seed germination.	Wen et al. (2002)
Brassinosteroid	*DWF1* *DWF4* *DWF5* *DWF7*	Brassinosteroid deficient	Brassinosteriod biosynthesis	Dwarf, brassinosteroid deficient.	Choe et al. (1999ab), Choe et al. (1998), Klahre et al. (1998), Choe et al. (2000)
	BIN3	Brassinosteroid insensitive	Brassinosteroid signalling	Response to brassinosteroid stimulus.	Yin et al. (2002)
	BIN2	Brassinosteroid insensitive	Auxin and brassinosteroid signalling pathways.	Dwarf, brassinosteroid insensitive.	Li et al. (2001)
	BRI1	Brassinosteroid insensitive	Brassinosteroids signal transduction	Extreme dwarf, accumulates high levels of brassinolide, sensitive to auxins, defective in development pathways that cannot be resuced with brassinosteroid treatment.	Mora-Garcia et al. (2004)
	DET2	Detiolated2	Brassinosteroid biosynthesis	Reduced gemination rates.	Steber & McCourt (2001)

Ethylene	*EIN2* *EIN3*	Ethylene insensitive	Ethylene signalling	Resistant to auxin and ethylene, germination inhibited by salt and osmotic stress. Reduced germination in the presence of mannitol.	Chao et al. (1997), Alonso et al. (1999), Ghassemian et al. (2000)
	EBF1	F-box protein, regulatory protein.	Ethylene signalling	Enhanced ethylene response. Overexpression of gene show ethylene insensitivity. Epistatic with *EIN3*.	Guo & Ecker 2003
	ETR1 *ETR2*	Ethylene insensitive	Ethylene signalling	Ethylene response.	Chang et al (1993), Rodriguez et al. (1999), Chiwocha et al. 2005. , O'Malley et al. (2005)
	CTR1		Ethylene signalling	Resistant to the negative effects of paclobutrazol (inhibitor of gibberellin biosynthesis on seed germination), response to ethylene and sucrose stimulus, slightly ABA resistant.	Kieber et al. (1993), Gibson et al. (2001)
Auxin and cytokinin	*AUX1*	Auxin-resistant	Auxin and ethylene signalling	Resistant to ethylene and auxin.	Pickett et al. (1990)
	AXR1 *AXR2* *AXR3* *AXR6*	Auxin-resistant	Auxin signalling.	Increased amplitude of auxin responses, resistant to auxin, less sensitive to brassinolide, ethylene and ABA. *AXR6* embryonic development ends in seed dormancy.	Abel et al (1995), Rouse et al. (1998) , Hobbie et al. (2000), Nagpal et al. (2000), Ouellet et al. (2001) , Liscum et al. (2002)
	ABP1	Auxin binding protein	Auxin binding protein	Cytokinin and auxin response.	Jones et al. (1998)
	IAR3	IAA-alanine resistant	IAA-amino acid conjugate hydrolase subfamily	Reduced sensitivity to IAA-Ala. Plants overexpressing *IAR3* are more sensitive to certain IAA-amino acid conjugates.	Davies et al. (1999)
	IAR1	IAA-alanine resistant	Member of IAA-alanine resistance protein 1.	Resistant to the inhibitory effects of several IAA-amino acid conjugates but remains sensitive to free IAA.	Lasswell et al. (2000)

	ILR1	IAA-leucine resistant	Hydrolyzes amino acid conjugates of the plant growth regulator indole-3-acetic acid (IAA), including IAA-Leu and IAA-Phe.	Resistant to indole-3-acetic acid-leucine conjugate, slightly resistant to IAA-phenylalanine.	Bartel et al. (1995), LeClere et al. (2002)
	CRE1	Cytokinin response	Cytokinin signalling	Reduced sensitivity to cytokinin.	Riefler et al. (2006)
	HYL1	Hyponastic leaves	dsRNA binding protein.	Altered vegetative morphology and response to hormones. Less sensitivity to auxin, cytokinin and hypersensitivity to ABA.	Lu & Federoff (2000)
Transcription Factors	*LEC1* *LEC2*	Leafy cotyledon	Transcription factor	Embryos sensitive to ABA, mature seeds are desiccation intolerant and germinate poorly.	Meinke et al. (1994), Lotan et al. (1998), Baumbusch et al. (2004)
	DAG2	GA insensitive	DOF transcription factor	Less sensitive to germination-promotive effect of gibberellins, more dependent on physical stimuli (light and cold treatment) for germination.	Gualberti et al. (2002)
	DAG1		DOF transcription factor	Altered response to red and far-red light. Germination by much lower red light fluence rates and germination reaches a point where it is independent of phytochrome signalling more quickly.	Papi et al. (2002)
	FUS3	Fusca3	Transcription factor	Seeds are desiccation sensitive, dry seed is shrunken and does not germinate, defective in late embryo development and germination, germination is inhibited by exogenous ABA.	Baumlein et al. (1994), Parcy et al. (1997), Gazzarrini et al. (2004)
	IMB1	Imbibition-inducible 1	Transcription factor	IMB1 expressed at low levels in dry seeds but markedly induced during seed imbibition. Transcript levels are down regulated during germination. Mutants show impaired cotyledon greening during germination in ABA.	Duque et al. (2003)

	PKL	Pickle	CHD3 protein, chromatin-remodelling factor, negative regulator of transcription.	Phenotype reminiscent of GA-deficient plant, reduced stature, root phenotype enhanced by inhibitors of GA biosynthesis.	Ogas et al. (1997)
Dormancy mutants	*ATPER1*	Fatty acid transporter protein	1-cys peroxiredoxin (extensin like gene)	Expression is limited to seed (aleurone and embryo). Not induced by ABA or drought.	Dubreucq et al. (2000)
	CTS/PXA1	Comatose	LRR Receptor Kinase	Germination potential highly reduced in mutant seeds. Inability to break down lipid bodies.	Russell et al. (2000), Footitt et al. (2002)
	RDO1 *RDO2*	Reduced seed dormancy 1 and 2		Reduced dormancy, does not require GA, high capacity for germination in the dark, freshly harvested seeds will germinate.	Leon-Kloosterziel et al. (1996a)

1.8.2 Analysis of natural variation

Although the use of mutants has been a powerful genetic tool in performing gene function analyses, there are limitations due to the small number of genetic backgrounds analysed in this approach (Koornneef et al. 2004). The mutant phenotypes produced represent the expressed alleles of the wild-type accession and may not be indicative of the variation in phenotypes that occur in natural populations. Alleles of the wild-type accession that are null or weakly expressed may not be detected (Koornneef et al. 2004). Epistatic interactions also influence the expression of phenotypes in certain genetic backgrounds (Koornneef et al. 2004). Therefore, genetic analysis of natural variation is necessary to better define the functions of individual genes and identify the mechanisms operating under natural conditions and in specific environments.

Naturally occurring genetic variation for complex traits such as seed dormancy is often of a quantitative nature, showing a continuous (phenotypic) distribution between the extremes of the trait in segregating populations (Langridge et al. 2001, Koornneef et al. 2004). This distribution pattern arises from the effects of alleles at various loci, each locus contributing to the overall phenotypic effect, and/or the influence of the environment on the various loci. Different combinations of alleles at these loci result in different phenotypes (Langridge et al. 2001). Thus, the resulting phenotype is derived from the combined action of all the alleles affecting the trait, the environmental effects and their interaction. The loci, referred to as quantitative trait loci (QTLs), are also often interactive with each other. Current and continuing advances in molecular genetics and statistical techniques make it possible to perform QTL mapping, enabling an estimation of the number, genome location and gene effect of the QTL for a segregating population (Koornneef et al. 2004).

1.8.2.1 Quantitative Trait Loci mapping

QTL mapping involves several steps: 1) the generation of experimental population for marker mapping, usually from homozygous, inbred parental lines that differ in the alleles that affect variation in the trait studied; 2) a linkage map of polymorphic markers is then constructed; 3) the individuals in the population are scored for their genotype with genetic markers as well as for the trait phenotype; 4) analysis is performed to find associations between the phenotypic values of the trait and genotypic classes of the polymorphic markers (Mauricio 2001, Koornneef et al. 2004). The more detailed the genetic linkage map is (i.e. the greater the number of polymorphic markers placed on the map), the better the mapping resolution of the QTL.

Mapping QTLs in a population can be performed using several techniques. In a simple linear regression, also known as single point analysis, the significance of the relationship between the phenotype and each individual marker tested is analysed (Tanksley 1993, Mauricio 2001). A molecular marker linkage map is not necessary for single point analysis (Tanksley 1993). Due to recombination events between the marker locus and QTL, the regression technique tends to underestimate the effect of the QTL as the strength of the effect is based on the distance between the QTL and the marker. There are also limitations in pinpointing the QTL location on the chromosome. These problems can be minimised by obtaining good coverage of the entire genome with multiple markers.

A standard approach for QTL mapping that is available in most software programmes is interval mapping using maximum likelihood (Lander & Botstein 1989) or regression methods (Hayley & Knott 1992, Martinez & Curnow 1992) A larger number of markers

is typically used in interval mapping, which uses sets of linked markers in the analysis and determines the probability that an interval between two markers is linked to a QTL affecting the phenotypic variation (Soller et al. 1979). This method compensates for recombination between markers and increases the statistical probability of finding a QTL. While interval mapping is more powerful than single-point mapping, the technique is not without limitations. It is assumed that markers are segregate randomly, however, in the event of segregation distortion, estimates of recombination and subsequent detection of QTL locations can be affected. Markers linked the trait of interest can affect the ability of the program to identify certain QTLs within the marker interval of interest or can result in the detection of false positive QTL (Mauricio 2001).

Composite interval mapping combines both interval mapping and multiple regression analysis (Jansen 1993, Zeng 1993, 1994). Composite interval mapping is considered to be more accurate as it measures the probability that a marker interval is associated with a QTL while taking into account the effects of other markers on the trait of interest (Mauricio 2001). The method attempts to assess QTLs independently of the genetic variation in other chromosomal regions (Mauricio 2001, Verbyla 2007).

Multiple interval mapping is an extension of composite interval mapping techniques (Kao et al. 1999, Zeng et al. 1999). This technique uses an algorithm to analyse multiple QTLs and the genetic architecture to characterise QTLs with individual effects and epistasis (Kao et al. 1999, Mauricio 2001). This method is able to identify the additive effects of multiple QTLs, their positions and the interactions that occur between them. A recent method using the framework of mixed linear modelling developed by Yang et al. (2008) integrates the additive effects of multiple QTLs,

epistasis, QTL by environment interactions and epistasis by environment interactions into a single mapping system.

As complex quantitative traits such as seed dormancy are often regulated by several genes and gene interactions, the number of QTLs and the possible epistasis between QTLs confound the ability to identify all regions associated with the phenotype (Doerge 2002). In general, large sample sizes and high resolution mapping with good coverage of all chromosomes increases confidence in QTL mapping (Doerge 2002). False positive QTLs or minor QTLs detected in several studies using different environments or populations would increase the reliability of these QTLs.

1.8.2.2 Quantitative Trait Loci populations

Various crossing methods are used to produce the segregating population for QTL mapping of a trait of interest. Populations that consist of virtually homozygous lines such as recombinant inbred lines (RILs, created by selfing the F1 progeny for several generations) or backcross inbred lines (BILs, created by crossing the F1 progeny to either or both of the parents) have the advantage of being able to be replicated with genetically identical material and retested, therefore making the measurement of a QTL more accurate and enabling testing under different environmental conditions (Tanksley 1993, Koornneef et al. 2004). Double haploid (DH) populations have proved to been useful in studying complex traits with over 130 quantitative traits having been mapped in nine crop species (Thomas et al. 2003). DH lines can be produced by subjecting immature pollen grains (haploids) to treatments that result in the doubling of the gamete genetic material and therefore pure homozygous lines. Double haploid populations are ideal for genetic mapping and for studying quantitative inheritance as linkage

disequilibrium is maximised (Lynch & Walsh 1998). The fixed nature of DH populations enables them to be phenotyped in a range of environments and for various traits (Lehmensiek et al. 2008). Generally, the larger a population is, the more conducive it is to representing the variation present in the trait and results in better map resolution (Lehmensiek et al. 2008).

1.8.2.3 Phenotyping

Methodology for genotyping and statistical analyses of complex traits are advanced, however, identifying the genetic components controlling these traits are reliant on the ability to accurately measure the phenotype of the trait. A major limitation is the variability that exists in the trait assessment. In cereals, seed dormancy measurements or grain susceptibility to preharvest sprouting has been generally been assessed based on alpha amylase measurements (Marquez-Cedillo et al. 2000, Zanetti et al. 2000), the presence of sprouted grains on cereal spikes (Anderson et al. 1993, Groos et al. 2002) or germination tests using harvested grain (Ullrich et al. 1993, Oberthur et al, 1995, Han et al. 1996, Takeda 1996, Thomas et al. 1996, Gao et al. 2003, Li et al. 2003a, Prada et al. 2004, Ullrich et al. 2008).

Alpha amylase is an enzyme that degrades starch to sugars. Preharvest sprouting, which is indicative of low seed dormancy levels, is associated with alpha amylase concentrations that are unacceptably high (Gale 1989). Alpha amylase activity is commonly measured using the Hardberg falling number (HFN) test (Hagberg 1960). The test measures the time it takes (in seconds) for a plunger to pass through a gelatinised wholemeal paste in hot water after a standard mixing time. A high HFN (>400 s) represents low alpha amylase activity in the grain whereas low HFN (62-400 s)

indicates excessive alpha amylase activity. While alpha amylase activity is associated with preharvest sprouting, it may not be entirely indicative of seed dormancy levels. Alpha amylase activity is correlated to water uptake and drying of the grain (Bentsink et al. 2007) and late maturity alpha amylase activity occurs independently of germination (Flintham 2000).

Assessment of germination or sprouting of grain on mature spikes to evaluate dormancy indirectly take into account rainfall, humidity, changes in temperature as well as spike structure, plant architecture and variation in spike maturity and disease effects (King & Richards 1984, King 1989, Mares 1989, 1993b). Measurement of sprouted spikes additionally includes the effects of the maternal vegetative tissue on dormancy (Mares & Mrva 2001). Heritability of germination differences in grains has been shown to be low in intact spikes compared to isolated grains (Trethowan 1995). Despite this, Mares & Mrva (2001) found a linear relationship between the time for intact spikes to sprout and the germination rate of isolated grains.

Most dormancy assays are based on germination tests that basically measure the absence of germination. While dormancy is a quantitative trait and can vary from being non-dormant to dormant and anything in between, germination is counted as an all or nothing event (Finch-Savage & Leubner-Metzger 2006). However, seed germination tests remain the most commonly used method of measuring dormancy and comparisons between different dormancy studies can only be made using the same assessment tool. Greater knowledge concerning the elements controlling dormancy increases the potential of finding a more accurate and discerning way of defining and measuring dormancy in the future.

A major problem in phenotyping seed dormancy is the interactive effect of the environment. Like most field studies, difficulties faced include the inability to maintain environmental control, particularly for seed dormancy, which is strongly affected by environmental conditions. In addition, variations in harvest ripeness, anthesis time, heading date, and the time of harvest can result in genotype by environment interactions expressed in the phenotype (Gale 1989). Dormancy needs to be observed under different environmental conditions to examine the genotypic effects under various environments. It is impossible to remove all extraneous effects, however, taking into account the genotype by environment interactions will enable better understanding of the regulation of dormancy under different conditions.

1.9 Seed dormancy Quantitative Trait Loci in model species

Recently developed molecular markers and their linkage maps have enabled QTLs of polygenic traits such as seed dormancy to be defined and mapped in some important species (Tanksley 1993). Seed dormancy QTLs have been identified in previous studies in Arabidopsis (van der Schaar et al. 1997, Alonso-Blanco et al. 2003, Clerkx et al. 2004), rice (*Oryza sativa* L.) (Miura et al. 2002b, Lin et al. 1998, Cai & Morishima 2000, Dong et al. 2003, Gu et al. 2004), wheat (*Triticum aestivum* L.) (Roy et al. 1999, Kato et al. 2001, Mares & Mrva 2001, Flintham et al. 2002, Groos et al. 2002) and barley (*Hordeum vulgare* L.) (Ullrich et al. 1993, Oberthur et al. 1995, Han et al. 1996, Larson et al. 1996, Thomas et al. 1996, Li et al. 2003a, Prada et al. 2004, 2005). The following sections of this review describe the QTLs identified in the model plants Arabidopsis and rice, as well as in wheat and barley. Table 1.2 summarises the QTLs identified in various studies of rice, wheat and barley.

1.9.1 Arabidopsis

There have been a limited number of studies conducted on seed dormancy QTLs in Arabidopsis (van der Schaar et al. 1997, Alonso-Blanco et al. 2003, Clerkx et al. 2004). van der Schaar examined QTLs in a population derived from a cross between (Ler) and another laboratory accession Columbia (Col) grown in three independent greenhouse experiments and under various germination conditions. Despite the small differences in germination behaviour (both ecotypes demonstrate low levels of seed dormancy), fourteen QTLs (of rather small effect) were found to be associated with seed dormancy control. Nine of these QTLs were detected in all the three different environments tested while five were detected only in specific environments. In a similar study, this time crossing the strongly dormant Cape Verde Islands (Cvi) accession with the Ler accession, seven QTLs termed DOG (Delay of Germination) were revealed in the genetic variation analyses (Alonso-Blanco et al. 2003). The seven QTLs appeared to coincide to similar genomic regions as those in van der Schaar's study, indicating that the limited number of loci might account for the natural variation in seed dormancy in Arabidopsis.

Table 1.2: Summary table of Quantitative Trait Loci (QTL) studies in rice, wheat and barley

REFERENCE	POPULATION	POPULATION TYPE	QTL CHROMOSOME	PHENOTYPE MEASURED
Rice				
Lin et al. (1998)	Nipponbare/Kasalath//Nipponbare	BC1F5	3, 5, 7(x2), 8	germination rate
Wan et al. (1997)	Milyang23/Todorokiwase	F2	3, 6, 7	germination percentatge, husk intact
Wan et al. (1997)	IR36/Nekken2//Miyukimochi	F1	6, 7, 12	germination percentatge, husk intact
Cai & Morishima (2000)	Pei-kuh/W1944	F2	1(x4), 2(x2), 3(x3), 5(x2), 6(x3), 7, 8, 9, 11(x5), 12	germination percentage (intact and dehulled)
Miura et al. (2002b)	Nipponbare/Kasalath	BC1F9	1, 3, 5, 7, 11	germination percentage
Cai & Morishima (2002)	*O. sativa* (Indica)/*O. rufipogon*	RILs	1(x3), 2, 3(x3), 5(x2), 6(x3), 8(x2), 9(x2), 11(x3), 12	germination percentage
Dong et al. (2003)	Asominori/IR24	RILs (F6)	1(x2), 4, 5, 7, 8	sprouted grains
Gu et al. (2004)	EM93-1//EM93-1/SS18-2	BC1	4, 6, 7(x2), 8, 12	germination percentage
Gu et al. (2005)	ss18-2/EM93-1	BC1	4, 7(x2), 8, 12	germination percentage
Wan et al. (2005)	IR50/Tatsumimochi//Miyukimochi	F1	1, 3, 7	germination percentage
Wheat				
Anderson et al. (1993)	NY6432-18/Clark's Cream	F5 derived RILs	1AS, 2S	germination on harvested spikes
Anderson et al. (1993)	NY6432-18/NY6432-10	F5 derived RILs	3BL, 4AL, 5DL, 6BL	germination on harvested spikes
Roy et al. (1999)	SPR8198/HD2329	RILs (F6)	6B, 7D	germination on harvested spikes
Zanetti et al. (2000)	Wheat/Spelt	F5 derived RILs	2A, 3B, 5A (x2), 6A, 7B	falling number and alpha-amylase activity
Kato et al. (2001)	AC Domain/Haruyutaka	DHLs	4ABDL	germination percentage
Mares & Mrva (2001)	Cranbrook/Halberd	DHLs	2AL, 2DL, 4AC/L	seed germination (weighted germination index)
Groos et al. (2002)	Renan/Recital	F7 RILs	3A, 3B, 3D, 5A	germination on harvested spikes
Miura et al. (2002a)	Zen/CS	BC1F3	1AD, 2BD, 3AD, 4ABD, 6AD, 7BD	germination percentage
Osa et al. (2003)	Zen/CS	RILs (F7)	3AL, 3AS	germination percentage
Kulwal et al. (2004)	W-7984 /Opata 85	RILs	2BL, 2DS, 3BL, 3DL	sprouted grains
Mori et al. (2005)	Zen/CS	RILs (F8)	3AS, 4AL, 4BC,	germination rate
Torada et al. (2005)	Kitamoe/Munstertaler	DHLs	4AL	germination rate (not clear if its percentage)
Nakamura et al. (2007)	*Triticum monococcum* L. (KT3-5)/*T. boeoticum* L. (KT1-1)	RILs	5AL, 3AL (x2), 4AL	seed germination (weighted germination index)
Ogbonnaya et al. (2008)	CN19055/Annuello	RIL(F8)	4AL	germination index, sprouting index, visual sprouting
Ogbonnaya et al. (2008)	CN19055/Annuello	RIL(F8)	4ACS	visual sprouting
Barley				
Ullrich et al. (1993)	Steptoe/Morex	DHLs	1H (x2), 2H, 4H, 5HL (x2), 7H	germination percentage
Oberthur et al. (1995)	Steptoe/Morex	DHLs	1H, 4H, 5HL (x2), 7H	germination percentage
Han et al. (1996)	Steptoe/Morex	DHLs	5H(x2)	germination percentage
Takeda (1996)	Harrington/TR306	DHLs	1H, 5H	germination percentage
Thomas et al. (1996)	Blenheim/E224/3	DHLs	2H, 3H, 5H, 6H, 7H	germination percentage
Marquez-Cedillo et al. (2000)	Harringtion/Morex	DHLs	4HS, 5HL	alpha-amylase activity
Gao et al. (2003)	Steptoe/Morex	DHLs	5HL	seed germination percentage
Li et al. (2003a)	Chebec/Harrington	DHLs	2HC, 5HL	seed germination percentage
Edney & Mather (2004)	Harrington/Morex	DHLs	2H, 5H(x2), 7H	germination percentage
Prada et al. (2004)	Triumph/Morex	DHLs	2H, 3H, 5H, 5HL	seed germination percentage
Zhang et al. (2005)	*Hordeum spontaneum* /*H. vulgare* cv.Mona	F4	1H, 2H, 4H, 5H(x2), 7H	germination rate
Ullrich et al. (2008)	Steptoe/Morex	DHLs	2HS, 3HC, 4HS, 5HC, 5HL (X2), 7HC	preharvest sprouting (mist chamber treatment)
Ullrich et al. (2008)	Steptoe/Morex	DHLs	2HS, 3HC, 4HS, 5HC, 5HL(x2)	alpha-amylase activity
Ullrich et al. (2008)	Steptoe/Morex	DHLs	1HL, 2HS, 3HC, 4HS, 5HC, 5HL(x2), 7HC	germination percentage

Population type: recombinant inbred lines (RILs), double haploid lines (DHLs), backcross lines (BC) and various crossing generations (F1-F4).

1.9.2 Rice

The number of rice seed dormancy QTLs and their genomic locations are varied across different studies, reflective of the wide variation in the degree of seed dormancy of rice cultivars (Sheshu & Sorrells 1986) and the influence of environmental and genetic factors. Practically all the rice chromosomes have been implicated in controlling seed dormancy, with more than 30 putative QTLs reported in studies using populations derived from cultivated rice (Wan et al. 1997, Lin et al. 1998, Miura et al. 2002b, Dong et al. 2003), wild rice (Cai & Morishima 2000) and weedy rice (Gu et al. 2004). In an early study, Wan et al. (1997), using simple linear regression for QTL analysis, established that isozyme loci Pgi1 on chromosome 3, Amp3, Est2 and morphological locus C for apiculus color on chromosome 6, Est9 on chromosome 7 and Acp2 on chromosome 12 were linked to genes for seed dormancy. Lin et al. (1998) using RFLP markers, reported on five QTLs affecting seed dormancy on chromosomes 3, 5, 7 (in which there were 2 QTLs) and 8 in BILs derived from a cross between the japonica cultivar Nipponbare and the indica cultivar Kasalath. However, the threshold level employed was low and only the QTL on chromosome 3 was associated with a relatively high LOD score. Comparisons between the studies are difficult due to the different markers used. Miura et al. (2002b) followed this study with another QTL analysis on the same BILs and similarly detected QTLs on chromosomes 3, 5 and 7 and additional QTLs on chromosomes 1 and 11. The QTLs on chromosomes 5 and 7 from both studies appeared to coincide, however, although both studies detected a QTL on chromosome 3, these did not co-locate to the same chromosomal region. The remaining QTLs were not in common between the two studies. The contrasting QTL locations may be due to the effect of environmental conditions on the genotype, considering the studies were conducted 4 years apart. The seeds of Nipponbare in the study by Miura et

al. (2002b) also demonstrated a stronger dormancy in comparison to the seed used in the study by Lin et al. (1998).

Cai & Morishima (2000) studied the QTL effects of hull-imposed and kernal dormancy over different years and post harvest conditions using RILs obtained from a cross between cultivated Taiwanese indica rice, Pei-kuh, and wild Asian common wild rice, *O. rufipogoon* Griff. They used mostly RFLP markers in their QTL analysis which was performed using simple interval mapping followed by composite interval mapping to lessen the influence of residual variation due to the genetic background. The resulting 17 dormancy QTLs were detected in all chromosomes with the exception of chromosome 4 and 10, many QTLs occurring more than once on a chromosome. There appeared to be no clear distinctions between the QTLs depicted in intact seed compared to de-hulled seed, however, the loci that were significant only under one condition were assumed to be responsible for type of dormancy induced. Dormancy QTLs detected in this study seemed to co-locate with at least 3 seed dormancy QTLs found by Wan et al. (1997) and 4 by Lin et al. (1998) although further research is needed to confirm this.

More recently, six QTL associated with seed dormancy were observed in an RI population derived from a cross between the weakly dormant Asominori (japonica cultivar) and the strongly dormant IR24 (indica cultivar) using RFLP markers (Dong et al. 2003). A combination of both composite interval mapping and simple linear regression was also used for the QTL analysis. These QTL were mapped to chromosomes 1 (two loci), 4, 5, 7 and 8. The QTL on chromosomes 1, 4 and 5 were detected in the three environments tested. The loci on chromosomes 4 and 5 coincided with or were linked to QTLs observed previously in Miura et al.'s (2002b) study.

Likewise, the putative QTLs on chromosomes 7 and 8, which were only found in certain environments, correlated to the QTLs observed by Lin et al. (1998).

Strong intergenic interactions were observed between the six dormancy QTLs detected in a cross between a cultivated (low dormancy) and a weedy (high dormancy) cultivar (Gu et al. 2004), giving an indication of the complexity of the gene network regulating seed dormancy. Two QTLs were located on chromosome 7, the remaining loci found on chromosomes 4, 6, 8 and 12. The QTLs on chromosome 12 and 7 demonstrated the largest effect and were depicted in the 3 after-ripening days tested. Comparisons with other studies suggested that the second of the two loci on chromosome 7 corresponded to the QTLs identified in the four previous studies on cultivated rice populations (Wan et al. 1997; Lin et al. 1998; Miura et al. 2002, Dong et al. 2003) and both loci on chromosome 7 appeared to correspond to the two linked QTLs in Lin et al.'s (1998) study. The first loci on chromosome 7 was also found to be tightly linked to the red pericarp colour gene Rc (Gu et al. 2004). Gu et al. (2005) also found interrelations between seed dormancy and the weedy characteristics of seed shattering, awn and black hull colour.

1.9.3 Wheat

The germination-restrictive properties of the seed coat have been positively correlated with seed coat colour produced by phenolic compounds in a number of plant species (Debeaujon et al. 2000). Dormancy in wheat has been related to grain colour in several studies. Strong dormancy is associated with red seed coat colour, while wheat with white seed coats have been shown to be less dormant and more susceptible to preharvest sprouting (Gfeller & Svejda 1960, Mares 1993a). In red-grained wheat, the

genes controlling the red testa pigmentation (*R*) are located on the homoeologous group 3 chromosomes (Flintham & Gale 1996) and have been shown to enhance grain dormancy (Mares 1993a, 1999; Flintham et al. 1999, Warner et al. 2000, Himi et al. 2002). Flintham (2000) used near isogenic lines of white-grained NS-67 wheat carrying a single *R* gene on one of the group 3 chromosomes, and associated increased levels of grain dormancy to the presence of the *R* gene on chromosome 3. Mares (1999) found the phenotypes of red grained progeny from a dormant, red grained and non-dormant white wheat cross ranged from dormant to non-dormant, while no dormant white grained progeny resulted from the same population. In addition, mutation of the *R* gene produces white-grained wheat with reduced dormancy (Mares 1999, Warner et al. 2000). QTLs for grain colour and PHS were mapped in a Renan (red grained, strong resistance to PHS) x Recital (white grained, highly susceptible to PHS) population and the detected QTLs on the group 3 chromosomes and on the short arm of 5A collocated for the two traits (Groos et al. 2002). Himi et al. (2000) disagreed with speculations that the *R* genes increased grain dormancy by accumulating germination inhibitors (Miyamoto & Everson 1958, Stoy & Sundin 1976) by demonstrating that there was no difference in the amount of inhibitors in water extracts of AUS 1490, a red-grained line, and *EMS-AUS*, its white grained mutant. Instead, the authors suggested that the *R* gene may increase embryonic sensitivity to ABA, resulting in greater dormancy levels. The specific function of the *R* genes in enhancing seed dormancy is not yet clear and whether the relationship between seed dormancy and grain colour is a result of pleiotropy between the *R* genes or due to linkage between the genes controlling the two traits, remains to be clarified.

The maize *Viviparous-1* (*Vp1*) gene, which encodes the transcription factor VIVIPAROUS-1, is known to be involved in the control and maintenance of dormancy in maize (McCarty et al. 1991) and regulates the expression of a subset of abscisic acid (ABA)-regulated genes during seed maturation (Hill et al. 1996). The gene's homologue *afVp1* in wild oat (*Avena fatua*) (Jones et al. 1997) and the closely related gene *ABI3* in Arabidopsis (Giraudat et al. 1992) have also been shown to play an active role in dormancy control. In wheat, the *Vp1* homologues (*taVp1*) were mapped to a similar region as the *R* genes on the long arms of the group 3 chromosomes (Bailey et al. 1999). The two loci appear to be loosely linked but are genetically distinct, suggestive of separate roles in coat imposed dormancy mechanisms and embryo dormancy, respectively (Bailey et al. 1999). Several studies have implicated the *taVp1* gene in the control of seed dormancy in wheat (Nakamura & Toyama 2001, Groos et al. 2002, McKibbin et al. 2002, Wilkinson et al. 2002). The level of expression of the *taVp1* gene was found to differ between dormant and non-dormant cultivars of wheat and was positively correlated with level of seed dormancy and embryo sensitivity to abscisic acid (Nakamura & Toyama 2001). McKibbin et al. (2002) demonstrated that the embryos of transgenic wheat grains containing the *Avena fatua Vp1* were more sensitive to applied abscisic acid and PHS was reduced in transgenic wheat plants. In the same study, the incorrect splicing of the *taVp1* transcript in wheat was shown to result in greater susceptibility to PHS (McKibbin et al. 2002). Although QTLs have been mapped in the vicinity of the *taVp1* gene, no direct associations have been found between the gene and a QTL in wheat. In a study by Miura et al. (2002a) using backcross reciprocal monosomic population, the chromosome 3A was predominantly implicated in influencing seed dormancy variation in the highly dormant red grained wheat Zenkoujikomugi (Zen), followed by the group 4 chromosomes. In subsequent

studies under controlled and field conditions (Osa et al. 2003, Mori et al. 2005), a major QTL and a QTL of minor effect were confirmed on the short arm and long arm of 3A, respectively, which clearly accounted for differences in seed dormancy and were contributed by the highly dormant Zen parent. The effect of the QTL on 3AS was found to be independent of the *taVp1* and the *R-A1* genes (Osa et al. 2003). The minor QTL on the long arm of the chromosome, which was only evident at the dormancy-breaking stage, was loosely linked to *taVp1* but also remained distinct (Osa et al. 2003).

There have been several studies citing QTLs associated with preharvest sprouting tolerance and/or seed dormancy on the homoeologous group 4 chromosomes in wheat (Anderson et al. 1993, Kato et al. 2001, Mares & Mrva 2001, Miura et al. 2002a, Mori et al. 2005, Torada et al. 2005, Ogbonnaya et al. 2008). QTL analysis of seed dormancy in DHL population derived from a cross between AC Domain (Canadian hard red spring wheat with a strong degree of dormancy) and Haruyutaka (Japanese hard red spring wheat associated with low levels of seed dormancy) identified three QTLs on the group 4 chromosomes using interval mapping methods with RFLPs (Kato et al. 2001). The three QTLs accounted for over 80% of the phenotypic variation observed. The QTLs on chromosomes 4B and 4D were located in the terminal region of the long arms and appeared to be homologues, while the QTL on chromosome 4A located to the proximal region of 4AL and the authors suggested the possibility of a homoeologous relationship between this QTL and the barley gene SD4. Similarly, in addition to the QTL on chromosome 3A discussed earlier, Mori et al. (2005) also confirmed the detection of QTLs on the long arm of 4A and on the centromeric region of the long arm of 4B in association with seed dormancy in the red grained Zen wheat. Mares & Mrva (2001) also identified a QTL on chromosome 4AL, which appeared to

collocate to the QTL in Kato's et al. (2001) study, and also identified two other QTLs on chromosomes 2AL and 2DL in a DHL population from a Cranbrook (extremely dormant) and Halberd (intermediate dormancy) cross. The three QTLs had rather small individual effects on the dormancy phenotype, however, the LOD scores were greater than 3 and LRS values were significant ($P<0.05$). All three QTLs were also contributed by the susceptible parent Cranbrook leading to speculation that the alleles derived from Cranbrook are involved in promoting rapid germination (Mares & Mrva, 2001). Mares et al. (2005) also established the QTL on 4A in two white grained wheat genotypes and one red grained genotypes of diverse origin. Intermediate dormancy was associated with the QTL on 4A. Mares et al. (2005) concluded that alone, the QTL will not produce the dormancy phenotype characteristic of the dormant parents, however, combined with the R gene and/or other additional dormancy genes, the dormant phenotype is likely to result. Even more recently, Ogbonnaya et al. (2008) detected a major QTL on 4AL measuring both seed dormancy and preharvest sprouting. The detection of the 4A QTL in various genetic backgrounds indicates that it is a critical component of dormancy. Noda et al.'s (2002) study indicated the presence of a gene on 4AL associated with embryo sensitivity to ABA and dormancy. It is uncertain whether there is any correlation in the 4AL QTL and ABA sensitivity chromosomal regions. Their study additionally suggested chromosome 2D may also play a role in ABA sensitivity.

Earlier studies on QTL analyses showed evidence that other wheat chromosomes were involved in the control of seed dormancy. However, the methods may not have measured seed dormancy as such. Anderson et al. (1993) identified four preharvest sprouting QTLs on chromosomes 1A, 3B, 4A, 5D, 6B and on the group 2 chromosomes

in each of the two recombinant inbred line populations of white grained wheat, some of which were specific to certain environments. Roy et al. (1999) used microsatellite and STS markers to analyse an RIL population developed from across between SPR8198 (red grained wheat with prehavest sprouting tolerance) and HD2329 (white grained wheat susceptible to preharvest sprouting tolerance) and found linkage of an microsatellite marker and a STS marker on chromosomes 6B and 7D respectively with pre harvest sprouting tolerance. However, as preharvest sprouting tolerance was assessed while the grains were still on the spike in both these studies, the phenotypic effects measured would have included the effects of the maternal tissue of the spike (Mares & Mrva 2001). QTLs have been located on chromosomes 2A, 3B, 5A, 6A and 7B in a study by Zanetti et al. (2000), who used falling numbers and alpha-amylase activity in harvested grain from field environments to measure differences in preharvest sprouting tolerance rather than visible sprouting. Effects of environmental differences, grain maturity, spike morphology and late maturity amylase effects would be also be included (Mares & Mrva 2001, Flintham et al. 2002).

1.9.4 Barley

Classical genetic studies on barley, similar to Arabidopsis, rice and wheat; have shown that several genes, the environment and the interaction between them control the expression of dormancy in barley. Generally, barley is more resistant to pre-harvest sprouting than wheat genomes, however, this varies significantly depending on the cultivar and environmental conditions (Hoyle 1999). Due to the agronomic implications related to seed dormancy and the lack of it, there have been several comprehensive studies investigating seed dormancy QTLs in barley.

The majority of QTL analyses on seed dormancy in barley have been conducted using populations derived from a Steptoe x Morex cross, a major barley mapping cross of the North American Barley Genome Project (NABGP), due to the saturated molecular marker linkage map available (Kleinhofs et al. 1993). Steptoe is a strongly dormant, six-row, feed barley (Muir & Nilan 1973) while Morex is a six-row malting cultivar that exhibits low dormancy (Rasmusson & Wilcoxson, 1979). In an early study, Ullrich et al. (1993) reported on 27 loci spanning across the seven barley chromosomes associated with seed dormancy in several field environments, giving an indication of the complexity of the trait. Several studies that followed Ullrich et al. (1993), on populations of the same cross and using interval mapping techniques demonstrated four regions of the barley genome on chromosomes 7(5H) (two loci), 1(7H) and 4(4H) with prominent gene effects associated with the control of seed dormancy (Ullrich et al. 1993; Oberthur et al. 1995, Han et al. 1996, Larson et al. 1996). Of the four verified loci, designated SD1-SD4 by Han et al. (1996), SD1 in the centromere region of barley chromosome 7(5H), demonstrated the largest effect (accounting for over 50% of the phenotypic variance) and was consistent over several post harvest periods tested (Oberthur et al. 1995, Han et al. 1996). SD2 located in the telomere region on the long arm of chromosome 7(5H) explained approximately15% of the variation, while the QTLs on chromosomes 4(4H) and 1(7H) each accounted for 5% of the seed dormancy variation (Romagosa et al. 1999). Significant epistatic interactions were observed between SD1 and the other dormancy loci (Oberthur et al. 1995, Han et al. 1996; Larson et al. 1996). Romagosa et al. (1999) investigated the nature of the epistatic interactions between these QTLs during seed development and after ripening and found evidence that the QTLs were associated with the release of dormancy at the end of the development process and during after ripening.

Since the verification of the SD1-SD4 QTLs, several studies using different crosses have identified seed dormancy QTLs that coincide with the SD1 and SD2 chromosomal locations. A major dormancy QTL in the double haploid populations of a Harrington/TR306 cross (Takeda 1996) and a Chebec/Harrington cross (Li et al. 2003a, 2004), accounting for 58% and 61% of the phenotypic variation respectively, was located on the long-arm telomere region on chromosome 7(5H). A seed dormancy QTL was again detected in the same vicinity in a Triumph/Morex population, (lesser effect of 19% of variation, Prada et al. 2004). Prada et al. (2004) and Li et al. (2003) additionally identified a QTL in the centromere region on chromosome 7(5H) which corresponded to the SD2 locus.

Genomic regions in other barley chromosomes have also been implicated in the control of seed dormancy. Thomas et al. (1996) identified a QTL on chromosome 6(6H) that was consistently detected over 4 years in a European DH population of Blenheim/E22/3 and other QTLs on chromosomes 2(2H), 3(3H), 5(1H) and 7(5H) in certain environments. Studies by Li et al. (2003) and Prada et al. (2004) both detected QTLs on 2(2H) contributing small effects on seed dormancy variation. In addition to the QTL on 5H discussed previously, Takeda (1996) located a QTL on chromosome 5(1H) contributing to seed dormancy control in the Harrington/TR306 DH population he studied.

In addition to mapping the chromosome locations of the loci effecting seed dormancy variation, QTL analyses have enabled inferences on the timing of their activity. SD1 is influential in the regulation of dormancy during grain development while SD2 appears to act as a dormancy-release allele post-harvest (Prada et al. 2004).

1.10 Comparative genomics

In broad terms, comparative genomics involves assessment of the similarity between two or more species that may or may not be taxonomically related at the gene or genome level (Appels et al. 2004, Francki & Appels 2007). Nucleotide and protein similarity is a tool commonly used in comparative genomics to identify related genes that have descended from a common ancestral gene (homologs). Genes that have diverged by the event of speciation are defined as orthologs while paralogs are genes that have diverged by the event of genetic duplication (Francki & Appels 2007). Orthologs will usually (but not always) retain the ancestral function throughout evolution while paralogs may evolve new functions with divergence of the species. The problem faced in identifying whether two genes are orthologs or paralogs lies in the difficulty of establishing whether divergence of the genes has occurred before or after a speciation event.

Comparative genomics encompasses comparisons between gene structure, gene sequence (protein or nucleotide), gene density and gene/marker order on chromosomes. Syntenic regions are chromosomal segments between two genomes that contain a high level of similarity in gene content and gene order (Appels et al. 2004). Syntenic regions containing genes of agronomic importance have been identified across different cereal species (Gale & Devos 1998). Comparative mapping of barley, wheat, oats, maize and rice using RFLP analyses have demonstrated conserved marker order over large linkage blocks (Ahn et al. 1993, Kurata et al. 1994a, Sherman et al. 1995, Van Deynze et al. 1995, Gale & Devos 1998, Smilde et al. 2001). Aligning regions known to control economically important traits such as seed dormancy across species may assist in identifying potential candidate genes involved in the expression of these traits. Provided

co-linearity of gene order is maintained, species such as rice and Arabidopsis can be used as models to predict potential candidate gene locations in related species with larger more complex genomes such as the barley genome (Devos et al. 1999). This approach is examined in detail later in this thesis (Chapter 2). While the extensive studies on gene functions in Arabidopsis may reveal crucial genes involved in the metabolic processes of higher plants, rice is currently the model plant for mapping cereal genomes (Izawa & Shimamoto 1996) due to its small genome size (430 Mb, the smallest among cereals), the availability of detailed linkage maps (Causse et al. 1994, Kurata et al. 1994b), the large scale sequencing of expressed sequences (Sasaki et al. 1994, Havukkala et al. 1995) and most importantly the large amount of synteny found between rice and other cereals (Ahn & Tanksley 1993, Kurata et al. 1994a, Kilian et al. 1995, Moore et al. 1995).

With the number of increasing QTL studies, comparative genomic attempts are being made to establish whether the same QTLs are observed in different studies, and whether these QTLs map across cereal species. Identification of same QTL in different environments would indicate a fundamental gene(s) controlling the trait while other minor QTLs may be more indicative of gene(s) that interact in specific environmental conditions. In another approach, candidate genes are mapped to determine whether or not they locate to chromosomal regions in the vicinity of seed dormancy QTLs. The Vp1 is one such gene that has been used in this approach, however, results demonstrated that the gene was not associated with a dormancy QTL in wheat and sorghum. Combining these techniques and identifying genes mapping to QTLs that are common across species may highlight candidate genes that that play a fundamental role in controlling a trait such as seed dormancy for further study.

1.11 General thesis direction

Seed dormancy is clearly a complex trait that involves numerous genes influenced by both environmental and developmental factors. There are still many gaps in the knowledge surrounding the nature of crucial regulator(s) that initiate germination and those that block embryo development processes associated with germination, and their interaction. This thesis aims to elucidate some of the features with putative roles in controlling seed dormancy in barley. The objectives of the current project include:

1) identify candidate genes involved in seed dormancy using comparative genomics
2) identify QTLs and QTL interactions in controlling dormancy in a Stirling and Harrington barley cross
3) associate candidate genes with the QTL network identified.

This study combined the data resources available on seed dormancy genes and QTLs in the model species of Arabidopsis and rice to identify candidate genes to further investigate in barley. In order to transfer this comparative genomic study to barley, a barley population segregating for seed dormancy was studied through QTL analyses to detect sources of seed dormancy control. Mapping of the candidate genes in this population to QTLs of interest was planned in order to pinpoint regions controlling seed dormancy in barley and provide the basis for identifying candidate genes for further investigation. The identification of candidate genes was of particular interest to provide breeders with a tool for selection against preharvest sprouting in breeding programs.

CHAPTER 2: CANDIDATE GENE IDENTIFICATION USING COMPARATIVE GENOMICS

2.1 Introduction

Model species provide a valuable resource to understand particular biological trends with the expectation that the knowledge gained will potentially be transferred to other related organisms. Extensive studies of the dicotyledonous plant Arabidopsis in plant development, physiology and biochemistry have generated large amounts of biological data, making it a useful model for dicots and other plants in general. In 2000, Arabidopsis became the first plant genome to be sequenced (The Arabidopsis Genome Initiative 2000). Soon after, rice, a major grain crop was completely sequenced (Goff et al. 2002, Yu et al. 2002) providing a reference for monocot crop species such as maize, wheat and barley. Since sequencing of the two genomes, large volumes of genetic and molecular information on the two model species has become available and housed in public domain databases such as The Arabidopsis Information Resource (TAIR), *Arabidopsis thaliana* Plant Genome Database (AtGDB) and MatDB (MIPS Arabidopsis Database) for Arabidopsis; and Gramene (http://www.gramene.org/), TIGR (http://www.tigr.org/tdb/e2k1/osa1/) and KOME (Knowledge-based Oryza Molecualr biological Encyclopedia website) for rice. These public databases offer sequence information, gene and marker mapping, gene ontology and function characterisation, gene structure, protein annotation and phenotypic data. These resources provide the opportunity to compare genes characterised in the two model species for identifying orthologs in crop species using comparative genomics.

Divergence between monocots and dicots is thought to have occurred approximately 170-235 Myr ago (Wolfe et al. 1989, Bowers et al. 2003). Therefore, it is expected that

significant variation exists between Arabidopsis and rice in terms of genome size and the gene density. The genome of Arabidopsis is 125 Mbp and has an estimated 26,422 genes (The Arabidopsis Genome Initiative 2000) whilst the rice genome is 420-466 Mbp with an estimated 50,000-55,615 genes (Goff et al. 2002, Yu et al. 2002). The large variation in genome size can be partially explained by differences in gene number multiplication arising from gene, chromosome segment and whole genome duplication or alternately gene loss (Tikhonov et al. 1999, Blanc et al. 2000, Grant et al. 2000, Ku et al. 2000, Vision et al. 2000, Wendel 2000, ,Bennetzen & Ramakrishna 2002). Gene duplication provides an important source for genetic novelty (Lynch & Conery 2000) and plays a significant role in the evolution of new biological functions (Hurles 2004). Gene duplication has given rise to many transcription factors and target genes involved in the regulatory networks of complex traits (Teichmann & Babu 2004).

Seed dormancy is one such trait that is under complex polygenic control, involving genes in biochemical, physiological and metabolic pathways. It has been well established that plant hormones are extremely important in the regulation and maintenance of seed dormancy and germination (Koornneef et al. 2002, Finkelstein 2004, Kucera et al. 2005). Dormancy and germination appear to be primarily controlled by the action and interaction of the plant hormones, gibberellic acid (GA) and abscisic acid (ABA). Gibberellic acid has a strong positive effect, enhancing germination in many species while ABA inhibits germination (Hilhorst & Karssen 1992, Bewley & Black 1994). Other hormones such as ethylene, brassinosteroids, auxins, cytokinins and other signalling molecules also play a role in the induction and maintenance of dormancy and germination (McCarty 1995, Foley et al. 2001, Bentsink & Koornneef 2002, Koornneef et al. 2002, Kucera et al. 2005, Bentsink et al. 2007, Feurtado &

Kermode 2007). Gene mutations have successfully identified some key elements controlling hormone biosynthesis, catabolism, perception and signalling, by demonstrating deficiency, overproduction, insensitivity or hypersensitivity to the hormones, resulting in variations in dormancy and germination response.

Quantitative Trait Loci (QTL) analyses have been used as a means to identify chromosomal regions containing alleles controlling complex quantitative traits such as seed dormancy. There have been limited studies conducted to identify QTLs controlling seed dormancy in Arabidopsis (van der Schaar et al. 1997, Alonso-Blanco et al. 2003, Clerkx et al. 2004, see Chapter 1 for details). In comparison, QTL studies in rice have been more extensive. The number of rice seed dormancy QTLs and their genomic locations are varied across different studies. This is reflective of the wide variation in the degree of seed dormancy of rice cultivars (Sheshu & Sorrells 1986) and the influence of environmental factors. All 12 rice chromosomes have been implicated in controlling seed dormancy, with more than 30 QTLs reported using populations derived from cultivated rice (Wan et al. 1997, Lin et al. 1998, Miura et al. 2002, Dong et al. 2003), wild rice (Cai & Morishima 2000) and weedy rice (Gu et al. 2004). The identification of QTLs in the two model species with genome sequencing completed provides an ideal opportunity to identify functionally associated genes for seed dormancy.

The aim of this chapter was to establish a bioinformatic strategy to identify candidate genes for seed dormancy in the model species Arabidopsis and rice. The hypothesis underpinning this bioinformatic strategy is that genes controlling dormancy in Arabidopsis will have rice orthologs associated with dormancy in rice. The

bioinformatic strategy provides a framework to extend a comparative genomics approach to identify functionally associated genes for cereal dormancy.

2.2 Materials and methods

2.2.1 General outline of bioinformatic strategy

The bioinformatic strategy used in this chapter to identify candidate genes contributing to seed dormancy control in Arabidopsis and rice is outlined in Figure 2.1. Briefly, Arabidopsis seed dormancy QTLs and seed dormancy genes were identified from the literature. The published map locations of the seed dormancy genes and associated markers for QTLs were integrated in a physical map. Arabidopsis seed dormancy genes that were located within a seed dormancy QTL were selected as candidate genes for further analyses. Protein BLAST (Basic Local Alignment Search Tool) searches were used to identify rice sequences with similarity to the Arabidopsis candidate genes. Rice seed dormancy QTLs identified from previously published studies were retrieved from Gramene (http://www.gramene.org/) and integrated on the physical rice map. The positions of the rice orthologs identified in the Arabidopsis-rice comparison were correlated to the rice seed dormancy QTL locations. Rice orthologs that were positioned within a rice seed dormancy QTL were selected for further analysis. Thus, the criteria for a candidate gene was defined as an Arabidopsis gene that was located within an Arabidopsis seed dormancy QTL, with a rice ortholog co-locating to a rice seed dormancy QTL.

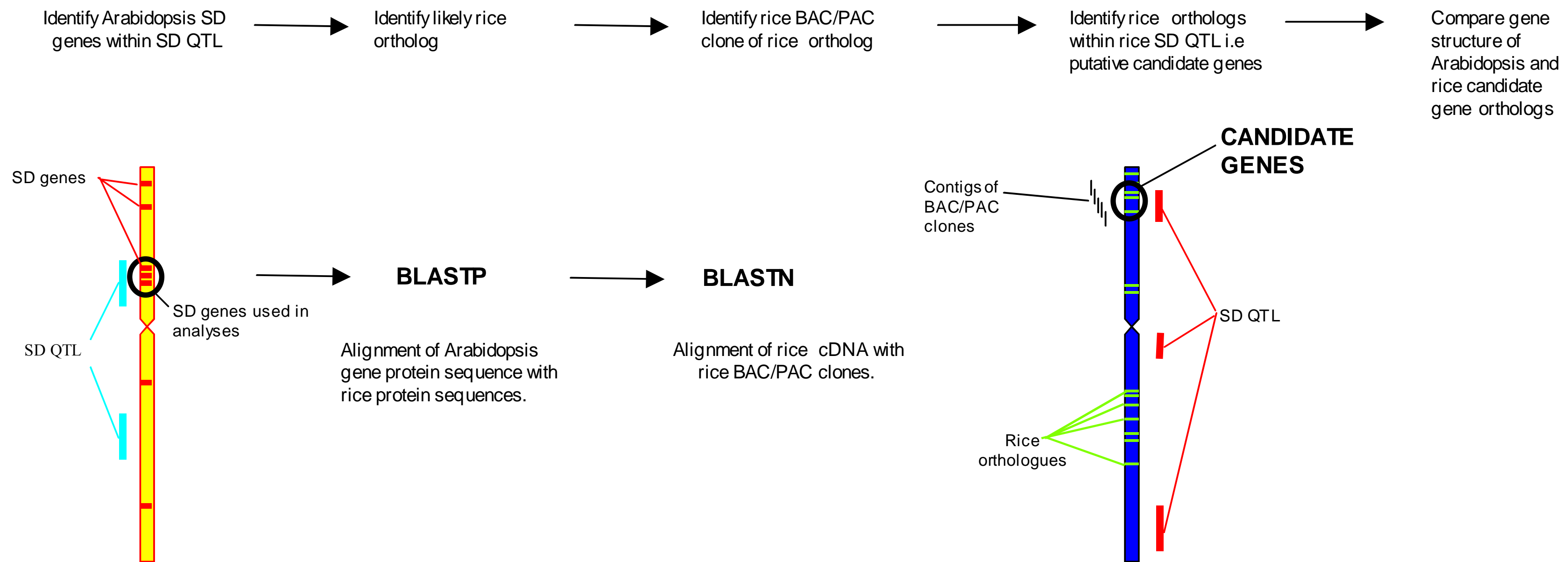


Figure 2.1: Schematic diagram of bioinformatic strategy used to identify candidate genes for seed dormancy (SD) using comparative genomics between Arabidopsis and rice.

2.2.2 Mapping of Arabidopsis seed dormancy QTLs

Markers linked to Arabidopsis seed dormancy QTLs were identified from the two studies by van der Schaar et al. (1997) and Alonso-blanco et al. (2003), who referred to the QTLs as DOG (Delay Of Germination) QTLs. The physical locations of the markers associated with the QTLs in the two studies, were identified on the Arabidopsis Genome Initiative (AGI) (http://genome.www.standford.edu/Arabidopsis/agi.html) sequence map available from the TAIR database (http://www.Arabidopsis.org/), enabling the mapping of Arabidopsis seed dormancy QTLs. Seed dormancy QTLs reported by van der Schaar et al. (1997) identified in all the environments tested were used. In some cases, flanking markers for a seed dormancy QTL had not been previously mapped on the AGI sequence map. A genetic map of these genes and markers was constructed based on the map coordinates from the Lister & Dean RI genetic map, also available from the TAIR website. Common markers between the physical and genetic map were used to bridge the two maps in order to establish the approximate location of QTLs on the physical map. Distances between common markers on the genetic map and the physical map were used to determine the estimated position of the QTL on the physical map. QTLs defined by only one marker were extended to include an additional distance of 500,000 bp (approximately 0.5 cM), on either side of the linked marker to incorporate ambiguities in QTL coverage.

2.2.3 Identification of Arabidopsis seed dormancy genes within QTLs

Arabidopsis genes with a known function associated with the control of seed dormancy and germination were identified from a survey of the current literature and "keyword" searches of the TAIR database, detailed in Appendix A. This study targeted genes functionally characterised by mutagenesis that were involved in plant hormone pathways that affected seed germination and/or dormancy, whereby a mutation resulted

in a defined response in germination and/or dormancy. Keyword searches were based on 'dormancy', 'germination', and the plant hormones abscisic acid, gibberellic acid, auxins, cytokinins, ethylene and brassinosteroids. The position of these genes on the AGI sequence map were determined and compared to the seed dormancy QTL locations detailed in section 2.2.2. Genes whose physical location co-located with a QTL were used for further analyses.

2.2.4 Identification of rice orthologs

Protein sequences of Arabidopsis genes co-locating with seed dormancy QTLs were obtained from the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov) database. The Arabidopsis protein sequences were aligned to rice protein sequences using a BLASTP search. Rice matches were selected for alignments that had at least 45% sequence identity for more than half the length of the Arabidopsis sequence, with an e value of $<e^{-50}$. Several genes were included as controls for the stringency of BLAST searches. These included the Arabidopsis *GAI* gene with orthologs in rice (*GAI*), wheat (*Rht*), maize (*Dwf8*) and barley (*SLN*) and the Arabidopsis *ABI3* gene with known orthologs *VIVIPAROUS-1*/*Vp1* in rice, maize and wheat. Rice cDNA sequences for the rice protein matches were obtained from NCBI and translated to ensure the rice cDNA translated to the protein sequence. The rice cDNA sequences were then used for BLASTN homology searches of rice genomic sequences. Rice BAC/PAC clones with greater than 95% nucleotide similarity for the entire length of the cDNA were identified on each rice chromosome. Rice cDNA-BAC alignment positions were examined for different cDNA accessions that aligned to the same BAC in order to elucidate whether the cDNA sequences were aligning to the same region on the BAC.

2.2.5 Mapping of rice seed dormancy QTLs and identification of co-location of rice orthologs with QTL positions

Rice seed dormancy QTLs from previously published studies were mapped on the Gramene TIGR Pseudomolecule Assembly of IRGSP Sequence 2005 (GR TIGR Assm IRGSP Seq 2005), a physical map containing contigs of rice BAC/PAC clones. Datasets of rice genetic maps from populations containing mapped seed dormancy QTLs were downloaded from the Gramene website. The populations and associated phenotyping data included F2 lines from a Nipponbare/Kasalath cross, F6/F7 RIL from a Asominori/IR24 cross, F6 recombinant inbred lines from a Pei-kuh/W1944 cross and backcross 1 lines from a EM93-1//EM93-1/SS18-2 cross. A consensus map of all the seed dormancy QTLs was generated by retrieving all markers that were located in a QTL for each map and combining them according to colinearity of marker order. The flanking markers for QTLs were located on the rice physical sequence map (TIGR Psudomolecule Assembly map). The seed dormancy QTLs were then integrated on the physical rice map using the common QTL markers between the maps. Putative rice orthologs identified from section 2.2.4 and co-located within a seed dormancy QTL were selected as candidate genes.

2.2.6 Confirmation of Arabidopsis candidate genes and putative rice orthologs by gene stucture similarity.

The intron-exon structures of the rice candidate genes were determined by aligning the rice BACs with corresponding full length cDNA sequences. In order to ensure that full length cDNA sequences were obtained, cDNA accessions were used in BLAST searches of the KOME database, containing full length cDNA sequences. Sequences retrieved from KOME were then subjected to translation using the ExPASy (Expert

Protein Analysis System) proteomic translation tool (http://au.expasy.org/tools/dna.html) and compared to the original protein sequence to ensure the full length cDNA translated to the entirety of the original protein accession. The rice gene structures were then compared to the intron-exon structure of the corresponding Arabidopsis gene provided by the TAIR database. Arabidopsis and rice genes were confirmed to be orthologous when intron-exon structures were similar. Amino acid residues at each splicing junction were also compared between rice and Arabidopsis genes (Appendix B) in order to identify orthologous exons.

2.2.7 Bin location of wheat orthologs

BLAST searches were conducted to identify mapped wheat ESTs with high similarity to the rice cDNA sequences of the four candidate genes, *ABI2*, *ERA1*, *HYL1* and *ETR1*. Both BLASTN and TBLASTX searches of the GrainGenes database containing binned wheat ESTs (http://wheat.pw.usda.gov/GG2/index.shtml) were performed. Sequence matches were considered to be significant when nucleotide similarity was at least 60% (e values < e^{-40}) or when amino acid identity was at least 80% (e values < e^{-80}) for at least half the length of the wheat EST sequence. When no significant EST matches were identified for the rice cDNA sequence, the search was extended to use the BAC clones of the rice cDNA sequences in the BLAST searches.

2.3 Results

2.3.1 Mapping of Arabidopsis seed dormancy QTLs

The seed dormancy QTLs from van der Schaar (1997) and Alonso-Blanco's (2003) studies were assigned to the Arabidopsis physical map based on the location of flanking markers. Figure 2.2 shows the genetic (Lister & Dean) and physical (AGI) Arabidopsis maps constructed with the published seed dormancy genes and QTLs and the common

markers between the two maps. Map locations of QTL1 and QTL3 on chromosome 1, QTL8 and QTL10 on chromosome 3 and QTL12 and QTL13 on chromosome 5, described by van der Schaar (1997) on the physical AGI map were estimated based on common markers between the genetic and physical maps, as flanking markers of the QTLs were not identified on the physical map (Figure 2.2). The only QTL that overlapped between the two studies was QTL1 from van der Schaar's (1997) study and the QTL termed DOG2 from Alonso-Blanco et al.'s (2003) study on chromosome 1, identifying this as a strong QTL for seed dormancy in Arabidopsis.

2.3.2 Identification of Arabidopsis seed dormancy genes and their location in relation to seed dormancy QTLs

Literature and keyword searches of the TAIR database for Arabidopsis seed dormancy genes resulted in the identification of 58 genes, functionally shown to be involved in various aspects of seed dormancy and/or germination (Appendix A). Of the 58 genes identified with the search, 56 had published map locations on the physical AGI map.

Mapping of the Arabidopsis seed dormancy QTLs with genes related to the control of seed dormancy and germination identified a total of 19 Arabidopsis genes that co-located with known seed dormancy QTLs (Figure 2.2). These included *ATRAB28*, *AXR3*, *HYL1*, *GAI*, *LEC2* and *AXR1* on chromosome 1; *ABH1* on chromosome 2; *DAG1* and *MARD1* on chromosome 3; and *ERA1* and *ABI2* on chromosome 5 co-located within an Arabidopsis seed dormancy QTL. An additional eight genes (*GA4*, *ABA3*, *LEC1*, *ETR1*, *RGL1*, *GA2*, *EBF1* and *ABI3*) were incorporated as candidate genes because they were located within 500,000 bp of a QTL defined by only one marker.

2.3.3 Identification of rice orthologs

The protein sequences for the 19 Arabidopsis genes that co-located to a seed dormancy QTL were used in BLASTP searches for rice orthologs. Rice orthologs, with greater that 45% similarity for more than half the length of the aligned Arabidopsis gene, were identified for 13 of the 19 genes (Table 2.1). The BLASTP search identified the published rice orthologs of the controls, *GAI* and the *ABI3* (*Vp1*) genes, indicating that the stringency level used was appropriate and the genes were no longer used in additional analyses (Table 2.1). There were no matches for the *ATRAB28*, *LEC1*, *LEC2*, *DAG1*, *MARD1*, and *GA4* genes at this level of stringency. These genes were therefore removed from further analyses. The remaining 13 candidate genes all had two or more significant alignments to rice protein sequences (shown in Table 2.1). Rice cDNA sequences for the 13 rice protein sequences (excluding the *GAI*, *ATRAB28*, *GA4*, *DAG1, MARD1, LEC1* and *LEC2* genes) were retrieved from NCBI and used in BLASTN searches to identify BAC sequences.

Single BAC sequences with greater than 95% identity across the entire length of the rice cDNA sequence were identified for each of the putative rice orthologs. The positions of the rice cDNA accessions that aligned to the same BAC (e.g. all the rice *GA2* cDNA accessions aligned to the BAC AL731610) were examined to elucidate whether they aligned to the same region on the BAC. No tandem duplications were observed (i.e all the rice cDNA accessions aligning to the same BAC, aligned to the same region on the BAC, and this was thoroughly checked). A summary of the BLAST searches is provided in Table 2.1. *GA2, ERA1, ABH1, ABI3, AXR1* and *EBF1* orthologs each aligned to one position on a single BACs indicating that they were single genes in rice. The cDNA of *AXR3, HYL1, ETR1* and *RGL1* significantly aligned to two or more BACs

and mapped to multiple positions on the same or different pseudomolecule, evidence of gene duplication in rice. The rice *ABI2* orthologs aligned to a BAC (AP003332) on chromosome 1 and to two separate BACs (AC130728 and AC097112) on chromosome 5. However, the two cDNA-BAC alignments on chromosome 5 did not match for the entire length of the cDNA sequences. The cDNA sequence encoding *ABI2* ortholog accession NP_001056498 did not align to the BAC clone, AC130728, in two regions of 55 and 86 bp in length. Similarly, there was a mismatch of 65 bp in the alignment between the cDNA sequence encoding the accession AAT39223 and the BAC accession AC097112.

Table 2.1: Summary of the BLASTP analysis in Arabidopsis (At) and rice (Os) identifying putative rice orthologs of the Arabidopsis candidate genes for seed dormancy. Table shows the rice BAC/PACs for the rice protein accessions and the e values for the Arabidopsis-rice protein sequence alignments. Percent identity (pid) and high scoring pair length (HSP length) are calculated from combined high scoring pairs of a single hit. Rice accessions that align to the same BAC all align to the same position on the BAC. BLAST results for the controls are presented at the top of the table (highlighted with an asterix) – *ABA3* and *GAI* Arabidopsis and rice orthologs are well established in the literature and serve to ensure the BLAST stringencies were discriminating.

At candidate	At accession	Os protein accession	e value	pid	At candidate length	HSP length	Os BAC/PAC clone	Os chr
*ABA3	NP_564001.1	ABH88164	0	56.8	819	831	AP003635	6
		Q655R6	0	51.3	819	832		
*GAI	CAA75492.1	AAX07462	1.00E-170	56.0	532	593	AC087797	3
		NP_001051032	1.00E-169	55.6	532	593	AC087797	3
		AAR31213	1.00E-103	52.6	532	371	AP003234, AP002910	1, 1
		AAY34500	3.00E-91	66.1	532	257	AC087797	3
		AAY34501	3.00E-91	66.1	532	257	AC087797	3
		AAY34509	3.00E-91	66.1	532	257	AC087797	3
		AAY34499	6.00E-91	66.1	532	257	AC087797	3
		AAY34502	3.00E-91	66.1	532	257	AC087797	3
		AAY34505	6.00E-91	66.1	532	257	AC087797	3
		AAY34497	3.00E-91	66.1	532	257	AC087797	3
GA2	AAC39443.1	AAQ72559	0	47.0	785	749	AL731610	4
		BAE72099	0	46.9	785	749	AL731610	4
		AAQ72560	0	46.7	785	749	AL731610	4
		AAQ72561	1.00E-180	46.0	785	750	AL731610	4
ERA1	AAA87585.1	BAD87023	1.00E-116	51.7	404	402	AP003218	1
ABI2	NP_200515.1	NP_001043415	1.00E-111	63.0	423	333	AP003332	1
		NP_001056498.1	3.00E-111	62.0	423	336	AC130728	5
		AAT39223	7.00E-87	51.6	423	343	AC097112	5
ABH1	NP_565356.1	AAG54079	0	57.1	848	855	AC134231	3
AXR3	NP_171921.1	NP_001067209	2.00E-59	47.8	229	278	AL837528	12
		Q75GK0	2.00E-55	50.8	229	242	AC145379	3
HYL1	NP_563850.1	AAY21792	1.00E-55	60.6	419	198	BX000493, BX000510, BX000492	11
		BAB00641	3.00E-52	59.3	419	199	BX000493, BX000510, BX000492	12
ABI3	CAA05484.1	BAA04066	2.00E-53	45.3	715	316	AP003436	1
AXR1	AAB59348.1	NP_001051717	0	68.0	540	522	AC104433	3
		AAO65877	0	67.4	540	522	AC104433	3
ETR1	AAA70047.1	AAB72193	0	69.8	738	613	AC091670	3
		AAL66363	0	63.0	738	621	AC087551	5
		AAV32219	0	67.0	738	512	AC087551	5
RGL1	AAL05911.1	AAX07462	1.00E-142	50.8	511	589	AC087797	3
		NP_001051032	1.00E-139	63.5	511	394	AC087797	3
		AAR31213	6.00E-91	49.3	511	371	AP003234, AP002910	1,1
		AAY34500	9.00E-84	61.3	511	256	AC087797	3
		AAY34501	9.00E-84	61.3	511	256	AC087797	3
		AAY34509	9.00E-84	61.3	511	256	AC087797	3
		AAY34499	9.00E-84	61.3	511	256	AC087797	3
		AAY34502	9.00E-84	61.3	511	256	AC087797	3
		AAY34505	9.00E-84	61.3	511	256	AC087797	3
		AAY34497	9.00E-84	61.3	511	256	AC087797	3
EBF1	AAL60026.1	BAD15849	1.00E-151	45.24207	628	599	AP004869	2
		NP_001046222	1.00E-123	46.08696	628	460	AP004869	2

2.3.4 Identification of rice orthologs within rice seed dormancy QTLs

Published rice seed dormancy QTL maps for the different rice populations were integrated on a genetic map. Common markers between the genetic map and the physical rice TIGR assembly map enabled 27 rice seed dormancy QTLs to be mapped on the physical map. The 13 putative rice orthologs identified on rice BACs (*ERA1, GA2, ABI2, ABA3, ABH1, AXR3, HYL1, ABI3, AXR1, ETR1, RGL1* and *EBF1*) were located on the rice physical TIGR assembly QTL map generated in section 2.2.5. Alignment of the rice genetic map with the physical map identified three rice gene orthologs that co-located with rice QTLs for seed dormancy, including *ABI2, ETR1* and *HYL1* (Figure 2.3). These genes co-located to seed dormancy QTLs in both Arabidopsis and rice fulfilling the criteria set for the identification of candidate genes. The rice *ERA1* ortholog aligned to a BAC positioned 200,268 bp outside a rice QTL (1 BAC interval). The close proximity of this gene ortholog to the QTL increased the likelihood that it was linked to the QTL, especially when discrepancies in marker locations that define the QTL when transferring from a genetic map to a physical map are taken into account. Furthermore, one of the boundaries of the QTL (QTL13) was not clearly defined due to a paucity of markers in the region of the *ERA1* gene. For these reasons, the *ERA1* gene was included as a candidate gene. *ABI2* and *ERA1* both mapped to rice chromosome 1; *ABI2* to QTL 2 and *ERA1* to QTL 4. *ETR1* mapped to the QTL 13 in rice chromosome 5 and *HYL1* was located on QTL30 on chromosome 11 (Figure 2.3). Further analyses on gene structure were performed on the four candidate genes *ABI2, ETR1, HYL1* and *ERA1*. The remaining rice orthologs that did not co-locate with a rice seed dormancy QTL were not further investigated.

2.3.5 Comparison of Arabidopsis-rice candidate gene structure

Full length cDNAs of the four candidate genes were aligned with their genomic counterparts in rice and Arabidopsis. Comparison of gene structure showed similar number of exons for *ABI2* and *ERA1* in rice and Arabidopsis (Fig 2.4). Variation in the *ERA1* genes size between Arabidopsis and rice (3393 bp compared to 3758 bp respectively) was mostly due to differences in intron size. Exon sizes were generally well conserved and differed by up to 7 base pairs with the exception of exon 1 and exon 12 which were 38 and 50 bp larger in rice respectively. The *ABI2* gene was substantially larger in rice compared to the Arabidopsis gene. This was mostly due to the third intron that was 4000 bp larger in rice than in Arabidopsis. In comparison, greater variation in gene structure was observed between the Arabidopsis *HYL1* and *ETR1* genes and their putative rice orthologs. The Arabidopsis *HYL1* gene has three exons while the gene structure of the two rice orthologs were made up of four exons. Only exon 2 was the same size between the Arabidopsis and rice *HYL1* genes. Examination of the splicing junctions revealed similarities in amino acid residues only for exon 2 and the 3' end of the third exon (Appendix B). The two *HYL1* rice orthologs were highly comparable to one another in both exon and intron size and base pair sizes differed only for the 5' terminal exons. Similarly to *HYL1*, the exon number differed for Arabidopsis *ETR1* gene and the identified rice orthologs. The Arabidopsis *ETR1* gene contains 6 exons while the two rice orthologs, cDNA accessions AK067813 and AK111696, contained 5 exons. The first four exons of the Arabidopsis *ETR1* gene and the two putative rice orthologs were highly similar in base pair sizes. The intron-exon junctions of the Arabidopsis *ERA1* gene and the rice orthologs also showed similar amino acids for the first four exons and the 3' end of exon 5. There were no amino acid similarities between the Arabidopsis and rice genes at the 5' terminal end of the genes.

2.3.6 Candidate gene alignment to binned wheat ESTs

BLASTN and TBLASTX searches were carried out to identify mapped wheat ESTs that had significant homology to the rice candidate gene orthologs. The results of the BLAST searches are shown in Table 2.2. BLASTN and TBLASTX searches of the cDNA sequences of rice *ABI2* orthologs consistently identified wheat EST accession BE445496, located on the long arms of wheat group 3 chromosomes. BLASTN and TBLASX searches using the rice cDNA sequence for the *ERA1* gene did not produce any significant matches to binned wheat ESTs. The BAC clone to which the *ERA1* gene aligned to in rice was then used in the BLAST searches. The BAC of the *ERA1* rice ortholog also had a significant alignment to an EST on the long arm of wheat chromosomes 3. The cDNA sequences of the rice *ETR1* ortholog had significant similarity to a wheat EST located on the long arm of chromosome 5B. Both BLASTN and TBLASTX searches of mapped wheat ESTs failed to result in significant alignments to the cDNA or BAC sequences of the rice *HYL1* ortholog.

Table 2.2: Results of BLASTN and TBLASTX alignments between rice candidate gene cDNA or BAC clones and bin mapped wheat ESTs. Table shows the chromosome location of the wheat EST hits and the wheat library that they belong to. Details of BLAST search are given, including e value, percent identity (pid) and high scoring pair length (HSP). The pid and hsp length are calculated from combined high scoring pairs of a single hit.

Candidate Gene	Rice accession	Wheat EST	Wheat Chr	Wheat Library	Wheat EST Length	BLASTN			TBLASTX		
						HSP e value	HSP pid	Total HSP length	e value	pid	Total HSP length
ABI2	AK242616 (cDNA)	BE445496	3ABDL	Wheat etiolated seedling root normalized cDNA library	585	E-144	87%	531	E-113	97%	551
	AK067627 (cDNA)	BE445496	3ABDL	Wheat etiolated seedling root normalized cDNA library	585	6.00E-49	83%	400	8.00E-81	81%	419
ERA1	AP003218 (BAC)	BE636979	3ABDL	Secale cereale anther cDNA library	357	3.00E-41	86%	330	no significant hits		
ETR1	AF013979 (cDNA)	BG605399	5BL	Wheat pre-anthesis spike cDNA library	563	5.00E-40	81%	301	6.00E-74	96%	347
	AK111696 (cDNA)	BG605399	5BL	Wheat pre-anthesis spike cDNA library	563	no significant hits			5.00E-55	85%	338

2.4 Discussion

Seed dormancy control involves multiple genes that are part of a network of biochemical, metabolic and physiological pathways. The increase in genome resources is facilitating the identification of genes contributing to complex traits such as seed dormancy. Computational methodology can be used to rapidly identify orthologs for genes that have been functionally characterised, for more discerning work. The bioinformatic strategy developed in the current study exploits the availability of gene sequences in the model species Arabidopsis and rice to identify candidate genes involved in seed dormancy control. The large number (58) of Arabidopsis genes identified to be contributing to seed dormancy control from the literature and keyword searches emphasises the complexity of the trait. This study relied on the publicly available genome sequence data of both species as of June 2004 and additional sequences and genome annotation may have revealed further information on other likely genes in either species. This study focussed on seed dormancy genes co-locating to published seed dormancy QTL regions in both Arabidopsis and rice, as these are potentially genes associated with natural variation for seed dormancy. Four genes fulfilled this criteria and were characterised as putative candidate genes with a function in seed dormancy control in cereals.

The four candidate genes that met the criteria of the bioinformatic strategy, *ERA1*, *ABI2, HYL1* and *ETR1*, have a range of roles involved in the hormonal response of Arabidopsis. *ERA1* and *ABI2* are important components of the Arabidopsis ABA signal transduction pathway, encoding a beta-subunit of protein farnesyl transferase (Culter et al. 1996) and a protein type 2C phosphatase (Leung et al. 1997) respectively. Genes involved in ABA signalling, response, biosynthesis and degradation pathways are key targets as the phytohormone has an established role in the maintenance of seed

dormancy. *ETR1* is involved in the early signal transduction pathway for the perception of ethylene (Hua et al. 1995, 1998, Bleecker et al. 1988, Sakai et al. 1998) while *HYL1* is associated with the hormone mediation of auxin, ABA and cytokinin (Lu & Fedoroff 2000). Hormones have long been known to be involved in the mediation of signals a plant receives from its environment (Trewavas & Malho 1997, Grill & Himmelback 1998, Solan & Ecker 1998, D'Agostino & Kieber 1999). Studies in the past have also alluded to a range of proteins involved in receiving, relaying and responding to external cues that commence the cascade of plant responses for development. These include receptor proteins, transcription factors and proteins such as protein kinases or farnesyltransferases that modify protein structure through phosphorylation or protein binding (Mulligan et al. 1997, Becraft 1998, Bonetta & McCourt 1998, Hooley 1998, Bleecker 1999, Thornton et al. 1999, Urao et al. 2000). Seed dormancy is a trait that is strongly influenced by the environment and would involve many of the types of proteins encoded by the candidate genes identified in this study to relay environmental cues indicating conditions are suitable for dormancy release and subsequent germination.

Orthologous regions have been identified in different species based on a high level of conservation in gene content and order (colinearity) at a macro level (macrosynteny) (Appels et al. 2004). Colinearity between genomes has primarily been observed between closely related species as it is correlated to evolutionary distance. In particular, members of the grass family, namely maize, rice, sorghum, barley and wheat demonstrate high overall conservation between genomes (reviewed in Gale & Devos 1998, Bennetzen 2000, Devos & Gale 2000, Keller & Feuillet 2000, Feuillet & Keller 2002). However, further characterisation of the colinearity at a sequence level

(microsynteny) has revealed macrosyntenic regions have inter-dispersed breaks in synteny from gene rearrangements such as deletions/insertions, duplications, translocations and inversions (Schmidt 2002, Bennetzen & Ma 2003). Colinearity is reduced between more diverse species due to a greater number of duplication events and gene rearrangements occurring during the time of their divergence. Not surprisingly, the genomic organizations of rice and Arabidopsis appear to be substantially different (Guyot & Keller 2004) with scarce collinearity in gene order (Laurie & Devos 2002, Liu et al. 2001, Devos et al. 1999, van Dodeweerd et al. 1999). Therefore, comparative mapping, between Arabidopsis and rice may not be useful in identifying orthologous genes. For this reason, a colinear study of syntenic regions across the Arabidopsis and rice genomes was not chosen. Instead, the strategy used was based on protein sequence similarity between Arabidopsis and rice and incorporated the co-location of seed dormancy QTLs to relate the annotated gene function to the trait of interest.

This study identified gene orthologs in Arabidopsis and rice on gene structure similarity and significant protein homology based on a criteria of minimum percent identities and e values. The stringency of the criteria set (>45% sequence identity for more than half the length of the Arabidopsis sequence, with an e value of $<e^{-50}$) and verification of a high degree of homology in terms of gene structure, provided a high level of confidence in the identification of true orthologs. Gene structures of two of the five candidate genes (*ABI2* and *ERA1*) identified in this study were largely comparable with similar sized and numbers of exons, further supporting them as true orthologs and indicating that they have been conserved since the evolution of a common ancestral lineage. Despite the evolutionary time separating the dicot Arabidopsis and monocot rice, comparative

analysis of their genomes have identified orthologous genes demonstrating well preserved gene structures and high sequence similarities (Liu et al. 2001). Genes that govern fundamental cellular process in plant development and growth may remain highly conserved across the monocot and dicot divide. Cereal homologues of the Arabidopsis gibberellic acid insensitive gene (*GAI*) were isolated in maize (*Dwarf 8*) and wheat (*Rht*) (Peng et al. 1999), demonstrating that orthologous genes can be conserved both in structure and in function in distantly related species. Another more recent example is the characterisation of the *MAX* genes in Arabidopsis and the identification of their orthologs in pea, petunia and rice (Zou et al. 2006). However, gene structure homology does not mean that the gene orthologs perform the same function. Gene function needs to be experimentally defined for reliable functional classification.

The candidate genes *HYL1*, *ETR1* and *ABI2* had more copies in rice than in Arabidopsis, indicating duplication events had occurred in these genes following the divergence of Arabidopsis and rice. The *ABI2* gene orthologs were identified on rice chromosomes 1 and 5. Duplicated segments between chromosome 1 and 5 have been previously reported (Kishimoto et al. 1994, Kurata et al. 1994a,). Figures accounting for the extent of rice genome duplication have varied from 15% (Vandepoele et al. 2003), 20% (Simillion et al. 2002), 45% (Wang et al. 2005), and 61.9% (Paterson et al. 2004). As a major food source, rice is grown in many regions of the globe and subsequently has been adapted to grow in a wide range of geographic habitats. The plant has undergone decades of continuous breeding and selection since the beginning of agriculture, thus increasing opportunities for genome expansion, contraction and rearrangements. Gene duplication enhances genetic diversity and provides an important source of novel genes

to enable plants to change and adapt to changing environmental conditions (Crow & Wagner 2006, Lynch & Conery 2000).

In addition to having multiple gene copies, the gene structure of the *HYL1* and *ETR1* rice genes differed from their Arabiopsis orthologs in terms of exon numbers. Changes in the gene structure of these genes, may be indicative of duplication events followed by divergence. While some gene copies are inactive due to functional redundancy, others may evolve a specialized function that is beneficial for plant survival in new environments through chance mutations (Blanc & Wolfe 2004, Francki & Appels 2007). For example, the *X2* gene encodes a transcription factor and has been sequenced in rice, maize and sorghum (Bennetzen & Ma 2003). Maize and sorghum orthologs are missing the first exon of the *X2* gene containing a zinc-finger domain involved in the protein-protein interaction (Bennetzen & Ma 2003). The authors suggest the lack of the first exon would enable the rice gene to interact with a greater number of protein partners compared to the maize and sorghum genes. Further studies would be required to determine whether the rice *ETR1* and *HYL1* orthologs are active genes, and whether changes in their structure since the divergence from Arabidopsis has had any effect on their function. The lack of amino acid similarity at the splice junctions of several exons when comparing the Arabidopsis and rice genes, particularly for *HYL1*, suggest that these genes are not true orthologs of the Arabidopsis *ETR1* and *HYL1* genes, at least not in function.

Mutations of duplicated genes generally leads to loss of function of the gene copy, except in some instances, where the mutation proves to be advantageous and is retained (Massingham et al. 2001). Thus, many gene copies are non functional or may vary in

their active role making it difficult to identify copies that have the ancestral function. BLAST searches at best may identify an ortholog or gene family member or even an unrelated gene with a comparable sequence domain (Devos & Gale 2000). Gene expression and regulation is additionally influenced by a loci's chromosomal location (Sasaki & Sederoff 2003), which can have biological implications. Thus, it is important to gather evidence correlating the gene function to the phenotype of interest. Investigations carried out in the current study, integrated phenotypic data (QTL data) from published seed dormancy studies for a targeted approach in identifying candidate genes with a high probability of having a conserved function in cereals. Locating candidate genes within seed dormancy QTLs provides further support for these genes to have a function related to the regulation or maintenance of seed dormancy. Genes such as *ABI2* and *ERA1* that have retained gene structure over the dicot-monocot divide and additionally map within seed dormancy QTLs are ideal candidate genes for further study.

In comparison to the divergence of monocots and dicots, speciation between cereals occurred more recently (approximately 50 Myr in the Poaceae family) (Paterson et al. 2004). Rice can therefore act as a bridging species to cereals and orthologs are expected to be more easily identifiable in other important cereal species such as wheat and barley. Comparative mapping of common molecular markers have demonstrated high levels of macrosynteny exists between rice and wheat. The BAC and cDNA sequences of *ERA1* and *ABI2* rice orthologs significantly aligned to mapped wheat ESTs in the long arm of the group 3 chromosomes. Comparative sequence analysis undertaken by La Rota & Sorrells (2004), have indicated that the rice chromosome 1, on which both the *ABI2* and *ERA1* rice orthologs are located, is closely related to wheat

group 3 chromosomes, across the whole chromosome. Recent mapping of the *ERA1* gene in a population of diploid wheat recombinant inbred lines derived from a cross between *Triticum monococcum* and *T. boeoticum* located the wheat ortholog (*TmERA1*) on wheat chromosome 3A (Nakamura et al. 2007). The authors additionally detected two minor seed dormancy QTLs on the chromosome 3AL. However, the QTL did not co-locate with the map location of *TmERA1*. Instead the map locations of *T. monococcum* orthologs of the *ABI8* and *ABI3*, other Arabidopsis ABA signalling genes, were found to correspond to these QTLs.

Seed dormancy QTLs have been previously detected on the long arms of the group 3 wheat chromosomes (Zanetti et al. 2000, Groos et al. 2002, Osa et al. 2003, Kulwal et al. 2004, Nakamura et al. 2007). Flanking microsatellite markers and markers located within the QTLs on the wheat group three chromosomes have been mapped to the same wheat deletion bin as the wheat ESTs aligning to the rice *ABI2* and *ERA1* gene orthologs (see Table 2.3 below). Therefore, these two ABA signalling genes co-locate to seed dormancy QTL regions in Arabidopsis, rice and wheat, providing substantial support to investigate these genes further as candidate genes. In order to confirm the predicted gene ortholog locations, genetic mapping should be undertaken using a population that segregates for the trait of interest.

Table 2.3: Wheat deletion bins containing wheat ESTs that align to rice orthologs of candidate genes ABI2 and ERA1. Table also shows wheat microsatellite markers that have been deletion bin mapped (Sourdille et al. 2004) and are associated with seed dormancy QTLs in wheat (Kulwal et al. 2004, Groos et al. 2002).

Candidate Gene	Wheat EST	Wheat Bin	Wheat QTL Markers	Reference
ABI2	BE445496	C-3AL3-0.42	Xgwm480, Xgwm666	Groos et al. 2002
		C-3BL2-0.22	Xfbb283, Xfba242, Xgwm131, Xcdo583, XATPase, Xgwm108, Xbcd1418, Xmwg818	Kulwal et al. 2004
			Xfbb110, Xgwm131	Groos et al. 2002
		C3DL2-0.27	Xcdo1406	Kulwal et al. 2004
ERA1	BE636979	3AL3-0.78-1.00	Xcdo482	Groos et al. 2002
		3BL2-0.22-0.50	Xfbb378	Kulwal et al. 2004
		3DL2-0.27-0.81	Xwg110, Xbcd288	Kulwal et al. 2004
			Xgwm383, Xgwm3	Groos et al. 2002

The outcome of the bioinformatic strategy demonstrated Arabidopsis seed dormancy genes had orthologous rice genes that colocated with seed dormancy QTLs, verifying the hypothesis tested. The genes *ERA1* and *ABI2* are ideal candidates for further study, given their function in the ABA signal transduction pathway and involvement in the early onset of dormancy in Arabidopsis. The fact that the structure of these two genes remains highly conserved despite large evolutionary distances between Arabidopsis and rice indicate that these genes are true orthologs. Their co-location to seed dormancy QTLs in both Arabidopsis and rice provide evidence that their function may also be conserved. Furthermore, the position of their wheat orthologs in the wheat group 3 chromosomes for which seed dormancy QTLs have been identified additionally support their potential as important candidate genes involved in seed dormancy in cereals. Genetic mapping of these genes in a mapping population segregating for seed dormancy would be necessary to establish whether these genes map to locations associated with the control of seed dormancy. The map locations of these genes within a seed dormancy QTL would justify further functional studies which remain pivotal to proving gene function is linked to the trait of interest.

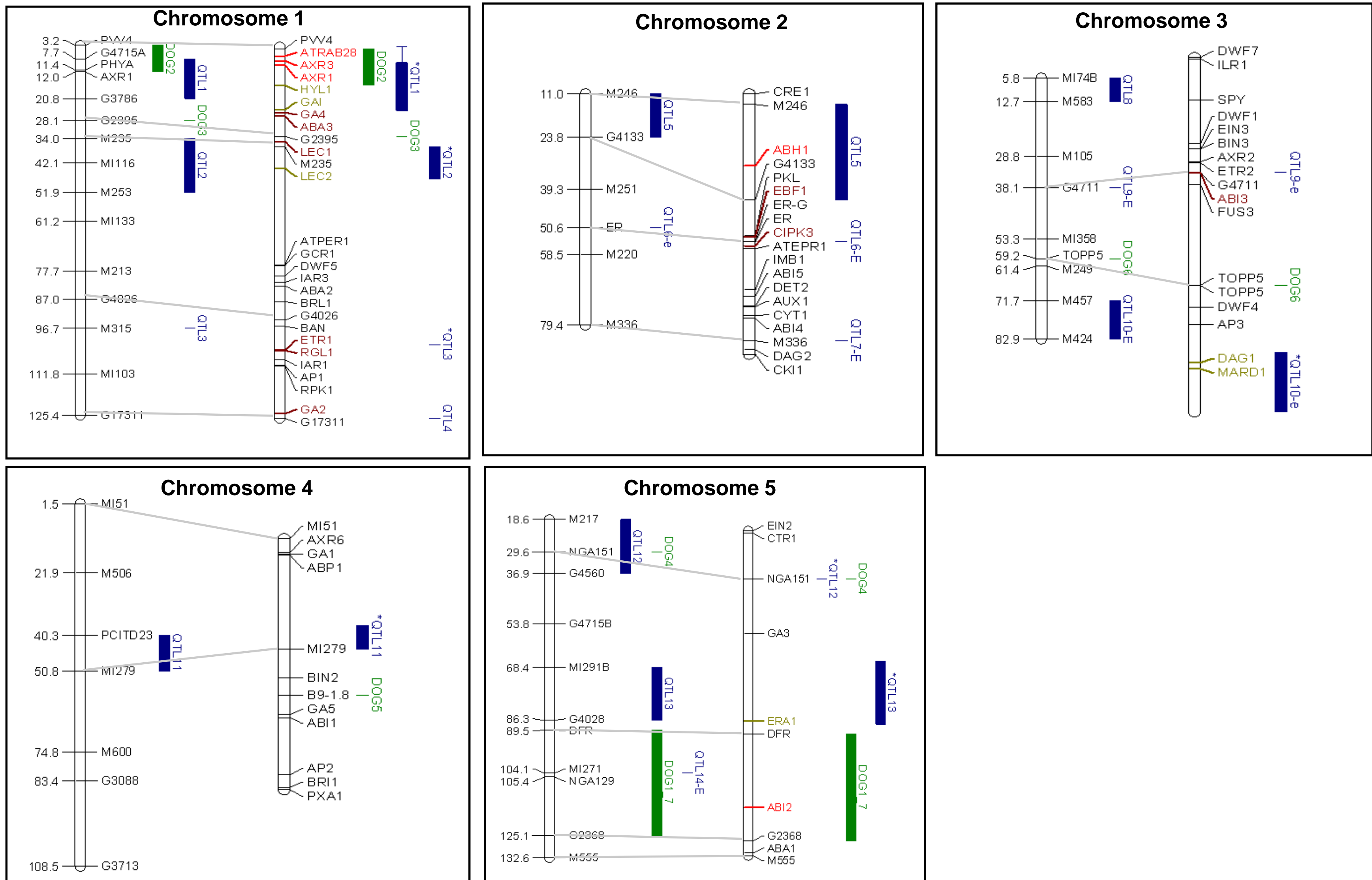


Figure 2.2: Arabidopsis genetic (left) and physical (right) maps showing published seed dormancy genes and QTLs and common reference markers between the maps (indicated by grey lines between the maps). Published seed dormancy QTLs are defined by green vertical lines (DOG QTLs from Alonso-Blanco et al, 2003) and blue vertical lines (van der Schaar, 1997). Seed dormancy genes are highlighted in red, khaki and brown. Genes in red are confirmed to locate within a seed dormancy QTL on the Arabidopsis physical map, khaki genes are those identified to locate within a seed dormancy QTL based on distances between common markers of the two maps. These QTLs are also marked with an asterix. Genes highlighted in brown are additional genes incorporated as candidate genes due to their close proximity, (within 500,000bp) to a seed dormancy QTL defined by only one marker.

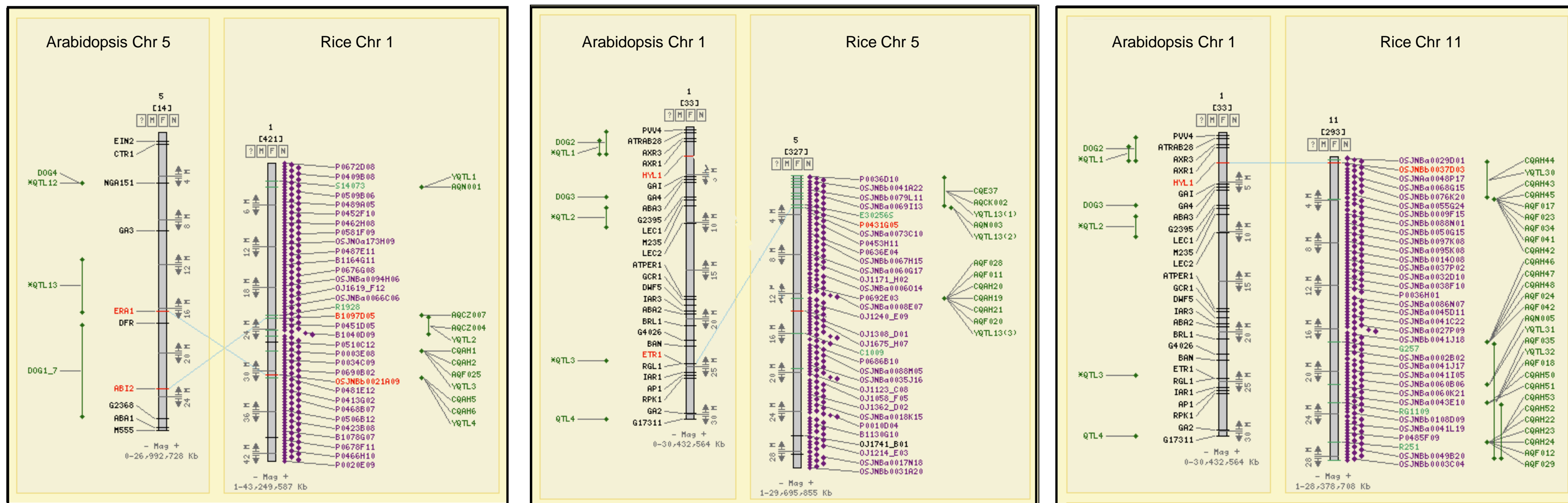


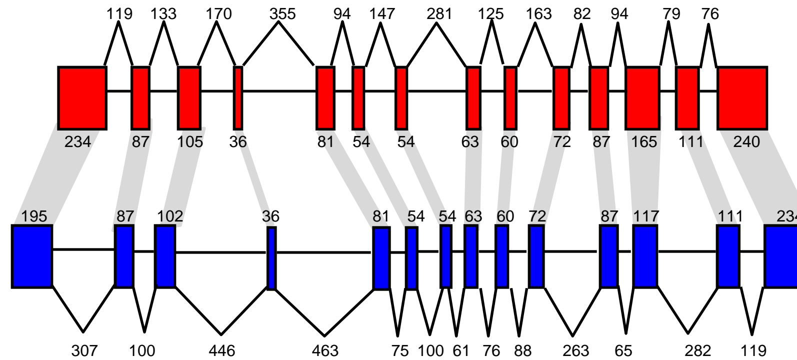
Figure 2.3: Arabidopsis physical map with published seed dormancy genes and QTLs with correspondences to rice physical map (Gramene Tigr Assembly map, pseudomolecule version 3) generated in CMAP (http://ccg.murdoch.edu.au). Rice physical map shows location of published rice seed dormancy QTL. Figure shows the location of the genes *ERA1, ABI2, HYL1* and *ETR1* within an Arabidopsis QTL as well as the location of their rice orthologs. These rice orthologs also locate to a rice seed dormancy QTL.

ERA1

Gene annotation: Encodes a protein phosphatase 2C (PP2C) and is involved in abscisic acid signal transduction

Arabidopsis (3393 bp)
cDNA : AF214106

Rice (3758 bp)
cDNA : AK073299
BAC - AP003218



ABI2

Gene annotation: Encodes a protein phosphatase 2C (PP2C) and is involved in abscisic acid signal transduction

Arabidopsis (1598 bp)
cDNA : Y08965

Rice (5895 bp)
cDNA : AK242616
BAC : AP003332

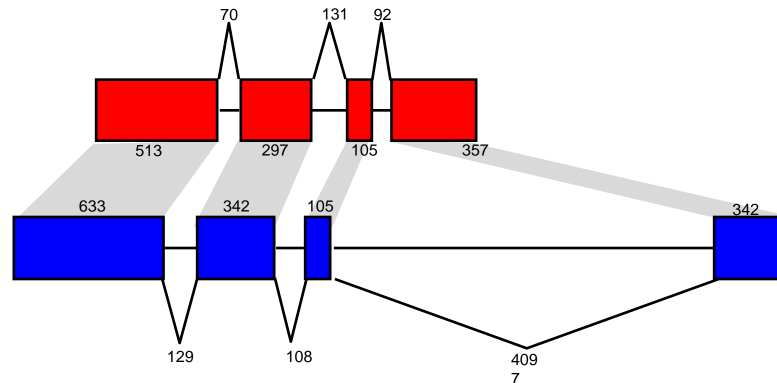


Figure 2.4: Intron-exon structure of Arabidopsis candidate genes (red) and putative rice orthologs (blue). Translated regions are represented by boxes, and introns by horizontal lines with the lengths of exons and introns displayed as number of base pairs. Grey shading indicates regions between the orthologs that are significantly aligned. GenBank accession numbers are given for the cDNA and BAC/PAC clones.

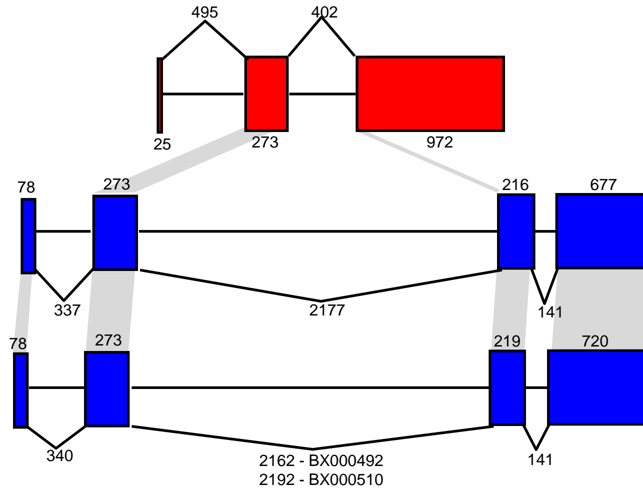
HYL1

Gene annotation: encodes a dsRNA binding protein. Among many other things, mutants exhibit less sensitivity to auxin and cytokinin

Arabidopsis (2167 bp)
cDNA : AF276440

Rice (3899 bp)
cDNA : DQ009988
BAC : BX000493

Rice (3792-3822bp)
cDNA : AB036988
BAC : BX000492/BX000510



ETR1

Gene annotation: encodes a histidine kinase and a response regulator domain. Ethylene receptor. Mutations affect ethylene, binding and metabolism of other plant hormones.

Arabidopsis (2664 bp)
cDNA : L24119

Rice (3239 bp)
cDNA : AK067813
BAC : AC091670

Rice (2651 bp)
cDNA : AK111696
BAC : AC087551

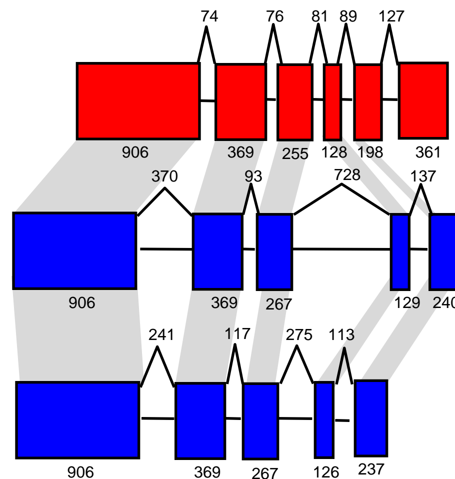


Figure 2.4 continued

CHAPTER 3: SEED DORMANCY IN A STIRLING/HARRINGTON CROSS: GENERATION OF A MOLECULAR GENETIC MAP AND QTL ANALYSES

3.1 Introduction

Preharvest sprouting is a world-wide problem for many barley cultivars, particularly malting barley. Pre-harvest sprouting often results from prolonged wet and damp conditions at grain maturity whereby the grain germinates while still attached to the mother plant prior to harvesting (Benech-Arnold 2002). The detrimental effects caused by preharvest sprouting result in substantial losses in grain quality and as a consequence severely limit the use of the grain in malting (Castor & Frederiken 1977, Bason et al. 1991).

Seed dormancy is an important trait for many cereal species because it affects resistance to pre-harvest sprouting. Seed that lack seed dormancy are susceptible to pre-harvest sprouting, however, strong seed dormancy in barley is undesirable as rapid germination upon imbibition is required for malting of the grain (Oberthur et al. 1995). The challenge is to develop barley cultivars that produce grain with sufficient levels of dormancy to withstand preharvest sprouting but allow efficient malting post harvest.

Seed dormancy is a difficult trait to dissect due to its polygenic nature and because it is strongly influenced by environmental conditions and naturally demonstrates large genetic variation (van Der Schaar et al. 1997). These compounding factors make it difficult for breeders to reliably phenotype and predict preharvest sprouting resistant lines. Seed germination percentage has been the predominant method of measuring seed dormancy as the quantitative trait cannot be directly measured (Zhang et al. 2005).

A number of QTL studies have highlighted chromosomal regions thought to control seed dormancy in barley. Major or minor QTLs identified in previous studies have implicated every barley chromosome in the control of seed dormancy, an indication of the trait's complexity. In particular, QTLs on chromosomes 2H (Thomas et al. 1996, Li et al. 2003a, Prada et al. 2004) and 5H (Ullrich et al. 1993, Oberthur et al. 1995, Han et al. 1996, Takeda 1996, Thomas et al. 1996, Li et al. 2003a, Prada et al. 2004) have been detected in several studies involving different environments.

Breeding programs in Australia have produced barley cultivars with relatively good levels of preharvest sprouting tolerance, including vars. Stirling, Gairdner, Franklin, Schooner and Sloop (C. Li, Department of Agriculture and Food Western Australia pers comm.). Stirling has been a dominant malting cultivar used in Australia since its release in 1982 and currently is still widely sown in Western Australia. The cultivar is popular due to the large plump grains and high level of diastase enzymes it consistently produces. It is widely adapted to both medium and low rainfall areas and more importantly, has strong resistance to preharvest sprouting. Harrington barley, originally from Canada, has been widely used as a parent within Australian breeding programs due to the high quality of malt it produces. Compared to Stirling, Harrington is superior in terms of malting characteristics. However, Harrington barley is virtually non-dormant and is therefore highly susceptible to preharvest sprouting. Using Stirling and Harrington as parents, a new barley cultivar Hamelin, was released in 2002 (Garlinge 2005). The cultivar Hamelin combines good agronomic traits from Stirling and high malting quality from Harrington. It is also susceptible to preharvest sprouting like Harrington.

The aim of this study was to characterise the features of dormancy in progenies from a Stirling/Harrington cross. The location and individual effects of QTLs associated with seed dormancy control was investigated and the epistatic interactions occurring between the QTLs were examined.

3.2 Materials and methods

3.2.1 Plant material

A population of 185 double haploid lines (DHLs) generated by anther culture from a Stirling/Harrington cross was kindly provided by Kirin Australia. The DHLs were used to construct a barley molecular genetic map and for QTL analyses.

3.2.2 DNA extraction

The DH population of the Stirling/Harrington cross and the parental lines were grown in a glasshouse, in trays using a standard potting mix. Phenol chloroform extraction of DNA using standard protocols adapted from Rogowsky (1991) was performed using the leaves of 2 week old seedlings for each of the 185 lines of the DH population and pure lines of Stirling and Harrington.

3.2.3 Molecular marker analysis

A genetic linkage map was generated using SSR markers covering the 7 chromosomes. Locus specific SSR primers obtained from the GrainGenes website (http://wheat.pw.usda.gov) with generic non-complementary nucleotide sequences at their 5' end were provided by the Molecular Plant Breeding CRC (MPBCRC) (Hayden et al. 2008). The generic forward and reverse nucleotide sequences at the 5' end were 5' ACGACGTTGTAAAA 3' and 5' CATTAAGTTCCCATTA 3', respectively. Tag primers, tagF and tagR with sequences 5' ACGACGTTGTAAAA 3' and 5'

CATTAAGTTCCCATTA 3’, respectively were also provided by the MPBCRC (Hayden et al. 2008). TagF primers were dye-labeled at the 5’ end with either a VIC, FAM, NED or PET fluorescent dye (Applied Biosystems). Polymerase chain reaction (PCR) was performed in a total volume of 6 µl containing 50 ng of template DNA, 0.2 mM dNTP, 1x ImmoBuffer (Bioline) (16 mM (NH4)2SO2, 0.1% Tween-20, 100 mM Tris-Hcl, pH8.3), 1.5 mM MgCl2, 75 nmol each of dye-labled tagF and unlabelled tagR primer, 0.15 U Immolase DNA polymerase (Bioline) and locus-specific primers at an optimum concentration determined by Hayden et al.. 2008. A total of 744 markers were screened for polymorphism on the parental lines Stirling and Harrington. Of these 211 markers were deemed to be polymorphic and were screened in the population for the genetic map construction.

Markers that varied in allele size were multiplexed and the marker sets amplified together. PCR reactions consisted of initial denaturing at 95°C for 10 min, followed by 5 cycles of denaturation at 92°C for 30 s, annealing at 50°C for 90 s and extension at 72°C for 30 s. This was then followed by 20 cycles of denaturation at 92°C for 30 s, annealing at 63°C for 90 s and extension at 72°C for 60 s, followed by another 40 cycles of denaturation at 92°C for 15 s, annealing at 54°C for 30 s and extension at 72°C for 30 s. A final elongation was carried out at 72°C for 10 minutes before a soak at 15 C. PCR products in each set of markers labelled with a different dye were mixed together in a ratio of 3:4:2:4 for VIC:FAM:NED:PET in order to account for fluorescent differences of each fluorophore. The mixed marker sets were transferred to an AcroPrep 384 filter plate (PALL Life Sciences cat#5077) and desalted by ultrafiltration. PCR products were then resuspended in sterile water to a final dilution factor of 1:100. Visualisation of PCR products was performed by capillary gel electrophoresis on an

ABI 3730 DNA capillary sequencer (Applied Biosystems). LIZ-labelled GeneScan500 (Applied Biosystems) was used as an internal size standard for capillary electrophoresis and GeneMapper v3.7 software (Applied Biosystems) was used for PCR fragment size analysis. Figure 3.1 shows analysis of marker screening as performed in GeneMapper.

Figure 3.1: View of marker screening performed in GeneMapper v3.7 demonstrating clear polymorphism in marker fragment size between DH lines of the Stirling/Harrington population.

3.2.4 Marker mapping and genetic map construction

Software package Map Manager QTX was used to determine initial linkage groups using the Kosambi map function (Kosambi 1944) and a search and linkage criterion of P<0.001. RECORD (Van Os et al. 2005) was used to optimise the marker order of linkage groups and the final map was then defined by Map Manager QTX. Chromosomal allocation of markers and marker order were compared to published

marker locations. Chi square analysis was performed on marker genotyping to ensure normal segregation of markers.

3.2.5 Phenotyping Trials

The Stirling/Harrington DH population was planted in 2005 at the Wongan Hills Research Station, Department of Agriculture and Food WA (DAFWA), 188 km northeast of Perth. The field plot was planted in a randomised complete block design with plots of 1 x 3 m^2. Control plots of Stirling (parental line) plants were sown in the first and last rows (rows 1 and 19) and Hamelin (progeny of Stirling x Harrington cross) control plots were sown in the middle of the field plot (row 11). The second parental line Harrington was planted in an individual plot in the next field. In 2005 the population was replanted in the Wongan Hills site as well as at the Mt Barker Research Station, Department of Agriculture located in the south west of Western Australia, 359 km south east of Perth. Field plots for the trials sown in 2006 were also planted in a randomised complete block design but with controls in a regular grid structure. The parental lines Stirling and Harrington and the cultivars Hamelin, Gairdner and Baudin (cultivars with known quality and performance traits) were used in the control. The grid structure of the controls ensured all DH line plots could be compared equally with each variety of control to account for spatial variation. Wongan Hills is a low rainfall site with a short growing season while Mt Barker is a high rainfall site with a long growing season. For both years and the two different sites, 20 intact heads from each line were harvested by hand at physiological maturity, indicated by the loss of all green colour from leaves and stems. These were placed in individual opened paper bags and allowed to air dry at room temperature for 7-10 days. Spikes were carefully hand threshed and cleaned using a rubber bicycle tube to prevent physical damage to the seed. Moisture content of the seeds was verified using near-infra red (NIR). The seed

with a moisture content of 11-13% (harvest ripeness) was stored at –20°C prior to seed germination tests to prevent any further physiological maturing. Seed germination tests were then carried out in batches until germination tests on all lines were performed.

3.2.6 Germination tests

Full grains that appeared to be healthy and free from any discolouration or damage were selected for germination tests. Germination tests followed the protocol of the Analytic EBC (European Brewery Convention 1987). The seed weights of two replicates of 100 seed for each DH line were measured. Germination tests on the seed of each replicate were carried out immediately after being removed from –20C storage. The seed was germinated on two Whatman No. 1 filter papers in a 9 cm petri dish. Four mL of deionised water was added to the petri dishes and the plates were incubated in the dark at 18 °C. Germinating seeds (visual sign of radicle emergence) were counted at 24 hour intervals for 72 hours. Seed dormancy was expressed as the percentage of the total germinating seed after 72 hours. Seeds that did not germinate within 3 days were considered to be dormant and defined as the seed dormancy phenotype. Germination tests were repeated using after-ripened seed – this was seed left at room temperature seven weeks after the initial germination test. Lines that demonstrated greater than 95% germination in the first germination test were not subjected to the second germination test as they were deemed to be essentially non-dormant. Germination tests on after-ripened seed were carried out in order to determine if QTLs for primary dormancy differed from after-ripening dormancy.

3.2.7 Data analysis

Germination percentage was analysed independently for the two germination tests performed at different after-ripening stages: seed germinated immediately after harvest, and the seed germinated 7 weeks post harvest. Analyses demonstrated no significant differences in germination or seed weight of replicates for all treatments. No outliers were detected and the correlation between replicates for germination and seed weight were 0.9925 and 0.93876 g respectively (data not shown here). Therefore replicate averages were used for all analyses hereafter.

Seed weight demonstrated a normal distribution while the distribution of seed germination was skewed. The means and variances of seed germination was calculated using a nonparametric method (Hozo et al. 2005) to account for the skewed distribution.

The controls for the Wongan Hills and Mt Barker sites in 2006 were used in analyses performed in GenStat to identify spatial variation in the field. Unfortunately, due to the design of the Wongan Hills 2005 trial, spatial analysis could not be performed for this site. For the two 2006 sites, REML variance analysis was carried out using a mixed model regression consisting of a random model to account for field variation by including the column and row terms as random terms and a fixed model using seed weight, germination date and cultivar as fixed terms. A probability level of 0.05 was used to determine significance. The REML analysis produces Best Linear Unbiased Predictor (BLUP) values. The analysis can cause a small amount of shrinkage in the predicted values depending on the size of the dataset, giving values of greater than 100%. In the current study, shrinkage effects were minimal due to the large size of the trials. Predictions slightly higher than 100% were substituted with a value of 100%.

Predicted values accounting for the factors that were statistically significantly affecting germination were then used in the QTL analyses that followed.

3.2.8 QTL mapping

QTL analyses were performed using the software package QTLNetwork (Yang et al. 2008); in this software the QTL effects are estimated by the Monte Carlo Markov Chain method/mixed linear model approach. The QTLNetwork also allows the data to be analysed for QTLs with additive effects (A), epistatic interactions (additive by additive, AA) and environmental effects (additive by environment effects AE and epistasis by environment effects, AAE). Permutation tests (Doerge & Churchill 1996) were carried out using 1000 iterations at 1 cM intervals. A minimum separation of 10 cM ('filtration window') was used to define individual adjacent QTLs. The QTLNetwork calculates a P value for significance value and in the present study a threshold of $P<0.05$ was used to declare a significant QTL.

3.3 Results

3.3.1 Trial climate

Average daily temperatures varied slightly for the two different locations (Wongan Hills and Mt Barker) and for the two trial years (2005, 2006). Wongan Hills in 2006 experienced warmer temperatures by 3-4°C in the last 3 months before harvest compared to the same trial site the year before. Daily temperatures (averaged for the month) were comparable between Wongan Hills 2005 and Mt Barker 2006, differing only by 1-2°C for the months in which the trials were planted. The two years at Wongan Hills received a similar total amount of rainfall over the 5 months (32.44 and 33.8 mL respectively), however, the rainfall trend in the 5 months differed. Rainfall at the Wongan Hills 2005 trial decreased over the last 3 months and in the final month

only 7.4 mL of rain was received. In 2006, the rainfall for the last 3 months was relatively consistent and high, with an average monthly rainfall of 43.1 mL. Mt Barker being in the higher rainfall zone experienced a higher monthly rainfall average of 43.44 mL for the 5 months of the trial. The rainfall also decreased in the last two months at Mt Barker. Prior to harvest, conditions were very dry with less than 1 mL of rainfall in Wongan Hills 2005 and Mt Barker 2006. Similarly for Wongan Hills 2006 there was no rainfall for 11 days before harvest, however, on the day of harvest there was 13.2 mL of rainfall (Figure 3.2), possibly inducing preharvest sprouting damage in some of the DH lines with low dormancy.

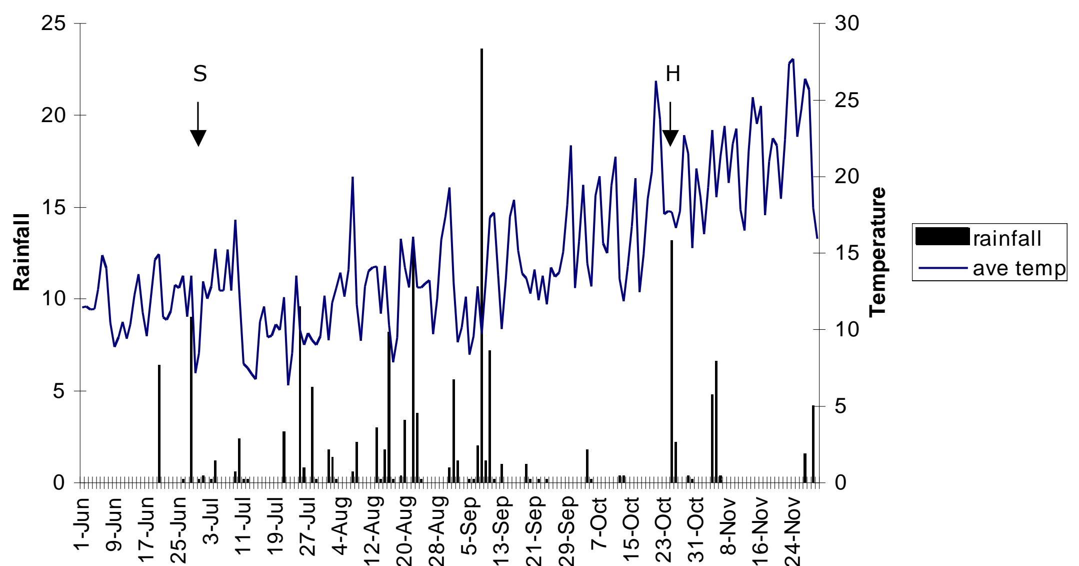


Figure 3.2: Daily rainfall and average daily temperature for the Wongan Hills 2006 site. Sowing date (S) and harvest date (H) are shown.

3.3.2 Seed germination

In the Wongan Hills 2005 and Mt Barker 2006 trials, Harrington seed from the control plots consistently had a higher germination percentage than Stirling within the 3 day period of the germination tests. All the Harrington seed germinated from 90-100% for both sites demonstrating no dormancy. Only 1% of Stirling seed from 2 out of the 10 control plots of the 2005 Wongan Hills site germinated, although the seed was viable, while the germination percentage of all the Stirling seed from the 2006 Mt Barker site ranged from 5-23%. Germination of Stirling and Harrington controls in the 2006 Wongan Hills site demonstrated relatively large variation possibly an indication of preharvest sprouting damage due to the large amount of rainfall just prior to harvest (Figure 3.2). Harrington germination percentages in all the control plots varied from 5-95% germination while germination of seed from all the Stirling control plots ranged from 18-68%.

The distribution of germination of the 2005 Wongan Hills seed suggested a single major gene controlled germination in this population with 89 DH lines (nearly half the population) having 0-10% germination (79 of these DH lines demonstrated no germination within the three day period of the germination test in one or both of the seed germination replicates) and 86 DH lines demonstrating 90-100% germination (Figure 3.3). In the two sites, Wongan Hills and Mount Barker, the following year, most DH lines germinated at 90-100% (112 and 97 DH lines respectively) while the rest of the DH lines varied relatively evenly from 0-90% germination (Figure 3.3). Four and six of the 185 DH lines showed no germination in the Wongan Hills 2006 and the Mt Barker 2006 trials, respectively.

Figure 3.3: Distribution of seed germination for germination tests carried out immediately after harvest in the Stirling/Harrington DH population from the two different sampling years and sites.

3.3.3 Statistical analysis

Mean seed weight of seed collected for the germination tests were comparable between the three sites, with MB2006 producing the highest seed weight mean of 4.95 g (per 100 seed) and WH2006 demonstrating the lowest seed weight mean of 4.15 g. The mean seed weight of WH2005 was 4.16 g. Variance appeared to increase with increasing seed weight. All three sites demonstrated a medium spread of variance (coefficient of variance ranging from 10.3 – 14.3%). Mean seed germination percentage was similar between WH2006 and MB2006 (74% and 71.5%, respectively). The mean seed germination was lower for WH2005 (57.75%). Variance decreased with lower germination and the spread of data was much higher for seed germination compared to seed weight. Statistical analysis demonstrated significant spatial trends in the two field trials of Wongan Hills 2006 and Mt Barker 2006 with t-test scores > 1.96 (10.55 and 10.64, respectively). Table 3.1 summarises the statistical data for seed weight and seed germination under the three environments tested.

Table 3.1: Mean, variance, standard deviation and coefficient of variation given for seed weight and seed germination under the three environments tested.

	*ENVIRONMENT		
	WH2005	**WH2006**	**MB2006**
SEED WEIGHT			
Mean (g/100 seed)	4.61	4.15	4.95
variance	0.26	0.18	0.5
st. dev	0.51	0.43	0.71
Coefficient of variance	11%	10.3%	14.3%
SEED GERMINATION			
Mean (%)	57.75	74	71.5
variance	219.4	220.75	220.54
st. dev	14.8	14.9	14.8
Coefficient of variance	25.6%	20.1%	20.8%

*Environments: Wongan Hills 2005 (WH2005), Wongan Hills 2006 (WH2006) and Mount Barker 2006 (MB2006).

3.3.4 Genetic map

The construction of the Stirling/Harrington genetic map shown in Figure 3.4 and described in Table 3.2 included 128 markers covering approximately 1145.3 cM with an average marker spacing of 8.9 cM. An additional 15 markers were not included due to severe segregation distortion, missing data or lack of linkage detection. A further 37 markers co-located with other more reliable markers and hence were removed from the map, in particular only 12 of the 28 markers on 5H had unique locations. This was mostly due to 13 markers being mapped to the same position as the marker HVM0030. There were gaps of more than 30-39 cM on each chromosome with the exception of 1H whose largest gap was 24.7 cM. Chromosome 5H had the largest gap from the centromeric region to the long arm which spanned 51 cM between GBM1399 and E6041XQ. Marker order and approximate map distances between markers were comparable between the current map and published consensus barley maps (Karakousis et al. 2003, Varshney et al. 2007).

Table 3.2: Number of mapped markers, total genetic distance and maximum distance (genetic) between marker intervals for each chromosome of the genetic map constructed from the DH population derived from the Stirling/Harrington cross (see Figure 3.3 for map diagram).

Chromosome	No. of marker positions	Length of linkage group	Maximum genetic distance between adjacent markers
1H	9	71.5	24.7
2H	24	248.8	39.2
3H	14	169.2	33.3
4H	33	223.6	33.1
5H	12	97.7	51
6H	24	185.1	31.5
7H	12	149.2	34.4

3.3.5 Detection of QTLs for seed dormancy

The QTLNetwork analysis indicated that phenotypic variation in seed dormancy was mostly influenced by genetic effects (data not shown). Genetic effects were 2.71 times more important than GxE interactions and 3.41 times more important than environmental effects. Random errors were relatively low compared to the above effects.

QTL analyses identified a major QTL and minor QTLs with significant additive effects on chromosomes 1H, 3H and 5H (Figures 3.4 and 3.5). Two QTLs were identified on chromosome 5H, one on the distal region of the long arm (5Hqb) and another near the centromere of 5H (5Hqa). The 5Hqb QTL was considered to be a major QTL as it contributed a high proportion of variation in seed dormancy, with an additive effect of 28.756 and accounting for 52.17% of the phenotypic variation (Table 3.3). Significant GxE interactions were observed for this locus in all the three environments tested essentially differentiating germination by 14.87%. Dormancy was consistently due to the Stirling allele at this locus. Significant GxE interactions were also demonstrated by the second locus on 5H (5Hqa) in the three environments (Table 3.3). This QTL displayed an additive effect of 7.658, with the Harrington allele generally increasing

germination with the exception of the WH2005 site. Although the heritability of this locus was relatively low (3.7%), its genetic effect on seed dormancy was greater than that of the 5Hqb locus in the MB2006 environment (AE = -5.23 compared to 2.59). Epistatic interactions detected between the two loci on 5H were also influenced by the environment (AAE interactions) (Table 3.3, Figure 3.3 - see inset). The epistatic interactions between the 5H loci had a negative effect on germination reducing germination by 7.66%.

Minor QTLs detected on chromosomes 1H (1Hq) and 3H (3Hq) were expressed only additively, with no environmental interaction (Table 3.3). Heritability of both these QTLs was small, 1.66% and 2.16% respectively and dormancy was conferred by the Harrington allele at both these loci (Table 3.3). 1Hq was located relatively close to the centromeric region while the QTL on 3H was positioned below the centromere on the long arm (according to correlation of markers in the consensus map). The 3Hqa QTL covered a relatively large genomic region of 27.3 cM. Resolution of this QTL was not achieved due to a lack of markers mapped to this region. Additive by additive (AA) interactions were observed between 1Hq and 3Hqa, with a small negative effect on germination of 3.92%. The QTL on 3H additionally displayed a negative epistatic main effect with the 5Hqa QTL (AA = -2.94%, Table 3) that was additionally influenced by the environment (AAE interactions) for the WH2005 site.

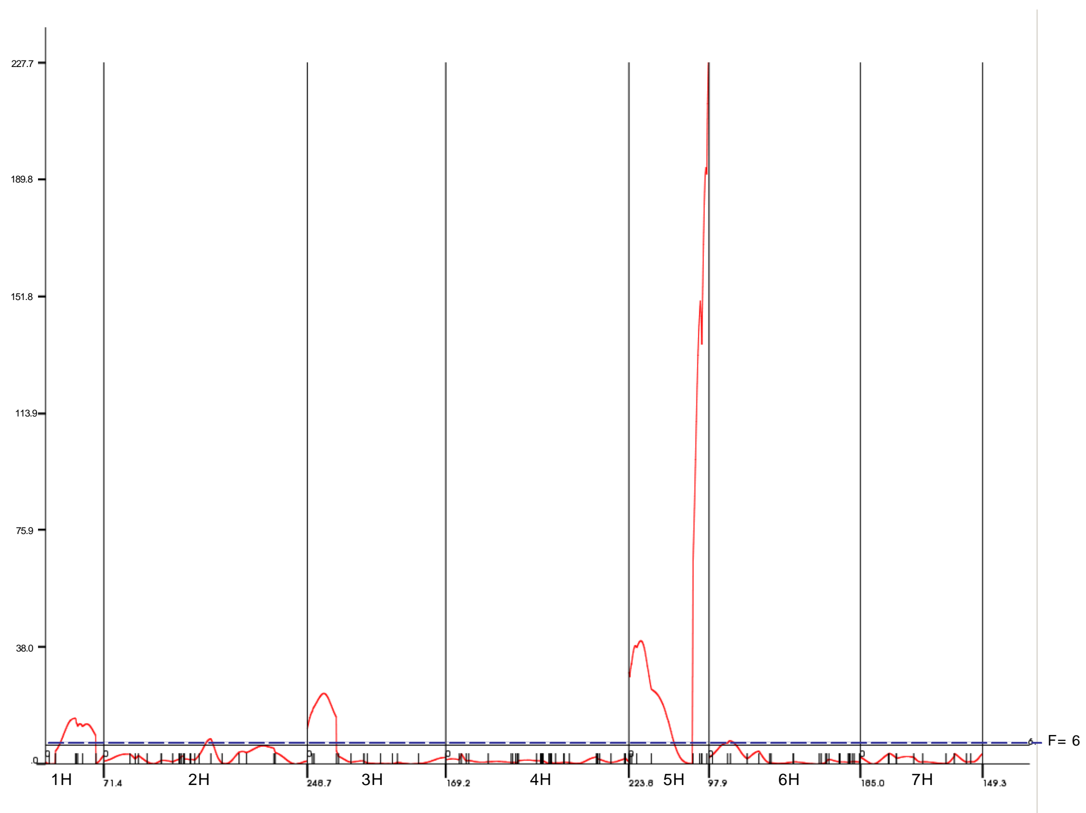


Figure 3.5: QTL-likelihood curves (in red) generated by QTLNetwork, showing the locations of QTLs on 1H, 3H and two QTLs on 5H for the first germination tests carried out immediately after harvest, taking into account all three environments, Wongan Hills 2005, Wongan Hills 2006, Mt Barker 2006. The horizontal black line indicates the QTL threshold equivalent to a critical F value of 6 (P<0.01).

Small and inconsistent effects were further expressed by some minor QTLs on 2H, 3H, and 4H detected only in epistatic interactions (described in Table 3.3 and shown in Figure 3.4). A QTL on the long arm of 2H demonstrated additive by additive interactions (AA = 1.328) with a locus situated on the distal end of the long arm of 3H (3Hqb, Table 3.3). The environment had no significant effect on the interactions between these loci. A QTL on the distal end of the 4H short arm demonstrated interactions with two loci on the long arm of the same chromosome (Table 3.3). One of these interactions exhibited an epistatic main effect (not environmentally influenced) while the other interaction was expressed in epistatic by environment interactions.

3.3.6 Seed dormancy QTL analysis on after-ripened seed

QTL analyses for the after-ripened seed germinated 7 weeks after harvest, was carried out separately for the Wongan Hills sites in both years (2005, 2006). Similar QTL locations were identified on 5H in the second germination test (using after-ripened seed) (Table 3.4). In the germination test using seed from the Wongan Hills 2005 planting, the same QTL was located on the long arm of 5H, however, the QTL on the short arm of 5H was identified one marker interval higher than in the first germination test. A QTL on 1H was also observed to be at the distal end of the short arm, 32.4 cM above the QTL located in the earlier analysis. Similar to the previous 1H QTL identified, the Harrington allele also conferred dormancy at this locus. Germination tests using after-ripened seed from Wongan Hills 2006 only detected the QTL on the long arm of 5H. The phenotypic variation that these QTLs accounted for was much lower than in the previous germination test (heritability of 9-12.6%), presumably because most of the seed had lost all dormancy. There were no QTLs detected for seed dormancy in the after-ripened seed of Mt Barker 2006. This was due to the lack of dormancy remaining in the seed after the 7 weeks, as germination percentages were greater than 95%.

Table 3.3: Description of significant ($P<0.05$) seed dormancy QTLs detected in the Stirling/Harrington DH population in three environments (two locations and two seasons). QTL heritability, additive main effects and epistatic and environmental interactions are shown.

	QTL	Chr	Position (cM)	Range (cM)	Marker interval	QTL effect	A	$h^2(A)$	GxE	Allele	Epistasis	Epistatic i	$h^2(AA)$	GxE
	1Hq	1H	37	28.3-38.6	Hvgluend-Awbms80	A	5.13	1.66%	-	H	1Hq-3Hq	I	0.97%	
*	2Hqa	2H	130.8	130.8-136.8	GBMS244-Bmag174				-	H	2Hqa-3Hqb	I	0.11%	
	3Hqa	3H	20.3	13.3-26.3	GBM1043-Bmag0013	A	5.85	2.16%	-	H	3Hqa-5Hqa	I	0.55%	
*	4Hqa	4H	0	0-7.0	GBM1501-Bmag741				-	H	4Hqa-4Hqb	I+IE	0.18%	+/-0.44%
											4Hqa-4Hqc	I	0.23%	
	5Hqa	5H	14.6	9.6-18.6	Bmag0337-GBM1399	AE	-7.66	3.7%	+/- 5.02	S	5Hqa-5Hqb	I+IE	3.71%	+/- 5.11%
	5Hqb	5H	97.1	96.1-97.1	scssr09041a-scssr03907	AE	-28.76	52.17%	+/- 14.87	S				

QTLs only detected in epistatic interactions are marked with an asterisk. QTL effects are defined by A (additive) or AE (additive by environment). GxE values indicate the increase or decrease in effect due to genotype x environment interaction in the three environments tested. Allele represents preferable allele at locus for seed dormancy. Epistatic interactions are defined by I (epistatic main effect only), IE (epistatic by environment interaction) and I+IE (epistatic main effect and epistatic by environment interactions).

Table 3.4: Description of seed dormancy QTLs detected in the individual three environments tested using after-ripened seed (seed stored for 7 weeks after harvest). Chromosome assignment, marker interval, additive effect, heritability, standard error and P value are given for each QTL. Parental allele for dormancy or germination is given for each locus.

Env	Var(G)	Var(r)	QTL	QTL loc'n	Chr	Marker interval	QTL effect	$h^2(A)$	P value	SE	G(germ)	G(dorm)
WH2006	81	111	1Hqc	1H-1	1H	GMS21-Bmag0321	A	12.6%	0.00	0.77	S	H
			5Hqa	5H-4	5H	EBmac0965b-Bmag0337	A	9%	0.00	0.79	H	S
			5Hqb	5H-11	5H	scssr09041a-scssr03907	A	11.7%	0.00	0.78	H	S
WH2007	0.08	0.78	5Hqb	5H-11	5H	scssr09041a-scssr03907	A	9%	0.00	0.0661	H	S
MB2007				-								

Env = Environment, Var(G) = genetic variance, Var(r) = residual variance, $h^2(A)$ – heritability, P value = significance value, G(ger) and G(dorm) refers to the parent allele contributing the germination or dormancy phenotype respectively.

3.3.7 High seed dormancy genotype

A genotype with high seed dormancy could be predicted by observing the various combinations of parental alleles at the different QTL locations and examining the effect on the phenotypic variation. For maximum seed dormancy, two alleles were required from the Stirling parent for the two QTLs on 5H and two alleles were required from the Harrington parent for the 1H and 3H QTLs. Desirable alleles at all the loci increased dormancy from 60.64 to 71.78% depending on the environment, while alleles promoting germination at each loci was predicted to increase germination by 23.5 – 54.8%. Statistical analyses of the quantitative data (Table 3.5) demonstrated the importance of allelic interactions between the minor QTLs in the control of seed dormancy. The minor QTLs at 1H and 3H played a consistent role in controlling seed dormancy. In the three trials, conducted over 2 consecutive years and in two different locations, the Harrington allele at both the 1H and 3H QTL locations resulted in consistently and significantly low seed germination.

Table 3.5: Table of means and standard variation for genotypic combinations at minor QTL locations on chromosomes 1H and 3H for germination percentage in the three environments tested. R^2 values are shown for P<0.0001 (***). Allele 'A' is the Stirling allele and allele 'B' is the Harrington allele. The first allele of the genotypes is representative of the allele at the 1H QTL and the second allele is representative of the allele at the 3H QTL e.g. AB = Stirling allele at the 1H QTL and Harrington allele at the 3H QTL. MB = Mount Barker, WH = Wongan Hills.

	2006MB		2006WH		2005WH	
Genotype	Mean	St Dev.	Mean	St. Dev	Mean	St. Dev
AA	72.5	4.2	82.7	3.2	55.7	5.5
AB	86.6	5.4	92.9	4.2	72.6	7.1
BA	90.6	5.0	96.1	3.9	73.9	6.7
BB	26.6	4.3	40.6	3.3	11.5	5.7
R^2	0.41***		0.47***		0.28***	

3.4 Discussion

Analysis of the Stirling/Harrington DH population in this study identified two highly significant QTLs on chromosome 5H (5Hqa and 5Hqb). In previous studies four regions of the barley genome on chromosomes 4H, 5H (two loci) and 7H have been found to demonstrate prominent gene effects associated with the control of seed dormancy. Han et al. (1996) termed these loci as SD1 to SD4. In the current study, the locations of the two QTLs on 5H coincide with the SD1 locus near the centromere of 5H (Ullrich et al. 1993, Oberthur et al. 1995, Han et al. 1996, Prada et al. 2004, Zhang et al. 2005) and the SD2 locus on the long-arm telomere region of 5H (Ullrich et al. 1993, Takeda 1996, Li et al. 2003a, Prada et al. 2004, Zhang et al. 2005), respectively. The widely conserved and repeatable expression of the two QTLs on 5H is an indication of their fundamental involvement in the control of seed dormancy. Romagosa et al. (1999) suggested that the SD1 gene was the most important QTL in seed dormancy release, supported by the large effect of the loci displayed (accounting for over 50% of the phenotypic variance) in several different environments (Oberthur et al. 1995, Han et al. 1996). Although QTLs relative to SD1 and SD2 were detected in this study, QTL 5Hqb (effectively SD2) was found to be the major gene, accounting for the largest overall additive effect and displaying the highest heritability.

Minor seed dormancy QTLs involved in epistatic interactions with other seed dormancy QTLs were identified on the terminal regions of the long arms of chromosomes 4H and 3H. Co-linearity between the barley and wheat genomes has been well established. Seed dormancy QTLs have been previously described for the long arms of the group 3 wheat chromosomes (Groos et al. 2002, Zanetti et al. 2000, Kulwal et al. 2004) and on the long arms of the wheat 4B and 4D chromosomes (Kato et al. 2001), in regions that

may be syntenic to the minor QTL locations in barley chromosomes 3H and 4H and the major QTL identified in the present study as 5Hqb (Li et al. 2004). The *viviparous-1* (*Vp1*) (described in the literature review of Chapter 1) gene in maize and its homologues in wild oat (*Avena fatua*) and wheat, as well as the closely related *ABI3* Arabidopsis gene, has been shown to be involved in the seed dormancy control of those respective plants (McCarty et al. 1991, Jones et al. 1997, Giraudat et al. 1992). The wheat *taVp1* homologues are located on the long arms of the group 3 chromosomes near the centromeres (Bailey et al. 1999) within the vicinity of the QTL for dormancy (Groos et al 2002). The *Vp1* gene has not yet been mapped in barley. Therefore, the possibility that the minor QTL identified on the long arm of 3H (3Hqa) could be associated with the *Vp1* gene cannot be discounted. Interestingly, the bioinformatic analysis in Chapter 2 did not identify a consistent association of *Vp1*/*ABI3* gene with dormancy QTLs. Further work is needed to verify whether the *Vp1* gene is linked to the QTLs identified in this study.

Seed weight and seed germination were weakly associated. Although seed weight was normally distributed, seed germination was skewed, making it difficult to interpret the correlation between the two variables. Grain weight may be a factor influencing seed germination but it is difficult to predict how it affects the trait due to environmental influence. For WH2005 and MB2006, larger seed demonstrated greater germination. However, WH2006 produced the smallest mean seed weight out of the three trials and had the greatest mean germination percentage. As the WH2006 trial experienced heavy rainfall just prior to harvest, this may have initiated germination cues in the seed, explaining the high germination percentage in relation to the smaller seed. There have been several studies that examined the relationship between seed size and germination

(as well as subsequent seedling establishment) in crop species (Mathur et al. 1982, Lafond & Baker 1986, Kawade et al. 1987, Roy et al. 1996, Guberac et al. 1998, Larsen & Andreasen 2004, Willenborg et al. 2005) with varying results depending on the crop species. However, the mechanism by which seed size affects germination is not yet completely understood. The physical properties of a seed (e.g. waxy surface layer, area of seed coat, crease depth) may affect the water potential of the seed, thus affecting imbibition and subsequently germination. Alternatively, seed size may be an indirect result of certain developmental processes involved in seed germination. What is clear is that the situation is complex and that there are many factors that may inadvertently affect dormancy and germination. Interestingly, QTL analysis of the seed weight data identified QTLs near the centromere region of 2H and the short arm of 3H, one marker interval (cM) higher than the minor seed dormancy QTL discussed previously. The same gene may, therefore, contribute the QTL on 3H for seed dormancy and seed weight. The seed dormancy QTL in the centromeric region of 2H has previously been reported by Li et al. (2003a). This genomic region has also been associated with cereal developmental processes. The basic vegetative period (BVP) (Major 1980) or earliness per se (Hoogendoorn 1985) relates to the photoperiod insensitive phase that immediately follows germination, during which vegetative growth occurs from sowing to floral initiation. Boyd et al. (2003) found that loci marked by Bmy2 on 2HC had a dominant influence on heading date and BVP. The authors concluded that the locus codes for a developmental variable. Pillen et al. (2003) examined the marker-trait associations of 13 agronomic traits in an advanced backcross population (BC2F2) generated by crossing var. Harry with ISR101-23. The locus marked by GMS3 located in the centromeric region on 2H was associated with heading date, plant height, number of kernels per ear and yield. Both seed germination and vegetative growth (including

seed development) would be stimulated by developmental signals. It is possible that the various pathways of cellular and metabolic processes involved in seed germination and vegetative growth may overlap, and thus certain genes may be common between these development processes.

The 5Hqa and 5Hqb loci were additionally expressed in the after-ripened seed. Genetic pathways are likely to be common between primary seed dormancy release and seed after-ripening and may involve genes represented by these QTLs. Analysis of the after-ripened seed also detected a QTL on the distal end of the short arm of 1H was sufficiently distant from the 1Hqa detected earlier, suggesting they are two separate QTLs. This QTL may therefore be specific to release of dormancy following after-ripening or a quiescent period. Although epistatic interactions were observed between the 5H loci in the initial germination test, no interactions were observed using the after-ripened seed. Instead, both loci interacted additively with the 1H QTL. Variation in these interactions may indicate the different pathways involved in primary dormancy release and after-ripening dormancy release even though several of the same genes may be involved.

The QTL on 5HL coincides with a region known to be associated with malting attributes (Mather et al. 1997, Marquez-Cedillo et al. 2000, Barr et al. 2003), compromising the ability to breed cultivars that have the allele for preharvest sprouting tolerance in this region without causing changes in the malting qualities (Li et al. 2003). Therefore the identification of other important alleles contributing to seed dormancy, particularly in the genetic background of a malting cultivar such as Stirling, would provide greater options in breeding for preharvest sprouting tolerance. The QTL

identified here accounted for 61.84 – 80.91% of the phenotypic variation encountered. There is likely to be additional QTLs compromised by environmental effects and genotype by environmental effects and therefore not detected in this study, particularly in regions lacking substantial marker coverage. In order to discover all the genes involved in the control of seed dormancy and possibly some that demonstrate high heritability in certain environments, QTL analysis of larger populations and greater marker density is required. Although the heritability of minor QTLs is often low (as seen in this study), their combined additive effects, and their epistatic interactions, may provide substantial main effects on overall seed dormancy performance.

All the QTLs detected in this study were found to be involved in epistatic interactions and the majority of the significant epistatic interactions had an inhibitory effect on germination. The largest of the epistatic effects was observed between the two loci on 5H. Previous studies have reported on interactions between the two QTLs on 5H (SD1 and SD2 appear to correspond to 5Hqa and 5Hqb in this study) (Oberthur et al. 1995, Han et al. 1996, Larson et al. 1996, Romagosa et al. 1999). In contrast, no epistatic interactions were detected between these QTLs in the after-ripened seed but both QTLs acted additively with the QTL on 1H.

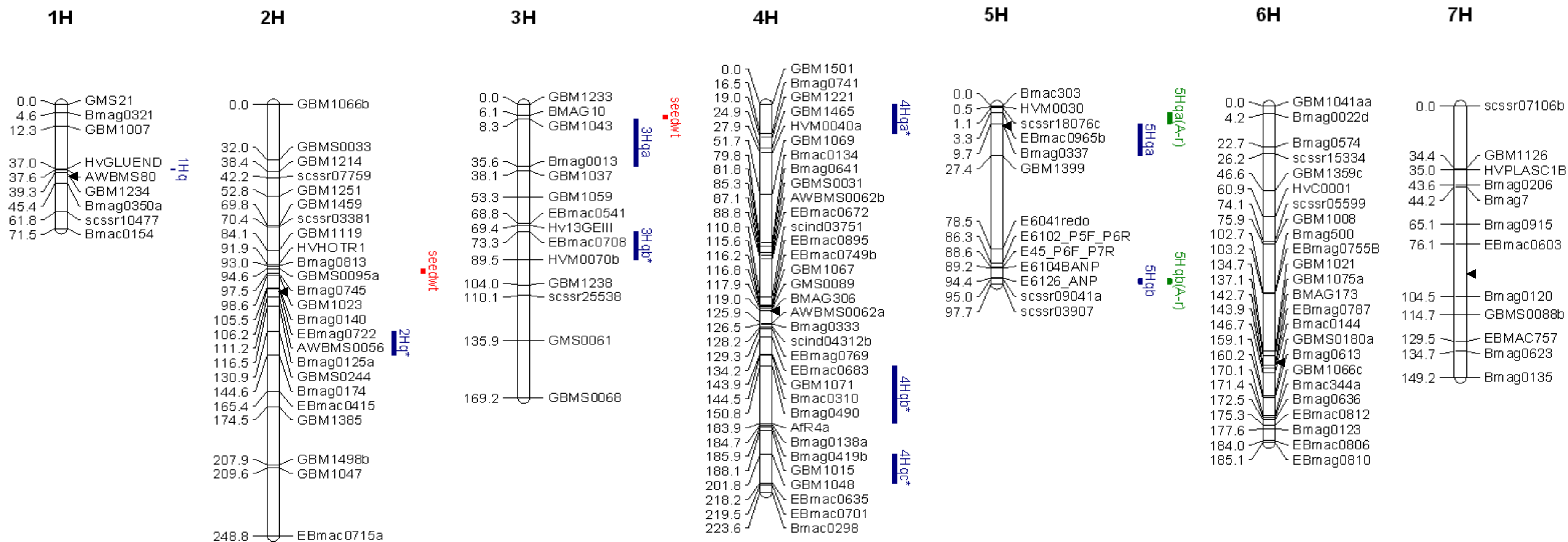
The range of genetic and environmental interactions observed in this study highlighted the complexity of the trait and emphasised the role that genetic by environment interactions play in determining the expression and effect of these loci. As seen in this study certain environments may enhance or reduce the expression of a gene and the additive effects of loci varied with the environment. The phenotypic variation accounted for by the identified QTLs ranged from 61.84 – 80.91% depending on the

environment. Some epistatic interactions were also only significant in certain environments. In addition, the environment also influenced whether the Stirling or the Harrington allele promoted or reduced germination for certain QTLs. For example, the presence of the Harrington allele at the 5Hqa increased germination for two out of the three environments tested. In the WH2005 site, the Stirling allele at this locus promoted germination although the parent Stirling is known to have strong dormancy. Similarly the additive by additive interactions between the two loci on 5H either reduced or enhanced germination depending on the environment. These two QTLs present a gene combination that produces variable effects in seed dormancy. Although the two loci were responsible for the majority of the phenotypic variation, the variability in their effect due to the environment would make it difficult for breeders to reliably use these loci in their routine breeding programs. Increased knowledge on the environmental conditions that would make these genes effective for seed dormancy is necessary and would need careful monitoring of the environment.

The current study demonstrated the significance of gene combinations in influencing a complex trait such as seed dormancy. By combining optimum alleles at QTL locations for dormancy a genotype with high levels of seed dormancy could be predicted, resulting in dramatic differences in dormancy and germination. In the environments tested, dormancy of this genotype could potentially be increased from 56.8 to 83.7% with specific allele combinations at the identified QTL locations. In the current study, the combination of Stirling alleles at the two 5H QTL locations and Harrington alleles at the 1H and 3H QTL locations significantly produced the greatest dormancy. While the behaviour of major genes contributing dormancy may be highly susceptible to environmental conditions (such as the 5HL QTL described in this study), minor genes

and their epistatic interactions play an important role in the control of the complex trait. Statistical analysis of the quantitative data showed that the Harrington allele at both the 1H and 3H QTL locations significantly and consistently resulted in greater seed dormancy in all three trials. Specific combinations of minor genes that are less sensitive to the environment may provide a more dependable source of preharvest sprouting tolerance. Establishing these gene combinations would require the identification of the genetic locations of all minor additive genes involved and large scale mapping studies. Molecular markers provide a powerful tool in which to track combinations of genes and examine their effect in order to potentially produce these superior genotypes in the future.

A key output of this Chapter was the generation of a molecular genetic map for the Stirling/Harrington double haploid population. This population has not previously been mapped and will provide a template for mapping other traits that are agronomically important in barley. The genetic map also provides a basis for mapping the candidate genes identified in Chapter 2 and other genes of interest in future studies as they arise.



QTL interaction	Interaction type(s)
5Hqa ⇔ 5Hqb	Additive and environmental
1Hq ⇔ 3Hqa	Additive
3Hqa ⇔ 5Hqa	Additive
2Hq ⇔ 3Hqa	Additive
4Hqa ⇔ 4Hqb	Additive and environment
4Hqa ⇔ 4Hqc	Additive

Figure 3.4: Genetic map of Stirling/Harrington double haploid (DH) population showing marker order and distances between them. Putative centromeric regions (based on published consensus maps, see text) are indicated by a black triangle. Chromosome regions that harbour QTLs are displayed by coloured vertical lines. QTLs defined by a blue line represent seed dormancy QTLs identified in the first germination test. QTLs marked by an asterix were only detected in epistatic interactions. QTLs defined by a green line represent seed dormancy QTLs detected using after-ripened seed 7 weeks post harvest. Main QTLs identified for seed weight are also shown (defined in red). The inset summarises QTL interactions and the interaction type.

CHAPTER 4: PHYSICAL MAPPING OF CANDIDATE GENES *ERA1* AND *ABI2* IN WHEAT-BARLEY ADDITION LINES AND DEVELOPMENT OF A CAPS MARKER FOR THE ERA1 LOCUS

4.1 Introduction

Chapter 2 of this thesis identified two candidate genes *ERA1* and *ABI2* with putative roles in the control of seed dormancy from comparative genomics using Arabidopsis and rice as model species. Both the candidate genes were aligned to wheat expressed sequence tags (ESTs) on the wheat group 3 chromosomes for which seed dormancy QTLs have been described in other studies. Chapter 3 identified a minor seed dormancy QTL on the long arm of 3H, syntenic to the wheat group 3 chromosomes, in a barley Stirling/Harrington population. Determining whether the *ERA1* and *ABI2* genes map to the seed dormancy QTL on 3H would provide additional support for these genes as candidate seed dormancy genes.

ESTs are partial transcribed nucleotide sequences derived from cDNA libraries produced from specific tissues. The National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov) database currently (January 2008) contains over a million wheat (*Triticum aestivum*) EST entries and almost half a million barley (*Hordeum vulgare*) ESTs. The large-scale generation of ESTs has provided a database for the discovery of new genes, gene structure analyses, gene homologue identification as well as molecular marker development. Polymorphic EST markers are a valuable resource for the generation of genetic markers that can be utilised to genetically map genes of interest.

A genetic tool that is used in cereals for determining the chromosomal locations of DNA markers and EST derived probes includes genomic chromosome stocks in wheat that have the addition of so-called ‘alien’ chromosomes. Foreign chromosome addition lines have numerous uses in plant genetics. They have been previously used to introduce chromosomal regions controlling valuable traits to the recipient species, examine the regulation of genes from the foreign parent, isolate individual chromosomes and genes under study and to scrutinise the behaviour of chromosomes in meiosis (Melzer et al. 1988, Islam & Shepherd 1990, Ananiev et al. 1998, Bass et al. 2000, Muehlbauer et al. 2000, Jin et al 2004, Cho et al. 2006). Chromosome addition lines have also been particularly useful for physical marker mapping and assignment of gene location. Scientists have been able to determine the physical map locations of genes and markers to a chromosome, chromosome arm or chromosome region on the basis of their presence or absence in the chromosome addition lines of various plant species (Garvin et al. 1998, Yan & Chen 2006, Masuzaki et al. 2007)

Wheat (*Triticum aestivum* L.) barley (*Hordeum vulgare* L.) chromosome addition lines have been developed through wide hybridisation using the hexaploid ($2n = 6x = 42$) wheat cultivar Chinese Spring as the recipient species and the diploid ($2n = 2x = 14$) barley cultivar Betzes as the donor species (Islam et al. 1975). The addition lines contain a complete set of wheat chromosomes and a single homeologous chromosome (disomic) or chromosome arm (ditelosomic) from barley (Islam et al. 1981; Islam 1983, Islam & Shepherd 1990, 2000). The unavailability of a wheat-barley ditelosomic addition line for the long arm of chromosome 5(1H) is due to the presence of at least one gene on 5L(1HL) resulting in sterility in wheat (Islam et al. 1981, Islam & Shepherd 1990, Taketa et al. 2002). Wheat-barley addition lines have proved to be

useful in the physical mapping of genes, isozymes and DNA markers to barley chromosomes and chromosome arms (e.g. Islam & Shepherd 1990, Garvin et al. 1998, Li et al. 2003b, Spielmeyer et al. 2004, Yan & Chen 2006).

The wheat barley addition lines provide a useful tool to determine the barley chromosome arm locations of the candidate genes, *ABI2* and *ERA1*, with putative roles in seed dormancy control identified in Chapter 2. The availability of wheat and barley EST sequences would also facilitate marker development for the genetic mapping of the genes. The aim of this chapter was to characterise and map the *ABI2* and *ERA1* genes in barley. The objectives include: a) sequence the *ABI2* and *ERA1* genes in barley cultivars Stirling and Harrington and compare the gene structures to their homologs in rice, b) compare *ABI2* and *ERA1* sequences in Stirling and Harrington for allelic variation, and c) develop EST derived markers for the *ABI2* and *ERA1* genes and physically map the location of the candidate genes in wheat barley additional lines. Detected polymorphisms would be utilised to genetically map the genes in the Stirling/Harrington doubled haploid population used to identify seed dormancy QTLs in Chapter 3. Markers developed for the genes that demonstrate linkage to a seed dormancy QTL could potentially be applied in breeding programs to select lines with seed dormancy, and hence be preharvest sprouting tolerant.

4.2 Materials and Method

4.2.1 Plant material

The plant material used in this chapter included *Triticum aestivum* L. cv. 'Chinese Spring' (CS) (chromosome recipient), *Hordeum vulgare* L. cv. 'Betzes' (B) (chromosome donor) and the six CS-B disomic addition lines harbouring 13 of the 14

barley chromosome arms, excluding the long arm of chromosome 1H. The plant material was kindly provided by Dr. Chengdao Li, Department of Agriculture and Food Western Australia. Sequencing of the candidate genes *ABI2* and *ERA1* was performed using cultivars Stirling (strongly dormant) and Harrington (non-dormant). The homozygous double haploid population used in Chapter 3 derived from a Stirling and Harrington cross and segregating for seed dormancy was used for genetic mapping of the candidate genes.

4.2.2 DNA extraction

DNA extractions of plant material were performed using standard protocols adapted from Rogowsky (1991). Leaf tissue of approximately 5 cm was collected from 2 week old plants and placed in a 1.5 mL eppendorf tube on ice before being stored at –80 ºC prior to DNA extraction. Samples were frozen in liquid nitrogen and crushed to a fine powder form with a micro-pestle. Samples were placed on ice and 600 µl of extraction buffer (1% laurel-N-Sarkosyl, 100 mM Tris HCl, 100 mM NaCl, 10 mM EDTA and 100 mM Na_2SO_3, pH 8.5) was added and homogenised using the micro-pestle. Samples were transferred to the fume hood and 600 µL of phenol/chloroform/iso-amy-alcohol, in the ratio of 25:24:1, was added and samples mixed by inversion prior to centrifugation for 10 mins at 14,000 rpm. The upper aqueous phase was transferred into a fresh 1.5 µL eppendorf tube. Phenol/chloroform/iso-amy-alcohol was added again (600 µL) and mixed. Samples were centrifuged for 5 mins at 14,000 rpm, the upper aqueous phase was transferred to a new 1.5 µL eppendorf and 60 µL of 3M sodium acetate (pH 4.8) and 600 µL of isopropanol was added. The samples were mixed by inversion and allowed to stand at room temperature for 2 mins. The precipitate was centrifuged for 5 mins at 14,000 rpm and the supernatent poured off. The DNA pellet was washed in 500

µL 70% ethanol, centrifuged for 2 mins at 14,000 rpm and the supernatent discarded. The ethanol wash was repeated and the pellet dried overnight before being suspended in 100 µL of R40 consisting of 40 µg/mL RNAse A in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). DNA quantification was performed using a Hoefer DyNA Quant 200 Fluorometer.

4.2.3 Primer design

The full length rice cDNA sequences of the *ABI2* and *ERA1* genes identified in chapter 2 and obtained from the National Centre for Biotechnology Information (NCBI) were used in BLASTN searches against the GrainGenes database (http://wheat.pw.usda.gov/GG2/index.shtml) of Triticeae ESTs. Wheat and barley ESTs that aligned to regions of the rice cDNA sequences with a nucleotide identity of 80% and an e value of < e-50 were selected for primer design. The rice cDNA sequences and the wheat and barley EST matches were aligned to genomic rice sequences to locate the intron splice junctions in the wheat and barley sequences. PCR primers were designed to conserved regions on either side of the predicted splice junctions to potentially amplify PCR products containing intron regions. This was done to exploit the increased variability that generally exists in introns. Designed primers were screened using the web-based software program Netprimer (http://www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html) to limit potential hairpins, primer dimers and cross dimers and predict the melting temperatures of the primers. Figure 4.1 displays the intron exon structure of the candidate genes and the locations of the primer pairs. The sequence data of the primers are shown in Table 4.1.

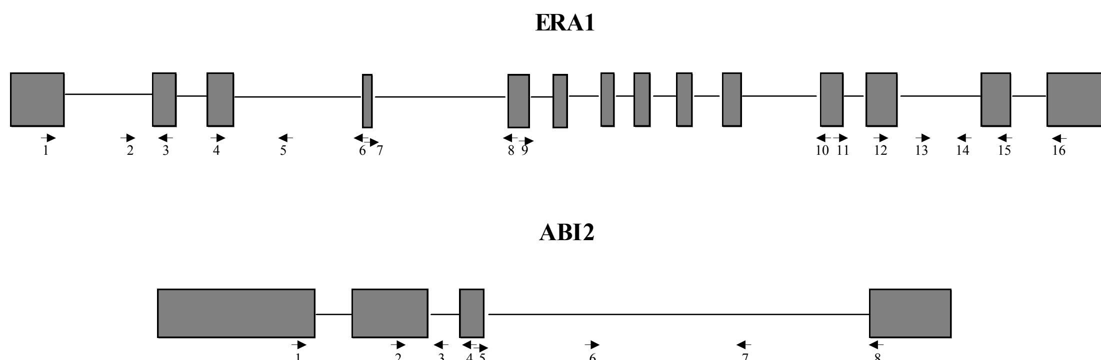


Figure 4.1: Diagram showing the exon (boxes) and intron (horizontal lines between boxes) structure of the candidate genes *ERA1* and *ABI2*. Gene structure is based on the rice *ERA1* and *ABI2* gene homologs. Diagram shows the primer locations used to amplify regions of the candidate genes and are represented by arrows. The number of the primers refers to the numbers used in Table 3.1, describing the primer sequence and name.

Table 4.1: Primers used to successfully amplify *ERA1* and *ABI2* candidate gene regions. The primer number corresponds to the numbers indicated in Figure 4.1. Accession numbers of wheat and barley ESTs from which the primers were designed from are provided.

Wheat/Barley sequence accession	Primer No.	Primer Name	Sequence (5' → 3')
ERA1			
AL506683	1	ERAi1F	CAGGTGGAGCAGATGAAGGT
AL506683	3	ERAi1R	CAGCCCTTTCGTCAAATACTC
AL506683	2	ERAi1_i3F	ATTCTGACAGGAACTGACTTG
AL506683	5	ERAi1_i3R	TTGTCTGAGCAACTTATTTATTAG
AL506683	4	ERAi3F	GAGAACGATATTGTGGACTTC
AL506683	6	ERAi3R	CACCACCATATCCACCAT
AL506683	7	ERAi4F	ATGGTGGATATGGTGGTGGA
AL506683	8	ERAi4R	GTGCTCTTTCACTCCCTATGGT
CK163448	9	ERAe5_11F	CCTCATCTCGCTACATCGTATGCTG
BY868683	10	ERAe5_11R	CCAAAAGGAGTAGCAACCATCAACC
BY868683	11	ERAe11_14abF	GTGGCATTTCGTCAAGGAGTGG
AL509620	16	ERAe11_14abR	GCTCAGGCAGTAACACGAGTGGTAG
*ERAe11_14abF fragment	12	ERAe11_14cF	ATGTTTGCCTGGTGGTGTTAG
*ERAe11_14abR fragment	15	ERAe11_14cR	TATCCAAGCACCCTCTTAGCAA
*ERAe11_14cF fragment	13	ERAe11_14dF	ATTTCAATGGATTTGGTGGACA
*ERAe11_14cR fragment	14	ERAe11_14dR	TCTATGTTTACTTTCCTTGCACCTC
ABI2			
AV835180	1	ABIe1_i2F	AGATGGAGGATGCGGTCGT
AV835180	3	ABIe1_i2R	GAAGACAAGAGCCCAGCAATG
BE445496 BG301135	2	ABIe2_3F	GTGCCGTTGTCAGTGGATC
BE445496 BG301135	4	ABIe2_3R	GGCAAGGACACCGAAAACT
BE445496 BG301135	5	ABIe3_4Fb	AGTCATACAGTGGAATGGCTACCG
BE445496 BG301135	8	ABI3_4Rb	GCAACATCGCATACCTCTTCATTC
*ABI3_4Fb fragment	6	ABIe3_4cF	AGGTTCTTATTCTTGCTCCACTTG
*ABI3_4Rb fragment	7	ABIe3_4cR	AGAGGGCTTCCACATAGAACAGG

*Based on sequencing analyses carried out in this thesis.

4.2.4 Mapping of candidate genes in Wheat Barley Additional lines – assigning barley chromosome arm location

The primers designed in Section 4.2.3 were used to amplify fragments in Stirling, Harrington, Betzes and Chinese Spring. PCRs were carried out in a total volume of 10 μL of 10 x PCR buffer containing 50 ng of template DNA, 1.5 mM $MgCl_2$, 100 μM dNTPs, 0.375 pmol of the forward and reverse primers and 0.45 units of *Taq* DNA polymerase (Fischer Biotech). Amplification of PCR products was optimised using Touchdown PCR. PCR conditions consisted of an initial denaturation step of 3 min at 94 ºC, followed by 15 cycles of 94 ºC for 30 s, touchdown from 65 ºC to 50 ºC, decreasing 1 ºC per cycle, 72 ºC for 30 s; then 25 cycles of 94 ºC for 30 s, 50 ºC for 30 s, 72 ºC for 30 and a final extension of 72 ºC for 3 mins.

PCR reactions for fragments exceeding 1 Kb, were amplified using the Qiagen LongRange enzyme (QIAGEN, Germany). Reactions with a total volume of 25 μL consisted of 50-100 ng of DNA, 2.5 mM $MgCl_2$, 500 μM of each dNTP, 0.4 μM of each primer and 2 units of the LongRange PCR Enzyme. Amplification of PCR products was performed using a PCR protocol consisting of 93 ºC for 3 min, 10 cycles of 93 ºC for 15 s, touchdown from 65 ºC to º55 C at a rate of 1 ºC per cycle for 30 s, 72 ºC for 4 min, followed by 25 cycles of 93 ºC for 15 s, 55 ºC for 30 s, 72 ºC for 4 min and a final extension of 72 ºC for 4 mins.

PCR products were separated on 1% agarose gels using 1 x TBE (0.9 M Tris, 0.9 M Boric acid, 25 M EDTA, pH 8.0) before being stained with ethidium bromide (10 mg/mL) and visualised. A Promega BenchTop 100bp DNA ladder was used as a size standard for determining fragment sizes of the amplicons. Betzes and Chinese Spring

PCR fragments were compared for polymorphisms. PCR fragments demonstrating differences in size or presence/absence between the cultivars were amplified in the wheat barley additional lines to establish the barley chromosomal arm location of the amplicons.

4.2.5 Candidate gene sequencing and development of CAPS marker

4.2.5.1 PCR Purification and sequencing reactions

PCR fragments amplified from Stirling and Harrington using the primer pairs listed in Table 4.1 were purified using Qiagen PCR cleanup kit (QIAGEN, Germany) following the manufacturer's manual. To ensure sequencing of a single fragment, the purified PCR products were electrophoresed on a 1% agarose gel and visualised by ethidium bromide staining. The purified PCR products were then directly sequenced in both the forward and reverse directions by performing 1/8 sequencing reactions with a total reaction volume of 10 µL. Sequencing reactions used 1 µL Big Dye (Applied Biosystems) dye terminator mix, 3.2 pmol/µL forward or reverse primer, 1.5 µL of 5 x sequencing buffer and 1-50 ng of purified PCR product. Reaction mixtures were made up to 10 µL with deionised water. The volumes of purified PCR products used in the reactions varied with fragment size and were adjusted accordingly. Sequencing reaction conditions consisted of an initial denaturation step at 94 ºC for 2 min, followed by 25 cycles of 94 ºC for 10 sec, 55 ºC for 35 secs and 60 ºC for 4 min.

4.2.5.2 Sequencing reaction purification and sequence analysis

Post-reaction purification was carried out on the sequencing reactions to remove salt and unincorporated dyes. Sequencing reactions were transferred to a 0.5 mL eppendorf tube and 1 µL of 125 mM EDTA, 1 µL of 3M Sodium acetate (pH 5.2) and 25 µL of

100% ethanol were added, mixed by pipetting and left at room temperature for 20 mins before being centrifuged for 30 mins at 14,000 rpm. The supernatant was discarded and the pellet rinsed with 125 µL of 70% ethanol and centrifuged for 5 mins at 14,000 rpm. The supernatent was discarded again and the pellet left to dry at room temperature for 15 mins. Sequencing reactions were then run on an ABI Prism 3700 capillary DNA analyser (Applied Biosystems, Foster, CA). The raw sequencing data was imported into the software program SeqEd (Applied Biosystems version 1.0.3) and analysed. Sequence fragments were aligned and consensus sequences constructed in MacClade 4.05 (Maddison & Maddison 1992). The Harrington and Stirling sequences were compared for polymorphisms.

4.2.5.3 Analysis and development of *ERA1* CAPs marker

Analysis of the sequence data revealed a single nucleotide polymorphism (SNP) between Stirling and Harrington in the first intron region of the *ERA1* gene. The sequence containing the SNP was screened using NEBcutter V2.0 available on the New England BioLabs website (http://www.neb.com) for a sequence specific restriction enzyme site for all Type II and commercially available Type III restriction enzymes.

Amplification of the region containing the intron 1 region was performed in the parental material and the population using primers ERAi1F and ERAi1R (primer pairs 1 and 3 for *ERA1*, Table 4.1) and the PCR conditions reported previously. For restriction analysis of the PCR fragments, 10 µL of the PCR product was digested by incubation for 1.5 h at 37C in a volume of 20 µL, containing 2 µL of 10x BSA, 2 µL 10x Restriction Buffer and 2 U of the restriction endonuclease. The incubation temperature was then increased to 80 ºC to deactivate the enzyme. The enzyme-digested PCR

product was then separated on a 1% agarose gel in 1x TBE buffer and visualised under UV light after ethidium bromide staining.

4.2.6 Genetic mapping of ERA1 candidate gene in Stirling/Harrington population

The genetic map of the double haploid Stirling/Harrrington population generated in Chapter 3 was used to genetically map the *ERA1* CAPS marker. The genetic map was made up of 128 microsatellite markers covering each of the 7 chromosomes (see Chapter 3). It should be noted that only the long arm of chromosome 3H was mapped based on the marker locations of published maps (Karakousis et al. 2003; Varshney et al. 2007). Marker data produced from screening the *ERA1* CAPS marker on the Stirling/Harrington population was imported into the Map Manager QTX file containing the genetic map of the population. The *ERA1* CAPs marker was assigned to a chromosome linkage group using the Distribution tool. Genetic map distances between the CAPs marker and other markers in the chromosome were calculated on the basis of the Kosambi map function (Kosambi 1944) with a search and linkage criterion of $P<0.0001$.

4.3 Results

4.3.1 Alignment of wheat and barley ESTs to candidate rice genes for seed dormancy

BLASTN searches identified significant wheat and barley EST alignments, for all the exons of the rice *ABI2* and *ERA1* genes. Table 4.2 shows the BLAST results for the wheat and barley EST alignments to the rice cDNA sequences and the gene exon regions covered by the alignments. The wheat and barley ESTs were used to design primers amplifying the intronic regions of the two candidate genes.

Table 4.2: Summary of BLAST alignment of wheat and barley ESTs to exon regions of rice candidate genes *ERA1* and *ABI2*. Table demonstrates BLAST e values, percentage coverage of the rice cDNA query, maximum identity of blast hit to wheat and barley EST accessions and the region of alignment on the *ERA1* and *ABI2* genes.

Gene	Alignment region on rice	EST species	EST Genbank Accession	Length (bp)	BLAST e value	Rice cDNA coverage	Max identity
ERA1 cDNA-AK073299	Exon 1 to 14	wheat	CK163448	1109	0	49%	87%
	Exon 1 to exon 6	barley	AL506683	675	2.00E-133	31%	84%
	Exon 1 to exon 5	barley	BE214204	715	1.00E-120	26%	85%
	Exon 1 to exon 5	barley	BE194182	582	5.00E-109	26%	84%
	Exon 4 to exon 11	wheat	CJ640246	662	0	40%	87%
	Exon 5 to exon 12, Exon 13 to 14	wheat	CK155788	863	0	45%	84%
	Exon 10 to exon 12	barley	BY868683	512	2.00E-84	16%	88%
	Exon 13 to exon 14	wheat	CJ897724	545	5.00E-98	21%	85%
	Exon 13 to exon 14	barley	AL509620	639	2.00E-103	20%	87%
ABI2 cDNA-AK242616	Exon 1 to exon 2	barley	AV835180	548	5.00E-106	30%	81%
	Exon 2 to exon 4	barley	BG301135	627	0	38%	86%
	Exon 1 to exon 4	wheat	BE445496	585	7.00E-171	35%	86%

4.3.2 Candidate gene sequencing and comparison of gene structure

All the primer pairs developed in Section 4.2.3 enabled successful PCR amplification and direct genomic DNA sequencing of the Stirling and Harrington parental lines, with the exception of *ABI2* primer pairs 1 and 3 (Table 4.1, Figure 4.1) in the Stirling cultivar. Comparison of the barley sequence data for the *ERA1* and *ABI2* genes and the aligned wheat and barley ESTs used to design the primers, confirmed that the targeted sequences were amplified and sequenced. Regions for which sequence data was not obtained included the 3’ and 5’ terminal ends of both the *ERA1* and *ABI2* genes. Approximately 50-100 bp was missing from exon 11 and the splice junctions at the 5’ end of exon 4, and the 3’ end of exon 5 in the *ERA1* gene. The area covering the splice junction of the 5’ end of exon 3 for the *ABI2* gene was also not sequenced.

4.3.2.1 Comparison of *ERA1* gene structure and sequence between rice and barley

In total, 7385 bp of the *ERA1* gene was sequenced. The size of the barley *ERA1* gene is therefore significantly larger than the rice (3758bp) homolog. A high degree of gene structure homology was observed between the coding regions of the rice and barley *ERA1* homologs as shown in Figure 4.2. Both the *ERA1* rice and barley homologs contained 14 exons of comparable sizes, differing at most by 5 bp in exons that were

completely sequenced. Variation was more prominent in the intron sizes between the rice and barley *ERA1* genes. With the exception of intron 2, most intronic regions were larger in barley than in their rice counterparts. In particular, intron 12 was 3695 bp in barley compared to 282 bp in rice.

Three single nucleotide polymorphisms (SNPs) were detected in the 7385 bp sequenced of the *ERA1* gene when comparing sequences from Stirling with those of Harrington (Figure 4.2). The SNPs were located in three of the four largest introns. One was located in intron one, another in intron 10 and a third in intron 12. The CAPS marker was developed from the SNP identified in intron 1. Figure 4.2 shows the location of the SNPs and highlights the location of the restriction endonuclease site used to develop the *ERA1* CAPS marker.

4.3.2.2 Comparison of *ABI2* gene structure and sequence between rice and barley

Based on alignments with wheat and barely ESTs the *ABI2* gene was expected to be approximately 3719 bp in length compared to the rice *ABI2* gene of 5738 bp. Figure 4.2 demonstrates the gene structure of the *ABI2* gene in barley and rice. 3114 base pairs of the *ABI2* gene were sequenced in the barley Stirling and Harrington cultivars, with the exception of the intron 1 region, which was only sequenced in Harrington. The barley *ABI2* gene appeared to be smaller in size compared to the rice homolog, predominantly due to smaller intron sizes. Intron 3 in the Stirling and Harrington cultivars was almost half the size as the rice counterpart (2289 bp compared to 4097 bp). Exon sizes in barley were comparable with the expected exon lengths identified in rice, differing in size by as much as 30 bp.

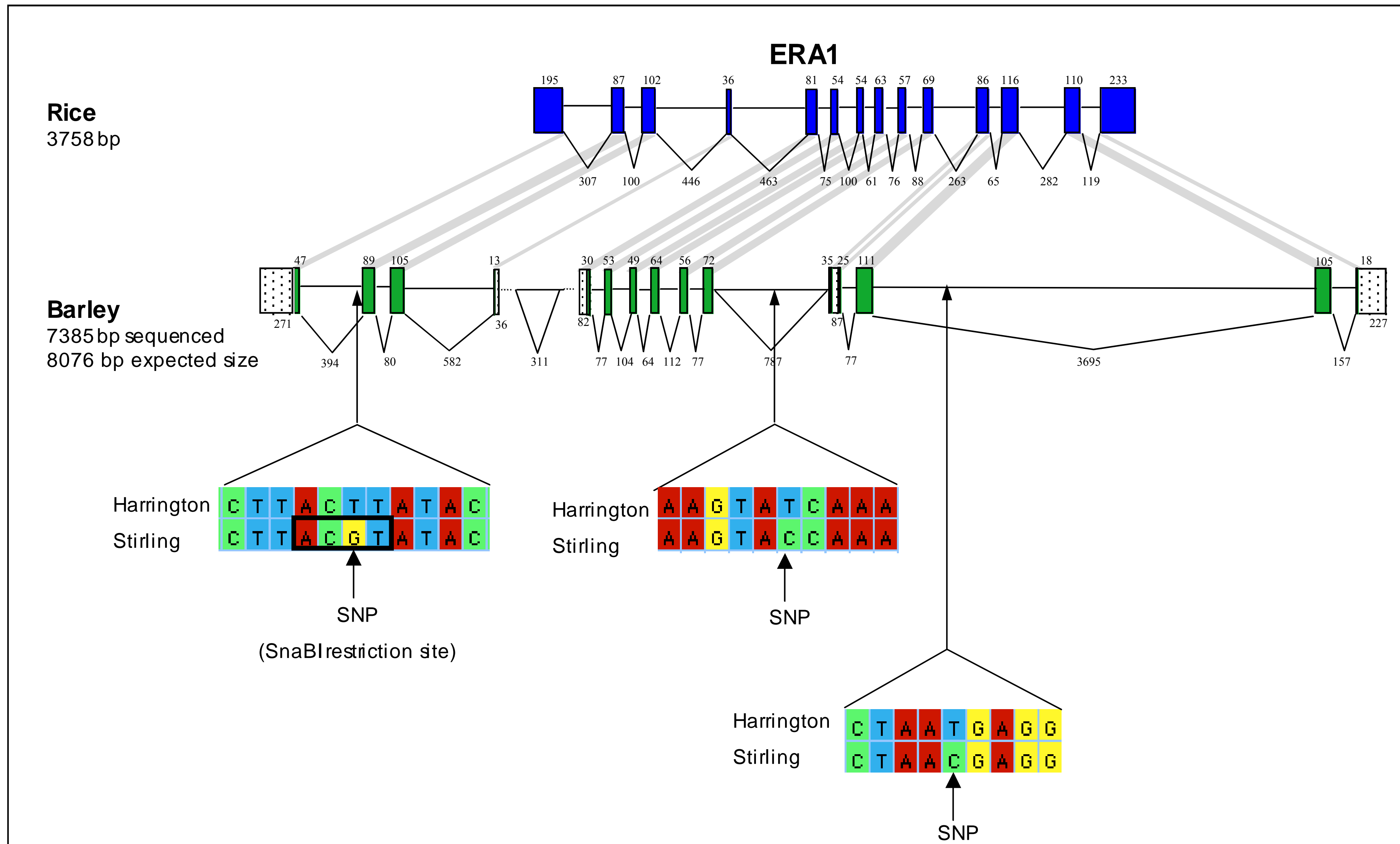


Figure 4..2 Intron-exon structure of candidate genes *ERA1* and *ABI2* in rice (blue) and sequenced barley (green) . Exons are represented by boxes, and introns by horizontal lines with the lengths of exons and introns displayed as number of base pairs. Base pairs displayed above the exons for the barley *ERA1* and *ABI2* genes represent the number of base pairs sequenced. Dotted exons and introns indicate regions without sequence data. The size of non-sequenced exon regions are based on aligned wheat and barley ESTs. Grey shading indicates regions between the orthologs that are significantly aligned. Diagram shows location and sequence of SNPs detected between Stirling and Harrington parental lines for the *ERA1* gene. The SnaBI restriction site for the *ERA1* CAPS marker is outlined.

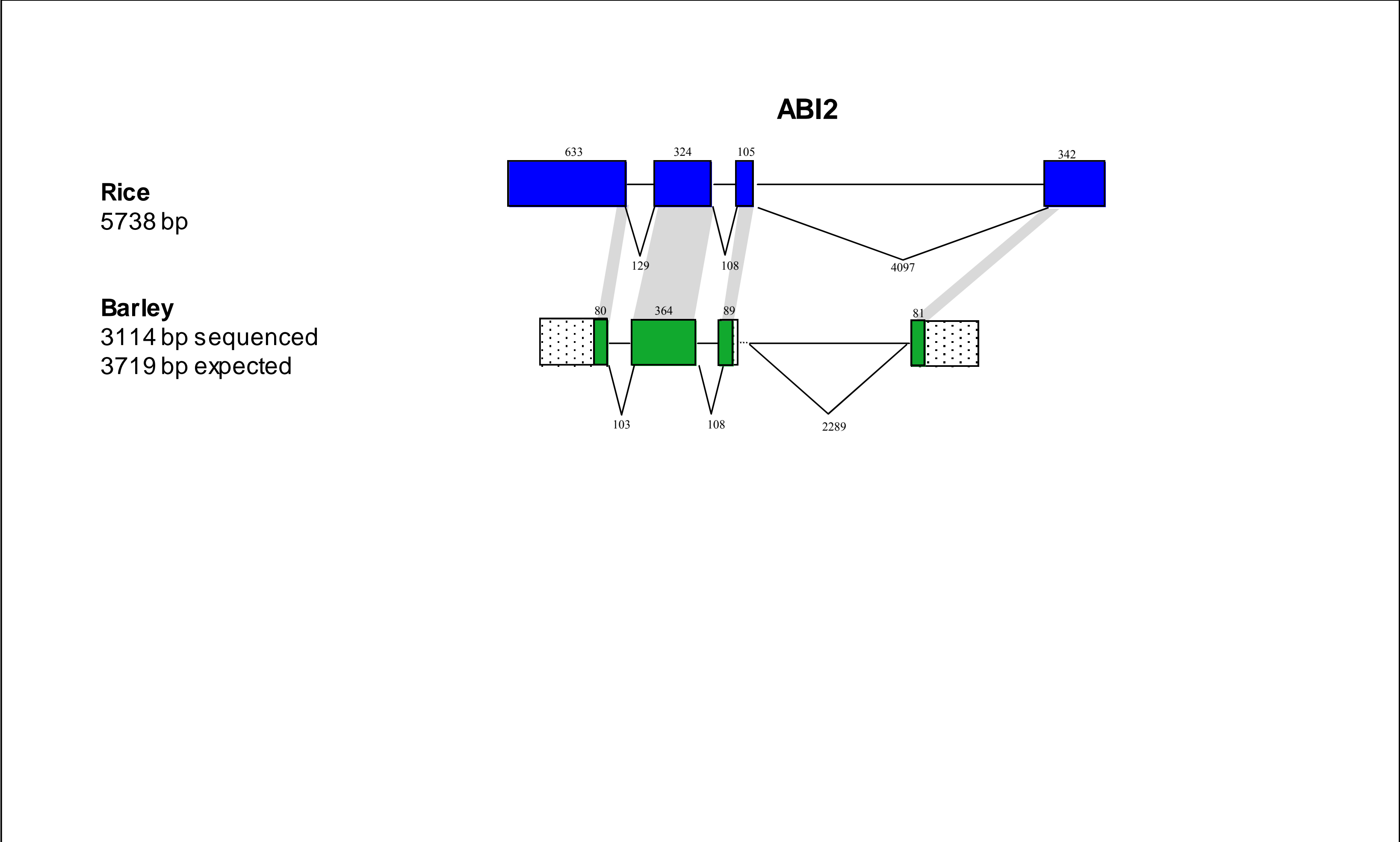


Figure 4.2 continued

Comparison of the available sequence data for the *ABI2* gene from Stirling and Harrington did not reveal any SNP differences between the two barley cultivars. If further studies indicate the *ABI2* gene to be of specific interest to the seed dormancy phenotype, more sequencing will be justified in order to genetically map the gene and examine its association to mapped seed dormancy QTLs.

4.3.3 Chromosome arm location of candidate genes

Amplification of the *ERA1* gene region spanning intron 1 to intron 2 using primers ERAi1F and ERAi1R (primers 1 and 3 for *ERA1*, Table 4.1), showed a presence/absence polymorphism between the Chinese Spring cultivar and the barley cultivar Betzes. A fragment of approximately 656bp was only amplified in the Betzes cultivar (A of Figure 4.3). The polymorphism was exploited to physically map the location of the *ERA1* gene in the wheat barley addition lines. Amplification of the fragment in the wheat barley addition lines assigned the *ERA1* fragment to the long arm of 3H.

Primers ERAi3 F and ERAi3R (primer pairs 4 and 6 for *ERA1*, Table 4.1) amplifying intron 3 of the *ERA1* gene additionally demonstrated the presence of a 643 bp fragment in the Betzes cultivar that was not present in Chinese Spring (B, Figure 4.3). Amplification of this fragment in the wheat barley addition lines assigned the fragment to the long arm of the 3H chromosome, additionally supporting the physical map location of the intron 1 fragment of the gene described above. A smaller sized band of approximately 550bp was observed in the Chinese Spring cultivar and all the wheat barley addition lines but was not amplified in Betzes.

A further polymorphism was demonstrated between Chinese Spring and Betzes when amplifying the *ABI2* gene region containing intron 1, exon 2 and intron 2. Amplification of this region using primers ABIe1_i2F and ABIe1_i2R (primer pairs 1 and 3 for *ABI2*, Table 4.1), produced a fragment of approximately 400bp in Chinese Spring while in the Betzes cultivar, the fragment amplified was approximately 725bp (C, Figure 4.3). Physical mapping of this fragment in the wheat-barley addition lines also located the fragment on the long arm of 3H.

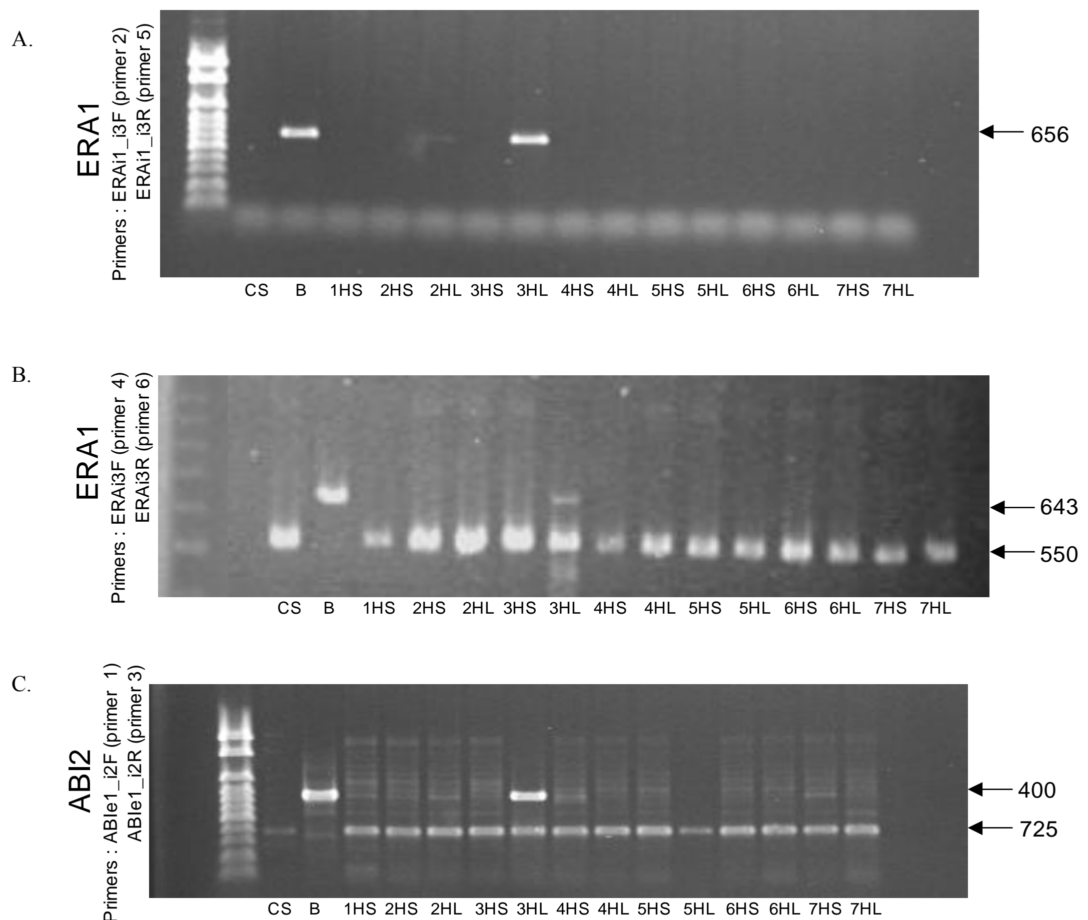


Figure 4.3: Physical mapping of candidate genes *ERA1* (A, B) and *ABI2* (C). Gel pictures of 1% agarose gels (1 x TBE) demonstrating the fragment sizes amplified in the wheat barley addition lines. Primer pairs used to amplify the fragments are indicated. The allocation of the *ERA1* gene and the *ABI2* gene to the long arm of 3H is based on the presence of a band in the Betzes (B) cultivar and the ditelosomic line for 3HL that is absent in Chinese Spring (CS) and the other wheat barley addition lines.

4.3.4 Development of SNP-CAPs marker for the ERA1 gene

Sequence comparison of genomic DNA from Stirling and Harrington parental cultivars revealed a single nucleotide polymorphism (SNP) in the first intronic region of the *ERA1* gene (Figure 4.2). A nucleotide guanine base was detected in Stirling, whereas Harrington contained a thymine base at the same location. The site of the SNP was additionally a restriction site for the restriction endonuclease SnaBI. The recognition site was 5' TACGTA 3', the enzyme cutting the fragment between the nucleotide cytosine (C) and guanine (G) bases. The restriction site was specific for Stirling, which contained a G at this site, while a thymine (T) base was detected in Harrington in the place of the G at the cutting site (Fig 4.2). Amplification of the intron 1 region of the *ERA1* gene with primers ERAi1F and ERAi1R (primer pairs 1 and 3 for *ERA1*), followed by digestion by the SnaBI enzyme produced two fragments of 428bp, and 93bp in Stirling and a single 521bp fragment of the Harrington allele. The polymorphism was easily distinguishable on a 1% agarose gel (Figure 4.4).

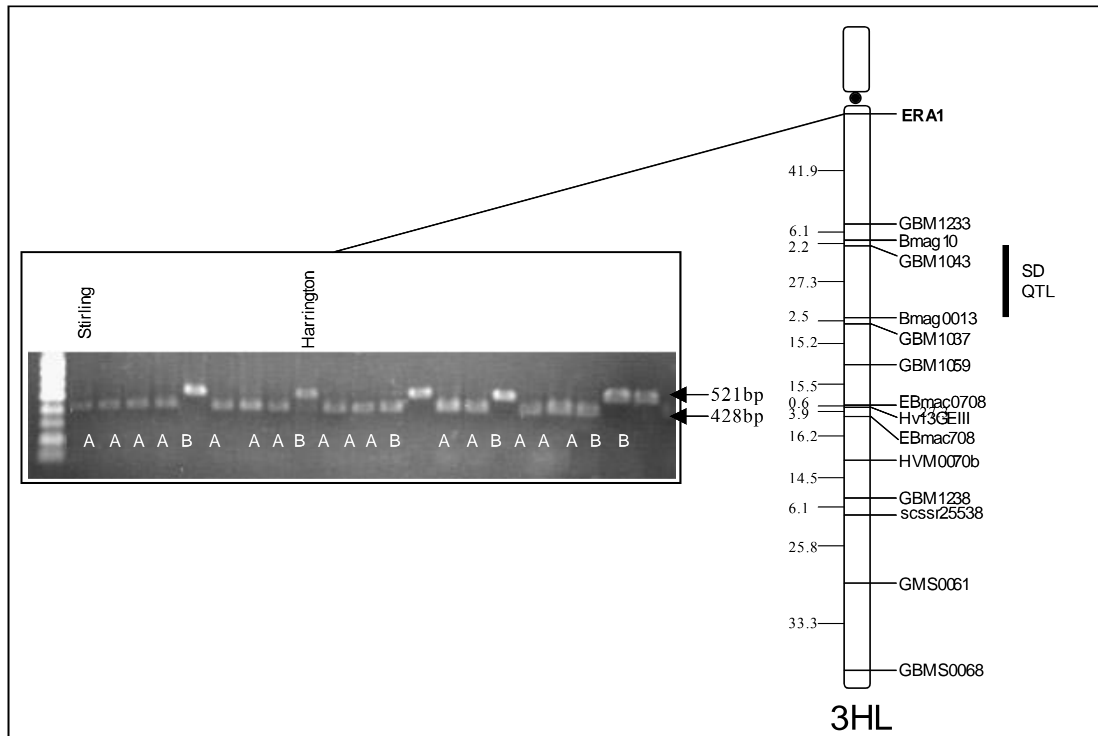


Figure 4.4: Genetic of mapping of the *ERA1* gene in the Stirling/Harrington double haploid population using the developed SNP CAPS marker. Mapping of the *ERA1* gene positions the locus at the top of the long arm of 3H. Location of the seed dormancy QTL identified in Chapter 3 is shown (SD QTL). CAPS loci were produced by amplification of genomic DNA using primers ERAi1F and ERAi1R followed by digestion of the PCR product by the restriction enzyme SnaBI and analysed on 1% agarose gel (shown in inset). Marker bands for Stirling (428 bp) and Harrington (521 bp) cultivars are indicated. Location of the centromere is approximate.

4.3.5 *Genetic mapping of ERA1 gene using the SNP-CAPS marker*

The *ERA1* CAPS marker was used to genotype the previously generated Stirling/Harrington population used to identify the seed dormancy QTL in Chapter 3. Although every line of the double haploid population was categorised by the marker, segregation of the marker was skewed in the population with 132 lines demonstrating the maternal Stirling genotype and only 52 lines genotyped with the paternal Harrington allele. Genetic mapping of the *ERA1* gene allocated the locus to the centromeric region of the long arm of the 3H chromosome, 41.9 cM above the GBM1233 microsatellite marker (Figure 4.4) with a LOD score of 4.0. The genetic mapping placed the *ERA1* gene proximal to the seed dormancy QTL region identified in Chapter 3 between the GBM1043-Bmag0013 marker interval (Figure 4.4).

4.4 Discussion

The large scale generation of publicly available ESTs has facilitated the identification of wheat and barley homologs to Arabidopsis *ERA1* and *ABI2* genes by *in silico* analysis. Comparison of the gene structures of the barley *ERA1* and *ABI2* genes (candidate genes identified in Chapter 2) with their rice homologs revealed that both the genes were largely conserved between the two plant species. The gene structures in rice and barley were highly comparable in terms of exon number and sizes. As expected, greater variation in size was observed in the intronic regions, and these were also the locations where single nucleotide polymorphisms (SNPs) were detected between the Stirling and Harrington cultivars for the *ERA1* gene. These SNPs were located in the intron regions. SNPs were not detected when comparing the sequences of the *ABI2* gene in Stirling and Harrington, although it is possible that polymorphic loci exist in the regions for which sequence data was not obtained. In the 7385 bp sequenced from the *ERA1* gene, three SNPs were detected, all of which were located in the intronic regions. Together with insertion and deletions, SNPs are one of the more frequent forms of genomic DNA variation (Cho et al. 1999, Griffin & Smith 2000, Brookes 1999). The frequency of SNPs in humans has been reported to be in the order of 1 in every 1000 base pairs (Sachidanandam et al. 2001), and even more abundant in maize with an occurrence of one SNP per 60-120 base pairs (Ching et al. 2002). Based on these frequencies, the number of SNPs observed in the *ERA1* and *ABI2* sequences is relatively low, indicating a high level of conservation.

The low allelic diversity and the results from the physical and genetic mapping performed in this study are consistent with single copies of the *ERA1* and *ABI2* genes. The conserved nature of the exons of these genes indicates that these genes have

maintained the ancestral gene function and may play a fundamental role in cellular processes of plant development. Indeed the protein phosphatase function encoded by the *ABI2* gene has been strictly conserved throughout its evolution, giving an indication of its essential nature (Rodriguez 1998). Reversable protein phosphorylation plays a fundamental role in the regulation of cellular processes in response to developmental, hormonal and environmental signals (Ohta et al. 2003). The *ERA1* gene is involved in farneslyation (Cutler et al. 1996), facilitating protein-protein or protein-membrane interactions (Zhang & Casy 1996). The function of farnesylation is ubiquitous and is involved in the regulation of various plant growth, development and signalling processes (Galichet & Gruissem 2003). Targets of the enzyme modified by *ERA1* are broad, however, its mechanism remains highly conserved (Galichet & Gruissem 2003).

Analyses conducted in Chapter 2 indicated that the *ABI2* gene had more than one ortholog in rice. Therefore, it was expected that more than one copy of the *ABI2* gene would be identified in barley. Although the barley genome is about 11-fold larger than the rice genome, they share a comparable level of complexity (Bennetzen & Kellogg 1997). However, contrary to what was expected, physical mapping of the gene in the wheat-barley addition lines indicated only one copy of the gene was present in barley. These results are an example of the expansion and contraction that occurs in gene families (Francki & Appels 2007). It appears that rice may have gained additional copies of the gene through polyploidy (whole genome duplication) or segmental duplication post speciation from barley, while barley has simply retained the original copy of the ancestral gene. Alternately, extra copies of the *ABI2* gene in barley may have been lost during localised contraction of the barley genome. While mechanisms facilitating genome growth have been the focus of numerous studies, the means by

which plant genomes are reduced are less understood. Bennetzen (2002) suggested illegitimate recombination through slipped-mispair or single-strand annealing might account for gene deletions in flowering plants. Reasons for gene deletions are unclear as most would have deleterious effects and would be selected against (Bennetzen 2002). Therefore, it seems more likely that *ABI2* gene was duplicated in rice after divergence with barley, rather than erosion of the extra barley *ABI2* copies.

The SNP located in the first intron of the *ERA1* gene served as a restriction site, facilitating the development of a CAPS marker for genetic mapping of the gene. There are several advantages associated with CAPS markers. Although the CAPS method requires an additional step of a restriction digest compared to simple PCR based markers, the method is simple and cost effective (Kunihisa et al. 2003). Inexpensive DNA extraction methods can be used and as the technique is performed using standard gel-based procedures, advanced techniques and expensive equipment are not required (Kunihisa et al. 2003). More importantly, CAPS markers are robust and reproducible (Halldén et al. 1996). PCR amplification produces the necessary DNA fragments therefore eliminating competing loci. CAPS markers are also co-dominant and can therefore distinguish heterozygotes (Miura et al. 2007) which would facilitate breeding programs that use various crosses other than double haploid populations like the one used in the present study. The *ERA1* CAPS marker could potentially be used in breeding programs to track the *ERA1* gene and observe whether it is associated with other traits of agronomic importance. The *ERA1* gene is involved in the abscisic acid signal transduction pathway in Arabidopsis (Culter et al. 1996). The abscisic acid hormone is involved in numerous processes such as those responding to abiotic stresses including drought, cool temperatures, salinity and fungal attack (Rock 2000). Thus, the

ERA1 gene may have diversified roles in plant development and cellular reactions affecting other important traits in barley.

Physical and genetic mapping of the *ERA1* and *ABI2* genes in the wheat barley addition lines placed both of the genes on the long arm of 3H. This validates the results of the bioinformatic study (Chapter 2), which predicted the location of these genes to the long arm of barley 3H chromosome based on alignments to binned wheat ESTs. Genetic mapping of the *ERA1* CAPS marker positioned the gene 41.9 cM above the GBM1233 marker. The published map position of GBM1233 is approximately 38 cM below the centromere on 3HL (Karakousis et al. 2003, Varshney et al. 2007). The distance between the GBM1233 and the *ERA1* CAPS marker is large and the gene demonstrates low linkage to GB1233 (LOD score of 4.0). Therefore, the exact position of the gene on 3HL was not resolved to a position with sufficient confidence. However, the combined results of the physical and genetic mapping in this study allow sufficient confidence that the *ERA1* gene is located on the long arm of 3H.

The *ERA1* CAPS marker demonstrated strong bias towards the maternal Stirling parent with 71% of the population demonstrating the Stirling allele at this locus. Segregation distortion is a common problem faced in mapping populations (Jiang et al. 2000) and has been reported in DH populations of tobacco (Kumashiro & Oinuma 1985), coffee (Lashermes et al. 2001), oilseed rape (Hoffmann et al. 1982), wheat (de Buyser et al. 1985, Winzeler et al. 1987), triticale (Charmet & Branlard 1985), maize (Lashermes et al. 1988, Dufour et al. 2001, Lu et al. 2002) and previously in barley (Graner et al. 1991, Zivy et al. 1992, Manninen 2000, Sayed et al. 2002). Segregation distortion has been attributed to the selective pressures of *in vitro* gametic selection in DH production

(Zivy et al. 1992) and in particular segregation distortion has been observed for 3H (Manninen 2000). At present it is difficult to determine exactly how estimates of recombination are influenced by skewed loci. Several authors have argued that segregation distortion affects calculations of map distances between loci and the order of markers on linkage groups (Garcia-Dorado & Gallego 1992, Lorieux et al. 1995a,b, Zhu et al. 2007). However, Hackett and Broadfoot (2003) in their investigation, concluded that distorted segregation ratios give difficulties in mapping only when there are two linked loci that both affect viability of the anther or microspore culture; otherwise they have little effect. Based on these contradicting findings, the precise position of the *ERA1* gene in the linear order of markers cannot be determined with absolute conviction. Mapping additional markers within closer distance to the *ERA1* gene would resolve it's positioning.

Arabidopsis *ERA1* mutants have been shown to be defective in meiosis. The development of pollen (microsporogenesis) is similar between *ERA1* mutants and the wild-type accession up to the late prophase stage of meiosis (Bonetta et al. 2000). While meiosis in the wild type is synchronous, in the *ERA1* mutants it appears to be asynchronous with microsporocytes observed at different stages of prophase, meiosis I or II or completion of meiosis. Some microsporocytes continue to undergo meiosis in later stages during callose degradation while others degenerate. This may result in unequal segregation of alleles. The pollen of *ERA1* mutants in the genetic crossing study were shown to clump together and plants that resulted from the crossings were semi-sterile (Bonetta et al. 2000). These observations may explain why the *ERA1* gene demonstrates segregation distortion in the Stirling/Harrington mapping population. Subtle variations in the *ERA1* allele of one parent could have an effect on pollen

viability and result in uneven representation of parental alleles in the generation of the double haploid lines. The addition of more markers mapped in the close vicinity of the *ERA1* CAPS marker would identify whether the segregation distortion is confined to the *ERA1* locus.

A minor QTL controlling seed dormancy in the Stirling/Harrington population was identified in the previous chapter (Chapter 3) on the long arm of 3H. As genetic mapping of the ABI2 gene on the 3H chromosome was not available, it is not possible to determine whether this gene is associated with the seed dormancy QTL identified in Chapter 3. Other genes involved in the abscisic acid signalling have been associated with QTL for seed dormancy in wheat (Nakamura et al. 2007). Nakamura et al. (2007) mapped wheat orthologs of Arabidopsis absisic acid signalling genes *ABF* and *ABI8* to seed dormancy QTL on chromosome 3A. Genetic mapping of the *ERA1* gene positioned the gene near the 3H seed dormancy QTL but it did not co-localize with the QTL. The lack of co-segregation of the *ERA1* gene with the seed dormancy phenotype indicates that the *ERA1* gene may not be directly associated in maintaining seed dormancy in the population sampled and under the environmental conditions tested. Whilst the *ERA1* gene does not appear to be directly linked to variation in the seed dormancy phenotype, the gene may play a more regulatory role within the seed dormancy gene network. Variation in seed dormancy phenotypes may also be associated with promoters and genes for transcription factors that mediate ABA responses rather than regulatory genes such as *ERA1*. The sequence data in this Chapter provides the basis for screening mutation populations to isolate barley lines that have lost *ERA1* and *ABI2* activity and thus further investigate their biological roles.

CHAPTER 5: GENERAL DISCUSSION

The aim of this thesis was to identify and characterise putative candidate genes involved in seed dormancy control and associated with seed dormancy QTLs in a new barley cross between the cultivars Stirling and Harrington. This investigation used a multidisciplinary approach, combining comparative genomics, field studies and molecular genetics. The first component of the research project provided a framework for bioinformatic analyses, using the model species Arabidopsis and rice, to identify candidate genes for the complex trait of seed dormancy. This strategy identified two candidate genes, *ABI2* and *ERA1*, which appeared to be highly conserved, as demonstrated by similar gene structures in Arabidopsis, rice and barley. The genes also displayed low allelic diversity between the barley cultivars Stirling and Harrington and mapping of the genes in barley indicated single gene copies. The *ERA1* and *ABI2* genes, known to be involved in ABA signalling in Arabidopsis, were found to co-locate to regions associated with seed dormancy QTLs in Arabidopsis, rice and wheat. The second stage of the project involved the identification of seed dormancy QTLs in a barley Stirling/Harrington cross and examination of the epistatic interactions occurring between the QTLs. The analysis detected a QTL on the long arm of 3H, a region syntenic to the wheat chromosome locations of ESTs that aligned to the *ERA1* and *ABI2* genes. This finding justified the mapping of the *ERA1* and *ABI2* genes in barley, undertaken in Chapter 4, to verify whether they co-located to the 3H QTL region. Physical mapping of the genes confirmed their location on the long arm of 3H but associations between the two genes and the 3H QTL needs to be investigated further.

Detailed analyses of gene structure and amino acid residues at intron/exon splice junctions of candidate genes *HYL1* and *ETR1* in this study showed evidence of duplication followed by divergence and this may result in altered functions in orthologous genes. The translocations, inversions and duplications that have occurred during the evolution of the Poaceae family, limit the ability to transfer genomic information from model plants to crop species and indicate the necessity for combining molecular, genetic and phenotypic evidence when identifying gene orthologs. The bioinformatic strategy used in the current study combined phenotypic (QTL) and functional (genes functionally characterised in Arabidopsis through mutagenesis) data and demonstrated that some seed dormancy genes could be tracked within seed dormancy QTLs across species as diverged as Arabidopsis and rice. As seed dormancy is a fundamental trait in the life histories of plants, it is likely that there are crucial genes involved in controlling seed dormancy that are common between even distantly related species.

The challenge for breeders is to develop a barley genotype that is dormant enough to withstand preharvest sprouting but not too deeply dormant so that the grain does not germinate during malting. Generally, there are two main genotypes in the expression of dormancy, those that release dormancy gradually and those that remain dormant for several months with a sudden release in dormancy. Ideally, a barley malting cultivar would be deeply dormant for the months during grain filling with a sudden release of dormancy immediately post harvest to reduce the cost of storage time (see Figure 5.1). Baudin is a relatively new malting cultivar that exhibits this phenotype (C. Li, Department of Agriculture and Food Western Australia, pers comm.). It could be speculated that cultivars displaying the ideal phenotype may have weak but effective

primary dormancy and a quick dormancy release such that they are unlikely to induce secondary dormancy again once they are in a quiescent state as discussed further below.

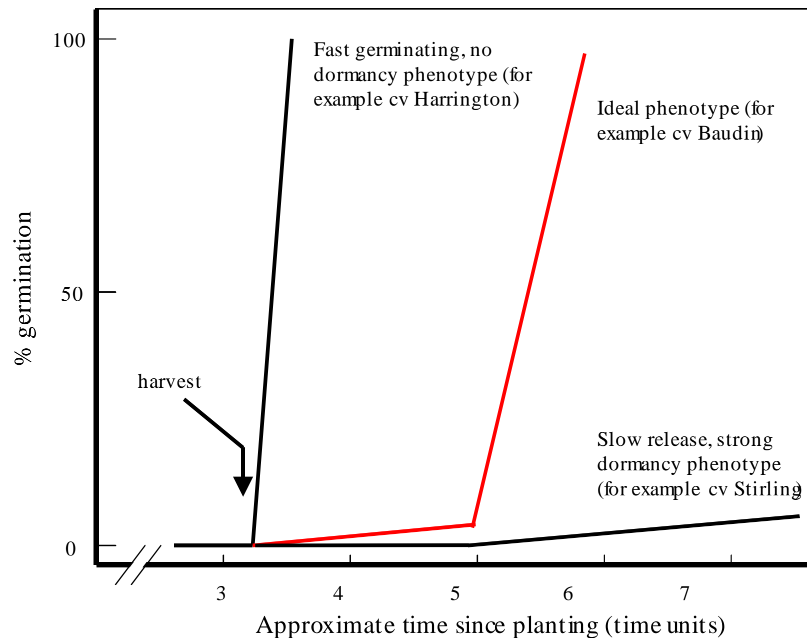


Figure 5.1: Hypothetical graph demonstrating germination behaviour of a fast germination phenotype with no dormancy such as the Harrington cultivar, a slow release phenotype with strong levels of seed dormancy such as Stirling and what may be the ideal seed dormancy phenotype displaying strong dormancy with a sudden timely release of dormancy such as Baudin.

An outcome of this investigation was the generation of a genetic map for a double haploid population derived from a cross between Stirling and Harrington that has not been previously mapped. QTL analyses on this population identified seed dormancy QTLs associated with a population generated from Australian malting cultivar Stirling and Harrington, a well-established Canadian malting cultivar. Stirling has been a major malting cultivar in Western Australia since the beginning of the 1980's and demonstrates strong seed dormancy. Stirling is also a parent of new malting barley cultivar Baudin, which combines both high levels of seed dormancy and malting quality.

Li et al. (2003a) emphasized that the major challenge in breeding for seed dormancy is linkage of the trait with inferior malting qualities. A number of studies have investigated seed dormancy QTLs in the barley feed cultivars Steptoe and Morex (Ullrich et al. 1993, Oberthur et al. 1995, Han et al. 1996). This thesis looks specifically at seed dormancy in the background of a malting cultivar, adding significant value to the current understanding of a trait that is a major concern for the malting industry.

The QTL analyses performed on the Stirling/Harrington population in this thesis identified two seed dormancy QTLs on the 5H chromosome, that were consistent with the SD1 and SD2 QTLs detected several seed dormancy studies using barley feed cultivars (Ullrich et al. 1993, Oberthur et al. 1995, Han et al 1999, Gao et al. 2003). The SD1 and SD2 loci have been proposed to have different physiological functions, with SD1 affecting the formation and maintenance of dormancy and SD2 associated with the release of dormancy (Prada et al. 2004). Although the location of the SD1 and SD2 loci are comparable to the QTLs identified on 5H in the current study, the Stirling dormancy allele could in fact represent a different gene(s) to the 5H QTL in the previous studies. The centromeric QTL on 5H (SD1) has been reported as the major QTL in previous studies and the QTL on the long arm of 5H (SD2) as a minor QTL. In the current study, the major QTL is on the long arm. This QTL may be therefore more relevant to malting cultivars and Australian germplasm. In a recent paper (Lin et al. 2008, *Crop Science* in press), the authors mapped QTLs for seed dormancy from a cross of Morex/Harrington. Although the same QTLs on chromosome 5H were identified, the non-dormancy allele from Morex in the studies by Ullrich et al. (1993) is the dormant allele in the Morex/Harrington population. Lin et al. (2008, *Crop Science* in press) additionally

suggests that there may be two different QTLs in the telomere region of 5HL. Thus, the nature of the QTLs for seed dormancy on chromosome 5H remains to be resolved.

Li et al. (2004) conducted comparative genomics using rice, wheat and barley in the region of the SD2 locus on the long arm of 5H and suggested the *gibberellin (GA) 20-oxidase* (encoded by the *GA5* gene in Arabidopsis) was the underlying gene for this QTL. Since then, further evidence implicating the role of *GA 20-oxidase* in seed dormancy has emerged. Appelford et al. (2006) observed high levels of expression of the enzyme as well as *GA 3-oxidase*, a member of the same gene family, in developing and germinating wheat grains. In addition, the *GA 20-oxidase* gene homoeologues map to wheat chromosomes 5BL, 5DL and 4AL, syntenic to the region of the 5HL QTL in barley (Appelford et al. 2006). Figure 5.2 summarises some key features of seed germination and in this theoretical overview the GA levels are postulated to be the primary 'switch' for germination.

Interestingly, although the seed dormancy QTL on the long arm of 5H is a prominent QTL in this study and several other barley studies mentioned above, the comparative genomics conducted in this research did not identify a candidate gene located in this region. The *GA5* Arabidopsis ortholog to the *GA 20-oxidase* gene did not co-locate to an Arabidopsis seed dormancy QTL in the bioinformatics investigation of this study and thus was not selected as a candidate gene. This may be due to the limited number of QTLs detected from the few seed dormancy QTL studies performed in Arabidopsis. Alternatively, the primary drivers for dormancy may vary between Arabidopsis and barley. For example, GA biosynthesis may be the first event inducing the seed germination process in barley and dormancy may be initially controlled by changes in

GA levels followed by ABA effects (see Figure 5.2). Arabidopsis may adopt a different strategy, for instance, regulating dormancy initially with light induced hormone interactions (Figure 5.2). The expression of genes involved with ABA and GA signaling and GA synthesis have been shown to be modulated by phytochrome (Yamaguchi et al. 1998a, 2001, Ogawa et al. 2003, Mazella et al. 2005). The sequenced Arabidopsis genome may therefore not always be informative for the biological processes that have evolved independently in other plant species such as barley. This is reflected in the different requirements plants need for germination. For example, Arabidopsis requires light and stratification (Penfield et al. 2005) for germination whereas barley seeds can germinate in darkness. Whilst functional studies in Arabidopsis have been pivotal in revealing genes that play a role in the plant's seed dormancy, similar molecular approaches need to be used in other seed model species to gain a broader understanding of the different strategies plants use to regulate germination processes.

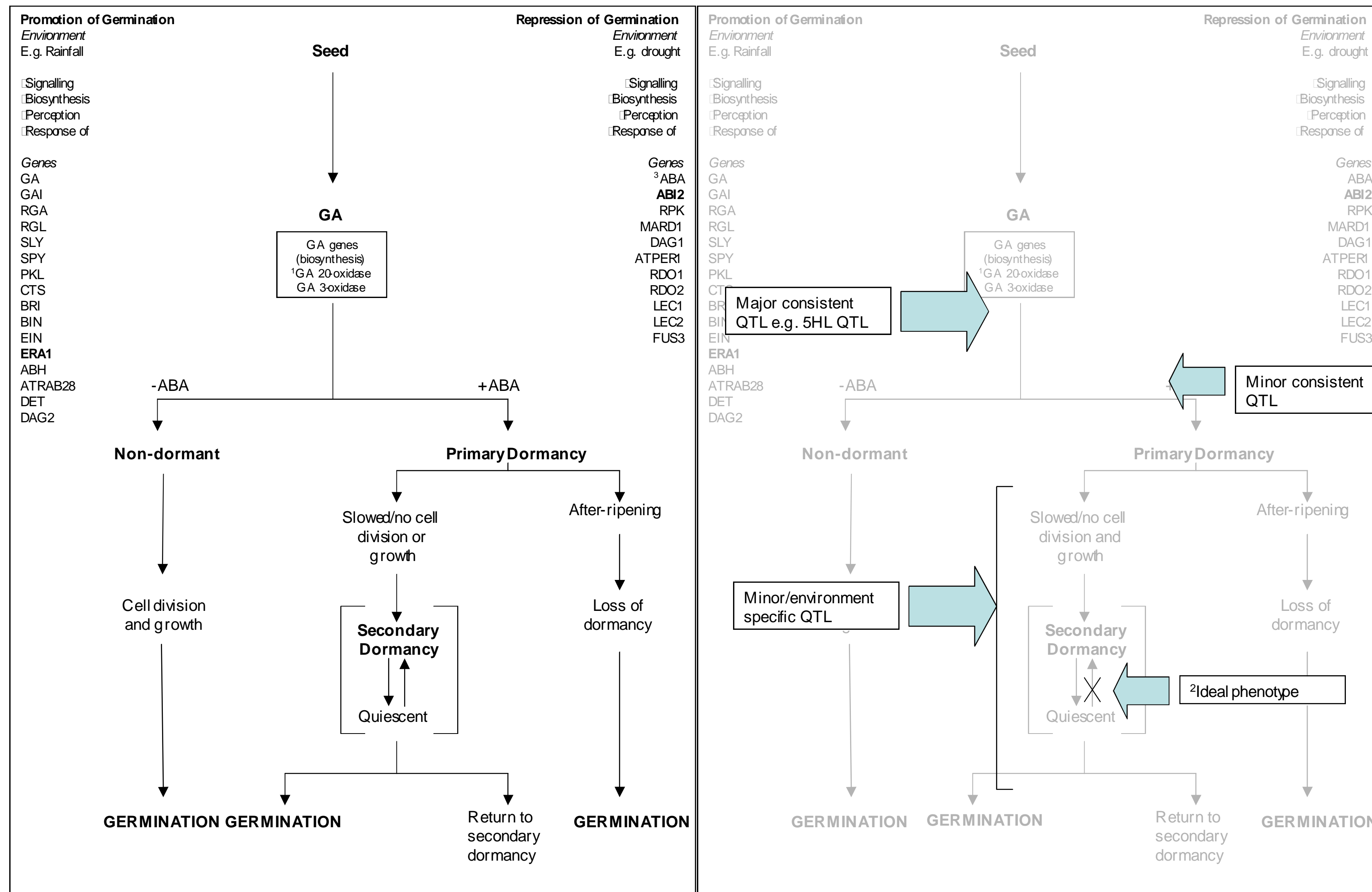


Figure 5.2: Aspects of seed germination pathway in barley. Examples are given, of environmental conditions conducive to germination or dormancy and highlighting seed dormancy genes (most of which are described in the literature review) functionally characterised in Arabidopsis that either promote or repress germination. Candidate genes ERA1 and ABI2 are in bold. Environmental conditions would be influential throughout the process from seed development to germination. Certain phases would be more sensitive to the environment, particularly after the loss of primary dormancy. The figure makes suggestions for where QTL may come into effect.

[1]GA 20-oxidase in cereals is orthologous to GA5 in Arabidopsis.

[2]See Figure 5.1 for a description of what may be an ideal seed dormancy phenotype in barley.

[3]Vp1 in cereals is orthologous to ABA3 in Arabidopsis.

Generally, a seed's main goal is to germinate and grow into a seedling. Germination could then be viewed as the default process with GA acting as the key player initiating and regulating germination while ABA counteracts to repress germination. As summarized in Figure 5.2, it is possible that the germination/dormancy pathway that a barley seed may experience in which the process is primarily driven by GA biosynthesis, involving genes such as the *GA 20-oxidase* and the *GA 3-oxidase* genes. Dormancy may be initially characterized by low levels of GA and then further influenced by the presence of endogenous ABA that would further repress germination. Low ABA levels would lead to non-dormant seeds that would continue cell division and growth, processes that would be slowed down or temporarily arrested in dormant seeds at higher levels of ABA. Dormant seeds in conditions of higher levels of ABA would either go through after-ripening and germinate or experience secondary dormancy and quiescence depending on the environmental conditions. The environment would be influential throughout the seed development from maturation to germination; however, certain parts of this process would be more sensitive to environmental conditions, particularly those controlled by genes modulated by factors such as temperature, light and moisture.

An inherent challenge in dissecting the seed dormancy trait is to understand what stage of development the phenotyping is targeting and thus what part of the seed germination pathway the detected QTLs represent. Figure 5.2 makes some suggestions of where QTLs may come into effect in the hypothetical germination/dormancy pathway. It is speculated that major and consistent QTLs such as the 5HL QTL in barley may represent genes, such as *GA 20-oxidase* and *GA 3-oxidase*, involved in independent and fundamental seed germination processes that are active "early" in the process from seed

maturation to germination. Chapter 3 identified alleles on 1H and 3H that consistently resulted in greater dormancy in all the environments tested, suggesting that some QTLs may represent genes involved in dormancy that are less receptive to the environment compared to those that are governed or triggered by environmental cues. Identifying the genes underlying these QTLs would prove to be beneficial for breeders in providing them with a more reliable source of dormancy. Environmental factors, such as moisture in particular, could be expected to have greater influence following the release of primary dormancy and thus affecting after-ripening or when the seed is cycling between secondary dormancy and a quiescent state. Germination of seeds in this phase predominantly relies on favourable environmental conditions. Minor QTLs or those detected only in certain environments may represent genes that are more responsive to the environment and become active in specific environmental conditions. Lee et al. (2002) characterised the first gene connecting environmental and endogenous interactions in seed germination. The Arabidopsis *RGL2* gene encoding a transcription factor is induced by moisture and commences signalling by GA (Lee et al. 2002). Identification of genes such as the *RGL2*, that act at the interface between environmental cues and hormone response would elucidate the pathways by which environmental factors influence phenotypic variation. Once all the mechanisms regulating dormancy and germination are identified for certain environmental conditions, it will be possible to attribute QTLs to certain parts of the dormancy pathway and/or target phenotyping to elucidate specific pathway nodes.

Future initiatives

The approach used in this thesis was to identify candidate genes for seed dormancy through comparative genomics and investigate whether these genes related to

chromosomal regions associated with seed dormancy detected in a barley population. An alternative approach would be conduct detailed comparative genomics of a chromosomal region in barley such as the 5HL region, known to be involved in seed dormancy control. *GA 20-oxidase* gene has been suggested as the underlying gene controlling dormancy in this region; however, the 5HL QTL itself may be the result of multiple loci. Considering the sequence data available for rice, identifying rice chromosomal regions syntenic to the 5HL locus in rice and subsequent data mining of these regions may reveal candidate genes with a role in seed dormancy in rice similar to the *GA 20-oxidase* gene.

The *ABI2* gene is worthy of further investigation in the two barley cultivars Stirling and Harrington. Although the gene was physically mapped to the long arm of 3H, its chromosomal location cannot be correlated to the QTL identified on 3H in Chapter 3 without further genetic mapping. Sequencing the entire gene plus neighbouring regions would reveal whether any polymorphisms exist between the Stirling and Harrington cultivars at this locus and enable the gene to be genetically mapped to verify whether it is associated with the 3H seed dormancy QTL.

QTL analyses can only be as informative as the genetic map used to detect the QTLs. Diversity Arrays Technology (DArT) is a hybridisation-based technology that is high-throughput and is capable of identifying hundreds to thousands of polymorphisms in a single assay (Wenzl et al. 2004). Genomic representations generated from genomic DNA samples are hybridised to a DArT array consisting of a library of the species and screened for the presence or absence of variable DNA fragments (Jaccoud et al. 2001). Originally developed for rice (Jaccoud et al. 2001), the technology has been successful

in generating some high-density genetic maps in barley and wheat (Akbari et al. 2006, Wenzl et al. 2004, Hearnden et al. 2007) amongst other plant species. This study conducted DArT analyses on 90 lines of the Stirling/Harrington population in preliminary investigations of the use of DArT markers to resolve regions of the SSR genetic map with low marker coverage, particularly in map areas containing seed dormancy QTLs. New polymorphic markers were identified for regions of the Stirling/Harrington genetic map that were previously devoid of markers (pilot map shown in Appendix C). As the analysis was performed on only half the population, marker order and positions were not reliable due to the large amount of missing data confounding statistical analyses. However, these preliminary results indicate DArT analysis to be a promising method to resolve regions of the Stirling/Harrington genetic map that lack molecular markers and possibly identify new seed dormancy QTLs.

Ultimately the role of a gene can only be reliably defined by conducting functional analyses. Mutational studies in Arabidopsis have enabled the elucidation of gene function for many important genes including those involved in ABA and GA biosynthesis, signalling and regulation pathways. Similar mutational studies using gene disruption techniques such as RNA-interence (RNAi) technology to inhibit or overexpress the *ERA1* and *ABI2* genes may reveal the roles of the genes in seed dormancy. Another method of mutagenesis involves heavy ion-bombardment has been used to introduce mutations in plant species (Tanaka 1999, Abe et al. 2000) particularly in ornamental plants to generate novel cultivars. This technology could be applied to barley cultivars with different seed dormancy. Those that demonstrate a significant change in their dormancy phenotype could be examined to identify whether alteration in the phenotype could be attributed to mutations in candidate genes such as *ABI2* and

ERA1. Although seed expression studies have been limited, most of the genes involved in ABA and GA biosynthesis have been identified, at least in Arabidopsis. Expression profiling is capable of demonstrating the expressed portion of genes involved in developmental processes (Alba et al. 2004, Xue et al. 2006) such as dormancy and germination. Gene expression profiling techniques include cDNA-AFLP, mRNA microarray analysis and another relatively new method of eQTL analysis, which measures the expression of genes in population genotypes using real-time RT-PCR and maps gene expression as a quantitative trait (Hazen et al. 2003, Potokina et al. 2006). Examination of the expression of the *ERA1* and *ABI2* genes at different developmental and growth phases of the seed and additionally in different seed tissue types would demonstrate whether these genes are active during seed maturity and germination (or lack thereof). These experiments would also provide a more comprehensive understanding of when these genes are active and where they are expressed. Finally, should expression studies demonstrate a strong involvement of the *ERA1* and *ABI2* genes in seed dormancy, transgenic approaches would enable scientist to observe the effect of these genes individually in different genetic backgrounds and under various environmental conditions.

REFERENCES

Aalen, R. B. (1999). Peroxiredoxin antioxidants in seed physiology. *Seed Science Research* **9**: 285-295.

Abe, T., C. H. Bae, T. Ozaki, J. M. Wang and S. Yoshida (2000). Stress-tolerant mutants induced by heavy-ion beams. *Gamma Field Symposium* **39**: 45-56.

Abel, S., M. D. Nguyen and A. Theologis (1995). The PS-IAA4/5-like family of early auxin-inducible mRNAs in *Arabidopsis thaliana*. *Journal of Molecular Biology* **251**: 533-49.

Achard, P., H. Cheng, L. De Grauwe, J. Decat, H. Schoutteten, T. Moritz, D. Van Der Straeten, J. Peng and N. P. Harberd (2006). Integration of plant responses to environmentally activated phytohormonal signals. *Science* **311**: 91-94.

Adkins, S. W., S. M. Bellairs and D. S. Loch (2002). Seed dormancy mechanisms in warm season grass species. *Euphytica* **126**: 13-20.

Ahn, S., J. A. Anderson, M. E. Sorrells and S. D. Tanksley (1993). Homoeologous relationships of rice, wheat and maize chromosomes. *Molecular and General Genetics* **241**: 483-90.

Ahn, S. N. and S. D. Tanksley (1993). Comparative linkage maps of the rice and maize genomes. *Proceedings of the National Academy of Sciences of the United States of America* **90**: 7980-7984.

Akbari, M., P. Wenzl, V. Caig, J. Carling, L. Xia, S. Yang, G. Uszynski, V. Mohler, A. Lehmensiek, H. Kuchel, M. J. Hayden, N. Howes, P. Sharp, P. Vaughan, B. Rathmell, E. Huttner and A. Kilian (2006). Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theoretical Applied Genetics* **113**: 1409-20.

Alba, R., Z. J. Fei, P. Payton, Y. Liu, S. L. Moore, P. Debbie, J. Cohn, M. D'Ascenzo, J. S. Gordon, J. K. C. Rose, G. Martin, S. D. Tanksley, M. Bouzayen, M. M. Jahn and J. Giovannoni (2004). ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. *Plant Journal* **39**: 697-714.

Allen, P. S., R. L. Benech-Arnold, D. Batlla and K. J. Bradford (2007). Modeling of seed dormancy. Seed development, dormancy and germination. K. Bradford and H. Nonogaki. Oxford, Blackwell Publishing Ltd**:** 72-112.

Alonso, J. M., T. Hirayama, G. Roman, S. Nourizadeh and J. R. Ecker (1999). EIN2, a bifunctional transducer of ethylene and stress responses in Arabidopsis. *Science* **284**: 2148-2152.

Alonso-Blanco, C., L. Bentsink, C. J. Hanhart, H. B. E. Vries and M. Koornneef (2003). Analysis of natural allelic variation at seed dormancy loci of Arabidopsis thaliana. *Genetics* **164**: 711-729.

Ananiev, E. V., R. L. Phillips and H. W. Rines (1998). Chromosome-specific molecular organization of maize (Zea mays L.) centromeric regions. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 13073-13078.

Anderson, J. A., M. E. Sorrells and S. D. Tanksley (1993). RFLP analysis of genomic regions associated with resistance to preharvest sprouting in wheat. *Crop Science* **33**: 453-459.

Appels, R., M. Francki, M. Cakir and S. M. Bellairs (2004). Looking through genomics: concepts and technologies for plant and animal genomics. *Functional Integrated Genomics* **4**: 71-73.

Appleford, N. E. J., D. J. Evans, J. R. Lenton, P. Gaskin, S. J. Croker, K. M. Devos, A. L. Phillips and P. Hedden (2006). Function and transcript analysis of gibberellin-biosynthetic enzymes in wheat. *Planta* **223**: 568-582.

Arenas-Huertero, F., A. Arroyo, L. Zhou, J. Sheen and P. Leon (2000). Analysis of Arabidopsis glucose insensitive mutants, GIN5 and GIN6, reveals a central role of the plant hormone ABA in the regulation of plant vegetative development by sugar. *Genes and Development* **14**: 2085-2096.

Badr, A., K. Muller, R. Schafer-Pregl, H. El Rabey, S. Effgen, H. H. Ibrahim, C. Pozzi, W. Rohde and F. Salamini (2000). On the origin and domestication history of Barley (*Hordeum vulgare*). *Molecular Biology and Evolution* **17**: 499-510.

Bailey, P. C., R. S. McKibbin, J. R. Lenton, M. J. Holdsworth, J. E. Flintham and M. D. Gale (1999). Genetic map locations for orthologous Vp1 genes in wheat and rice. *Theoretical and Applied Genetics* **98**: 281-284.

Ballard, L. V. (1973). Physical barriers to germination. *Seed Science Technology* **1**: 285-303.

Barr, A. R., A. Karakousis, L. R.C.M, S. J. Logue, S. Manning, K. J. Chalmers, J. M. Kretschmer, B. W.J.R, H. M. Collins, S. Roumeliotis, S. J. Coventry, M. D.B., B. J. Read, D. Poulsen, L. C.D, P. G.J., P. A. Inkerman, J. F. Panozzo, B. R. Cullis, A. B. Smith, P. Lim and P. Langridge (2003). Mapping and QTL analysis of the barley population Chebec x Harrington. *Australian Journal of Agricultural Research* **54**: 1125-1130.

Bartel, B. and G. R. Fink (1995). ILR1, an amidohydrolase that releases active indole-3-acetic-acid from conjugates. *Science* **268**: 1745-1748.

Baskin, C. C. and J. M. Baskin (1998). Seeds: ecology, biogeography, and evolution of dormancy and germination, San Diego, Academic Press.

Baskin, J. M. and C. C. Baskin (2004). A classification system for seed dormancy. *Seed Science Research* **14**: 1-16.

Bason, M., J. Ronalds and C. Wrigley (1991). Screening for pre-germination of barley at harvest. Proceedings of the 40th Australian Cereal Chemistry Conference, Melbourne, RACI.

Bass, H. W., O. Riera-Lizarazu, E. V. Ananiev, S. J. Bordoli, H. W. Rines, R. L. Phillips, J. W. Sedat, D. A. Agard and W. Z. Cande (2000). Evidence for the coincident initiation of homolog pairing and synapsis during the telomere-clustering (bouquet) stage of meiotic prophase. *Journal of Cell Science* **113**: 1033-1042.

Batlla, D. and R. L. Benech-Arnold (2005). Changes in the light sensitivity of buried *Polygonum aviculare* seeds in relation to cold-induced dormancy loss: development of a predictive model. *New Phytologist* **165**: 445-52.

Baumbusch, I. O., D. W. Hughes, G. A. Galau and K. S. Jakobsen (2004). LEC1, FUS3, ABI3 and Em expression reveals no correlation with dormancy in Arabidopsis. *Journal of Experimental Botany* **55**: 77-87.

Baumlein, H., S. Misera, H. Luerssen, K. Kolle, C. Horstmann, U. Wobus and A. J. Muller (1994). The FUS3 gene of *Arabidopsis thaliana* is a regulator of gene expression during late embryogenesis. *Plant Journal* **6**: 379-387.

Beaudoin, N., C. Serizet, F. Gosti and J. Giraudat (2000). Interactions between abscisic acid and ethylene signaling cascades. *Plant Cell* **12**: 1103-1115.

Becraft, P. W. (1998). Receptor kinases in plant development. *Trends in Plant Science* **3**: 384-388.

Benech-Arnold, R. L., R. A. Sanchez, F. Forcella, B. C. Kruk and C. M. Gheresa (2000). Environmental control of dormancy in weed seed banks in soil. *Field Crops Research* **67**: 105-122.

Benech-Arnold, R. L. (2002). Bases of pre-harvest sprouting resistance in barley: physiology, molecular biology, and environmental control of dormancy in the barley grain. Barley science recent advances from molecular biology to agronomy of yield and quality. G. A. Slafer, Molina-Cano J.L, R. Savin, J. L. Araus and I. Romagosa. New York, Food Products Press**:** 481-503.

Bennett, M. D. and J. B. Smith (1976). Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society of London, Series B* **274**: 227-274.

Bennetzen, J. L. and E. A. Kellogg (1997). Do plants have a one-way ticket to genomic obesity? *Plant Cell* **9**: 1509-1514.

Bennetzen, J. L. (2000). Comparative sequence analysis of plant nuclear genomes:m microcolinearity and its many exceptions. *Plant Cell* **12**: 1021-1029.

Bennetzen, J. L. (2002). Mechanisms and rates of genome expansion and contraction in flowering plants. *Genetica* **115**: 29-36.

Bennetzen, J. L. and W. Ramakrishna (2002). Numerous small rearrangements of gene content, order and orientation differentiate grass genomes. *Plant Molecular Biology* **48**: 821-7.

Bennetzen, J. L. and J. Ma (2003). The genetic colinearity of rice and other cereals on the basis of genomic sequence analysis. *Current Opinion in Plant Biology* **6**: 128-33.

Bentsink, L. and M. Koornneef (2002). Seed dormancy and germination. The Arabidopsis Book. The Netherlands, American Society of Plant Biologists**:** 1-18.

Bentsink, L., W. Soppe and M. Koornneef (2007). Genetic aspects of seed dormancy. Seed development, dormancy and germination. K. Bradford and H. Nonogaki. Oxford, Blackwell Publishing Ltd**:** 113-132.

Bewley, J. D. and M. Black (1994). Seeds: Physiology of Development and Germination. New York, Plenum Publishing Corporation.

Bewley, J. D. (1997). Seed Germination and Dormancy. *The Plant Cell* **9**: 1055-1066.

Biddulph, T. B., J. A. Plummer, T. L. Setter and D. J. Mares (2007). Influence of high temperature and terminal moisture stress on dormancy in wheat (*Triticum aestivum* L.). *Field Crops Research* **103**: 139-153.

Blanc, G., A. Barakat, R. Guyot, R. Cooke and M. Delseny (2000). Extensive duplication and reshuffling in the Arabidopsis genome. *Plant Cell* **12**: 1093-1101.

Blanc, G. and K. H. Wolfe (2004). Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. *Plant Cell* **16**: 1679-1691.

Bleecker, A. B., M. A. Estelle, C. Somerville and H. Kende (1988). Insensitivity to ethylene conferred by a dominant mutation in *Arabidopsis thaliana*. *Science* **241**: 1086-1089.

Bleecker, A. B. (1999). Ethylene perception and signaling: an evolutionary perspective. *Trends in Plant Science* **4**: 269-274.

Bonetta, D. and P. McCourt (1998). Genetic analysis of ABA signal transduction pathways. *Trends in Plant Science* **3**: 231-235.

Bonetta, D., P. Bayliss, S. Sun, T. Sage and P. McCourt (2000). Farnesylation is involved in meristem organization in Arabidopsis. *Planta* **211**: 182-190.

Borrell, A., M. C. Cutanda, V. Lumbreras, J. Pujal, A. Goday, F. A. Culianez-Macia and M. Pages (2002). *Arabidopsis thaliana* ATRAB28: a nuclear targeted protein related to germination and toxic cation tolerance. *Plant Molecular Biology* **50**: 249-259.

Bowers, J. E., B. A. Chapman, J. Rong and A. H. Paterson (2003). Unraveling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422**: 433-438.

Boyd, W. J. R., C. D. Li, C. R. Grime, M. Cakir, S. Potipibool, L. Kaveeta, S. Men, M. R. Jalal Kamali, A. R. Barr, D. B. Moody, R. C. M. Lance, S. J. Logue, H.

Raman and B. J. Read (2003). Conventional and molecular genetic factors contributing to the variation in the timing of heading among spring barley (*Hordeum vulgare* L.) genotypes grown over a mild winter growing season. *Australian Journal of Agricultural Research* **54**: 1277-1301.

Brewing and Malting Barley Research Institute (2005). Quality factors in malting barley, Winnipeg, Canada, Brewing and Malting Barley Research Institute.

Brocard-Gifford, I., T. J. Lynch, M. E. Garcia, B. Malhotra and R. R. Finkelstein (2004). The Arabidopsis thaliana ABSCISIC ACID-INSENSITIVE8 encodes a novel protein mediating abscisic acid and sugar responses essential for growth. *Plant Cell* **16**: 406-421.

Brookes, A. J. (1999). The essence of SNPs. *Gene* **234**: 177-186.

Cai, H. W. and H. Morishima (2000). Genomic regions affecting seed shattering and seed dormancy in rice. *Theoretical and Applied Genetics* **100**: 840-846.

Cai, W. and H. Morishima (2002). QTL clusters reflect character associations in wild and cultivated rice. *Theoretical Applied Genetics* **104**: 1217-1228.

Carreck, N. L. and D. G. Christian (1997). A study of grain dormancy and viability in spring barley. *European Journal of Agronomy* **6**: 155-161.

Castor, L. and F. Frrederiken (1977). Seed moulding of grain sorghums caused by *Fusarium* and *Curvularia*. *Proceedings of Annual Phytopathology Society* **4**: 151.

Causse, M. A., T. M. Fulton, Y. G. Cho, S. N. Ahn, J. Chunwongse, K. Wu, J. Xiao, Z. Yu, P. C. Ronald, S. E. Harrington, G. Second, S.R McCouch and S.D. Tanksley (1994). Saturated molecular map of the rice genome based on an interspecific backcross population. *Genetics* **138**: 1251-74.

Chang, C., S. F. Kwok, A. B. Bleecker and E. M. Meyerowitz (1993). Arabidopsis ethylene-response gene ETR1: similarity of product to two-component regulators. *Science* **262**: 539-544.

Chao, Q. M., M. Rothenberg, R. Solano, G. Roman, W. Terzaghi and J. R. Ecker (1997). Activation of the ethylene gas response pathway in Arabidopsis by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. *Cell* **89**: 1133-1144.

Charmet, G. and G. Branlard (1985). A comparison of androgenetic doubled-haploid and single-seed descent lines in *Triticale. Theoretical Applied Genetics* **71**: 193-200.

Chaussat, R., D. Bouinot and G. Mangeret (1983). Heterogeite de la germination de L'orge dormant, liee a la position du grain sur l'epi. (Heterogeneity of the germination of the dormant barley in relation with the grain position on the ear). *Bios.* **14**: 70-72.

Ching, A., K. S. Caldwell, M. Jung, M. Dolan, O. S. Smith, S. Tingey, M. Morgante and A. J. Rafalski (2002). SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genetics* **3**: 19.

Chiwocha, S. D. S., A. J. Cutler, S. R. Abrams, S. J. Ambrose, J. Yang, A. R. S. Ross and A. R. Kermode (2005). The ETR1-2 mutation in *Arabidopsis thaliana* affects the abscisic acid, auxin, cytokinin and gibberellin metabolic pathways during maintenance of seed dormancy, moist-chilling and germination. *Plant Journal* **42**: 35-48.

Cho, R. J., M. Mindrinos, D. R. Richards, R. J. Sapolsky, M. Anderson, E. Drenkard, J. Dewdney, T. L. Reuber, M. Stammers, N. Federspiel, A. Theologis, W. H. Yang, E. Hubbell, M. Au, E. Y. Chung, D. Lashkari, B. Lemieux, C. Dean, R. J. Lipshutz, F. M. Ausubel, R. W. Davis and P. J. Oefner (1999). Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. *Nature Genetics* **23**: 203-207.

Cho, S. H., D. F. Garvin and G. J. Muehlbauer (2006). Transcriptome analysis and physical mapping of barley genes in wheat-barley chromosome addition lines. *Genetics* **172**: 1277-1285.

Choe, S. W., B. P. Dilkes, S. Fujioka, S. Takatsuto, A. Sakurai and K. A. Feldmann (1998). The DWF4 gene of Arabidopsis encodes a cytochrome P450 that mediates multiple 22 alpha-hydroxylation steps in brassinosteroid biosynthesis. *Plant Cell* **10**: 231-243.

Choe, S.W., B. P. Dilkes, B. D. Gregory, A. S. Ross, H. Yuan, T. Noguchi, S. Fujioka, S. Takatsuto, A. Tanaka, S. Yoshida, F. E. Tax and K. A. Feldmann (1999a). The Arabidopsis dwarf1 mutant is defective in the conversion of 24-methylenecholesterol to campesterol in brassinosteroid biosynthesis. *Plant Physiology* **119**: 897-907.

Choe, S. W., T. Noguchi, S. Fujioka, S. Takatsuto, C. P. Tissier, B. D. Gregory, A. S. Ross, A. Tanaka, S. Yoshida, F. E. Tax and K. A. Feldmann (1999b). The Arabidopsis dwf7/ste1 mutant is defective in the Delta(7) sterol C-5 desaturation step leading to brassinosteroid biosynthesis. *Plant Cell* **11**: 207-221.

Choe, S.W., A. Tanaka, T. Noguchi, S. Fujioka, S. Takatsuto, A. S. Ross, F. E. Tax, S. Yoshida and K. A. Feldmann (2000). Lesions in the sterol delta reductase gene of Arabidopsis cause dwarfism due to a block in brassinosteroid biosynthesis. *Plant Journal* **21**: 431-443.

Clerkx, E. J., M. E. El-Lithy, E. Vierling, G. J. Ruys, H. Blankestijn-De Vries, S. P. Groot, D. Vreugdenhil and M. Koornneef (2004). Analysis of natural allelic variation of Arabidopsis seed germination and seed longevity traits between the accessions *Landsberg erecta* and Shakdara, using a new recombinant inbred line population. *Plant Physiology* **135**: 432-443.

Clouse, S. D. and J. M. Sasse (1998). BRASSINOSTEROIDS: Essential Regulators of Plant Growth and Development. *Annual Review of Plant Physiology and Plant Molecular Biology* **49**: 427-451.

Cohn, M. A. and D. L. Butera (1982). Seed dormancy in red rice (*Orysa sativa*). II. Response to cytokinins. *Weed Science* **30**: 200-205.

Cohn, M. A. (1996). Operational and philosophical decisions in seed dormancy research. *Seed Science* **6**: 147-153.

Colucci, G., F. Apone, N. Alyeshmerni, D. Chalmers and M. J. Chrispeels (2002). GCR1, the putative Arabidopsis G protein-coupled receptor gene is cell cycle-regulated, and its overexpression abolishes seed dormancy and shortens time to flowering. *Proceedings of the National Academy of Science USA* **99**: 4736-4741.

Crow, K. D. and G. P. Wagner (2006). What is the role of genome duplication in the evolution of complexity and diversity? *Molecular Biology and Evolution* **23**: 887-892.

Cutler, S., M. Ghassemian, D. Bonetta, S. Cooney and P. McCourt (1996). A protein farnesyl transferase involved in abscisic acid signal transduction in Arabidopsis. *Science* **273**: 1239-1241.

D'Agostino, I. B. and J. J. Kieber (1999). Molecular mechanisms of cytokinin action. *Current Opinion in Plant Biology* **2**: 359-364.

Davies, R. T., D. H. Goetz, J. Lasswell, M. N. Anderson and B. Bartel (1999). IAR3 encodes an auxin conjugate hydrolase from Arabidopsis. *Plant Cell* **11**: 365-376.

Debeaujon, I., K. M. Leon-Kloosterziel and M. Koornneef (2000). Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. *Plant Physiology* **122**: 403-413.

Debeaujon, I., L. Lepiniec, L. Pourcel and J. Routaboul (2007). Seed coat development and dormancy. Seed development, dormancy and germination. K. Bradford and H. Nonogaki. Oxford, Blackwell Publishing Ltd**:** 25-50.

de Buyser, J., Y. Henry and G. Taleb (1985). Wheat androgenesis: cytogenetical analysis and agronomic performance of doubled haploids. *Z Pflanzenzuecht* **95**: 23-34.

Devos, K. M., J. Beales, Y. Nagamura and T. Sasaki (1999). Arabidopsis-rice: will colinearity allow gene prediction across the eudicot-monocot divide? *Genome Research* **9**: 825-829.

Devos, K. M. and M. D. Gale (2000). Genome relationships: the grass model in current research. *Plant Cell* **12**: 637-646.

Doerge, R. and G. Churchill (1996). Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142**: 285-294.

Doerge, R. W. (2002). Mapping and analysis of quantitative trait loci in experimental populations. *Nature Reviews Genetics* **3**: 43-52.

Dong, Y. J., E. Tsuzuki, H. Kamiunten, H. Terao, D. Z. Lin, M. Matsuo and Y. F. Zheng (2003). Identification of quantitative trait loci associated with pre-harvest sprouting resistance in rice (*Oryza sativa* L.). *Field Crops Research* **81**: 133-139.

Dubreucq, B., N. Berger, E. Vincent, M. Boisson, G. Pelletier, M. Caboche and L. Lepiniec (2000). The Arabidopsis ATEPR1 extensin-like gene is specifically expressed in endosperm during seed germination. *Plant Journal* **23**: 643-652.

Duffus, C. M. and M. P. Cochrane (1992). Grain structure and composition. Barley: genetics, biochemistry, molecular biology and biotechnology. P. R. Shewry. Bristol, CAB International**:** 291-317.

Dufour, P., C. Johnsson, S. Antoine-Michard, R. Cheng, A. Murigneux and M. Beckert (2001). Segregation distortion at marker loci: variation during microspore embryogenesis in maize. *Theoretical Applied Genetics* **102**: 993-1001.

Duque, P. and N. H. Chua (2003). IMB1, a bromodomain protein induced during seed imbibition, regulates ABA- and phyA-mediated responses of germination in Arabidopsis. *Plant Journal* **35**: 787-799.

Edney, M. J. and D. E. Mather (2004). Quantitative trait loci affecting germination traits and malt friability in a two-rowed by six rowed barley cross. *Journal of Cereal Science* **39**: 283-290.

European Brewery Convention (1987). Analytica. Zürich, Switzerland, Brauerei und Getränke Rundschau**:** 271.

Feuillet, C. and B. Keller (2002). Comparative genomics in the grass family: molecular characterization of grass genome structure and evolution. *Annals of Botany (London)* **89**: 3-10.

Feurtado, J. A. and A. R. Kermode (2007). A merging of paths: abscisic acid and hormonal cross-talk in the control of seed dormancy maintenance and alleviation. Seed development, dormancy and germination. K. Bradford and H. Nonogaki. Oxford, Blackwell Publishing Ltd**:** 176-223.

Ficher-Iglesias, C. and G. Neuhaus (2001). Zygotic embryogenesis: hormonal control of embryo development. Current Trends in the Embryology of Angiosperms. S. S. Bhojwani. Dordrecht, Kluwer Academic Publishers**:** 223-247.

Finch-Savage, W. E. and G. Leubner-Metzger (2006). Seed dormancy and the control of germination. *New Phytologist* **171**: 501-523.

Finkelstein, R. R. (1993). Abscisic acid-insensitive mutations provide evidence for stage-specific signal pathways regulating expression of an Arabidopsis late embryogenesis-abundant (LEA) gene. *Molecular and General Genetics* **238**: 401-408.

Finkelstein, R. R., M. L. Wang, T. J. Lynch, S. Rao and H. M. Goodman (1998). The Arabidopsis abscisic acid response locus ABI4 encodes an APETALA2 domain protein. *Plant Cell* **10**: 1043-1054.

Finkelstein, R. R. (2004). The role of hormones during seed development and germination. Plant hormones - Biosynthesis, signal transduction, action! P. J. Davies. Dordrecht, Kluwer Academic**:** 513-537.

Finkelstein, R., W. Reeves, T. Ariizumi and C. Steber (2008). Molecular Aspects of Seed Dormancy. *Annual Review of Plant Biology* **59**: Epub ahead of print.

Flintham, J. and M. D. Gale (1996). Dormancy gene maps in homologous cereal genomes. Preharvest Sprouting in Cereals 1995. K. Noda and D. J. Mares. Osaka, Japan, Centre for Academic Societies**:** 143-149.

Flintham, J., R. Adlam, M. Bassoi, M. Holdsworth and M. Gale (2002). Mapping genes for resistance to sprouting damage in wheat. *Euphytica* **126**: 39-45.

Flintham, J. E., R. Adlam and M. Gale (1999). Seedcoat and embryo dormancy in wheat. Eighth International Symposium on Preharvest Sprouting in Cereals 1998, Detmold, Germany, Association of Cereal Reasearch, Federal Centre for Cereal Potato, and Lipid Research.

Flintham, J. E. (2000). Different genetic components control coat-imposed and embryo-imposed dormancy in wheat. *Seed Science Research* **10**: 43-50.

Foley, M. E. (2001). Seed dormancy: an update on terminology, physiological genetics, and quantitative trait loci regulating germinability. *Weed Science* **49**: 305-317.

Footitt, S., S. P. Slocombe, V. Larner, S. Kurup, Y. S. Wu, T. Larson, I. Graham, A. Baker and M. Holdsworth (2002). Control of germination and lipid mobilization by COMATOSE, the Arabidopsis homologue of human ALDP. *Embo Journal* **21**: 2912-2922.

Fox, G. P., J. F. Panozzo, C.D. Li, R.C.M. Lance, P.A. Inkerman and R. J. Henry (2003). Molecular basis of barley quality. *Australian Journal of Agricultural Research* **54**: 1081-1101.

Francki, M. and R. Appels (2007). Comparative genomics and crop improvement. Comparative genomics: fundamental and applied perspectives. J. R. Brown. New York, CRC Press, Taylor and Francis Publisher**:** 323-342.

Gale, M. D. (1989). The genetics of preharvest sprouting in cereals, particularly in wheat. Preharvest field sprouting in cereals. N. F. Derera. Boca Raton, Florida, CRC Press**:** 85-110.

Gale, M. D. and K. M. Devos (1998). Comparative genetics in the grasses. *Proceedings of the National Academy of Sciences of the USA* **95**: 1971-1974.

Galichet, A. and W. Gruissem (2003). Protein farnesylation in plants-conserved mechanisms but different targets. *Current Opinion in Plant Biology* **6**: 530-535.

Gao, W., J. A. Clancy, F. Han, D. Prada, A. Kleinhofs and S. E. Ullrich (2003). Molecular dissection of a dormancy QTL region near the chromosome 7(5H)L telomere in barley. *Theoretical and Applied Genetics* **107**: 552-559.

Garcia-Dorado, A. and A. Gallego (1992). On the use of the classical tests for detecting linkage. *Journal of Heredity* **83**: 143-146.

Garlinge, J. (2005). The crop variety sowing guide 2005. Department of Agriculture Western Australia, Bulletin. **4655:** 71-97.

Garvin, D. F., J. E. Miller-Garvin, E. A. Viccars, J. V. Jacobsen and A. H. D. Brown (1998). Identification of molecular markers linked to ant28-484, a mutation that eliminates proanthocyanidin production in barley seeds. *Crop Science* **38**: 1250-1255.

Gazzarrini, S., Y. Tsuchiya, S. Lumba, M. Okamoto and P. McCourt (2004). The transcription factor FUSCA3 controls developmental timing in Arabidopsis through the hormones gibberellin and abscisic acid. *Developmental Cell* **7**: 373-385.

Gfeller, F. and F. Svejda (1960). Inheritance of post-harvest seed dormancy and kernal colour in spring wheat lines. *Canadian Journal of Plant Science* **40**: 1-6.

Ghassemian, M., E. Nambara, S. Cutler, H. Kawaide, Y. Kamiya and P. McCourt (2000). Regulation of abscisic acid signaling by the ethylene response pathway in Arabidopsis. *Plant Cell* **12**: 1117-1126.

Gibson, S. I., R. J. Laby and D. G. Kim (2001). The sugar-insensitive1 (sis1) mutant of Arabidopsis is allelic to CTR1. *Biochemical and Biophysical Research Communications* **280**: 196-203.

Giraudat, J., B. M. Hauge, C. Valon, J. Smalle, F. Parcy and H. M. Goodman (1992). Isolation of the Arabidopsis ABI3 gene by positional cloning. *Plant Cell* **4**: 1251-1261.

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

Goldberg, R. B., G. Depaiva and R. Yadegari (1994). Plant Embryogenesis - Zygote to Seed. *Science* **266**: 605-614.

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.

Graner, A., A. Jahoor, J. Schondelmaier, H. Siedler, K. Pillen, G. Fischbeck, G. Wenzel and R. G. Herrmann (1991). Construction of an RFLP map of barley. *Theoretical and Applied Genetics* **83**: 250-256.

Grant, D., P. Cregan and R. C. Shoemaker (2000). Genome organization in dicots: genome duplication in Arabidopsis and synteny between soybean and Arabidopsis. *Proceedings of the National Academy of Science USA* **97**: 4168-4173.

Greenboim-Wainberg, Y., I. Maymon, R. Borochov, J. Alvarez, N. Olszewski, N. Ori, Y. Eshed and D. Weiss (2005). Cross talk between gibberellin and cytokinin: The Arabidopsis GA response inhibitor SPINDLY plays a positive role in cytokinin signaling. *Plant Cell* **17**: 92-102.

Griffin, T. J. and L. M. Smith (2000). Single-nucleotide polymorphism analysis by MALDI-TOF mass spectrometry. *Trends in Biotechnology* **18**: 77-84.

Grill, E. and A. Himmelbach (1998). ABA signal transduction. *Current Opinion in Plant Biology* **1**: 412-418.

Groos, C., G. Gay, M. R. Perretant, L. Gervais, M. Bernard, F. Dedryver and G. Charmet (2002). Study of the relationship between pre-harvest sprouting and grain color by quantitative trait loci analysis in a white x red grain bread-wheat cross. *Theoretical Applied Genetics* **104**: 39-47.

Gu, X., S. F. Kianian and M. E. Foley (2004). Multiple loci and epistases control genetic variation for seed dormancy in weedy rice (*Oryza sativa*). *Genetics* **166**: 1503-1516.

Gu, X. Y., S. F. Kianian, G. A. Hareland, B. L. Hoffer and M. E. Foley (2005). Genetic analysis of adaptive syndromes interrelated with seed dormancy in weedy rice (*Oryza sativa*). *Theoretical Applied Genetics* **110**: 1108-1118.

Gualberti, G., M. Papi, L. Bellucci, L. Ricci, D. Bouchez, C. Camilleri, P. Costantino and P. Vittorioso (2002). Mutations in the Dof zinc finger genes DAG2 and DAG1 influence with opposite effects the germination of Arabidopsis seeds. *Plant Cell* **14**: 1253-1263.

Guberac, V., J. Martincic and S. Maric (1998). Influence of seed size on germinability, germ length, rootlet length and grain yield in spring oat. *Bodenkultur* **49**: 13-18.

Guo, H. W. and J. R. Ecker (2003). Plant responses to ethylene gas are mediated by SCF (EBF1/EBF2)-dependent proteolysis of EIN3 transcription factor. *Cell* **115**: 667-677.

Guyot, R. and B. Keller (2004). Ancestral genome duplication in rice. *Genome* **47**: 610-614.

Hackett, C. A. and L. B. Broadfoot (2003). Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. *Heredity* **90**: 33-38.

Hagberg, S. (1960). A rapid method for determining alpha-amylase activity. *Cereal Chemistry* **47**: 218-222.

Haley, C. S. and S. A. Knott (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**: 315-324.

Halldén, C., a. Hansen, A. Hjerdin and T. Säll (1996). Competition as a source of errors in RAPD analysis. *Theoretical Applied Genetics* **93**: 1185-1192.

Han, F., S. E. Ullrich, J. A. Clancy, V. Jitkov, A. Kilian and I. Romagosa (1996). Verification of barley seed dormancy loci via linked molecular markers. *Theoretical and Applied Genetics* **92**: 87-91.

Han, F., S. E. Ullrich, J. A. Clancy and I. Romagosa (1999). Inheritance and fine mapping of a major barley seed dormancy QTL. *Plant Science* **143**: 113-118.

Haslekas, C., R. A. Stacy, V. Nygaard, F. A. Culianez-Macia and R. B. Aalen (1998). The expression of a peroxiredoxin antioxidant gene, AtPer1, in *Arabidopsis thaliana* is seed-specific and related to dormancy. *Plant Molecular Biology* **36**: 833-845.

Havukkala, I., H. Ichimura, Y. Nagamura and T. Sasaki (1995). Rice genome analysis by integration of sequencing and mapping data. *Journal of Biotechnology* **41**: 139-148.

Hayden, M. J., T. M. Nguyen, A. Waterman, G. L. McMichael and K. J. Chalmers (2008). Application of multiplex-ready PCR for fluorescence-based SSR genotyping in barley and wheat. *Molecular Breeding* **21**: 271-281.

Hazen, S. P., Y. Wu and J. A. Kreps (2003). Gene expression profiling of plant responses to abiotic stress. *Functional Integrated Genomics* **3**: 105-111.

He, Y. H. and S. S. Gan (2004). A novel zinc-finger protein with a proline-rich domain mediates ABA-regulated seed dormancy in Arabidopsis. *Plant Molecular Biology* **54**: 1-9.

Hearnden, P. R., P. J. Eckermann, G. L. McMichael, M. J. Hayden, J. K. Eglinton and K. J. Chalmers (2007). A genetic map of 1,000 SSR and DArT markers in a wide barley cross. *Theoretical Applied Genetics* **115**: 383-391.

Helliwell, C. A., A. Poole, W. J. Peacock and E. S. Dennis (1999). Arabidopsis ent-kaurene oxidase catalyzes three steps of gibberellin biosynthesis. *Plant Physiology* **119**: 507-510.

Hilhorst, H. W. M. and C. M. Karssen (1992). Seed dormancy and germination: the role of abscisic acid and gibberellins and the importance of hormone mutants. *Plant Growth Regulation* **11**: 225-238.

Hilhorst, H. W. M. (2007). Definitions and hypotheses of seed dormancy. Seed Development, Dormancy and Germination, Annual Plant Reviews. K. Bradford and H. Nonogaki. Oxford, Blackwell Publishing Ltd. **27:** 50-71.

Hill, A., A. Nantel, C. D. Rock and R. S. Quatrano (1996). A conserved domain of the viviparous-1 gene product enhances the DNA binding activity of the bZIP

protein EmBP-1 and other transcription factors. *Journal of Biological Chemistry* **271**: 3366-3374.

Himi, E., D. J. Mares, A. Yanagisawa and K. Noda (2002). Effect of grain colour gene (R) on grain dormancy and sensitivity of the embryo to abscisic acid (ABA) in wheat. *Journal of Experimental Botany* **53**: 1569-1574.

Hisamatsu, T., R. W. King, C. A. Helliwell and M. Koshioka (2005). The involvement of gibberellin 20-oxidase genes in phytochrome-regulated petiole elongation of Arabidopsis. *Plant Physiology* **138**: 1106-1116.

Hobbie, L., M. McGovern, L. R. Hurwitz, A. Pierro, N. Y. Liu, A. Bandyopadhyay and M. Estelle (2000). The AXR6 mutants of *Arabidopsis thaliana* define a gene involved in auxin response and early development. *Development* **127**: 23-32.

Hoffmann, F., E. Thomas and G. Wenzel (1982). Anther culture as a breeding tool in rape. 2 Progeny analyses of androgenetic lines and induced mutants from haploid cultures. *Theoretical Applied Genetics* **61**: 225-232.

Hong, S. W., J. H. Jon, J. M. Kwak and H. G. Nam (1997). Identification of a receptor-like protein kinase gene rapidly induced by abscisic acid, dehydration, high salt, and cold treatments in *Arabidopsis thaliana*. *Plant Physiology* **113**: 1203-1212.

Hoogendoorn, J. (1985). The physiology of variation in the time of ear emergence among wheat varieties from different regions of the world. *Euphytica* **34**: 559-571.

Hooley, R. (1998). Auxin signaling. Homing in with targeted genetics. *Plant Cell* **10**: 1581-1584.

Hoyle, F. (1999). A review on pre-harvest sprouting in cereals. Western Australia, Field Crops Research Group, Agriculture Western Australia.

Hozo, S. P., B. Djulbegovic and I. Hozo (2005). Estimating the mean and variance from the median, range, and the size of the sample. *BMC Medical Research Methodology*.

Hua, J., C. Chang, Q. Sun and E. M. Meyerowitz (1995). Ethylene insensitivity conferred by Arabidopsis ERS gene. *Science* **269**: 1712-1714.

Hua, J. and E. M. Meyerowitz (1998). Ethylene responses are negatively regulated by a receptor gene family in *Arabidopsis thaliana*. *Cell* **94**: 261-271.

Hugouvieux, V., J. M. Kwak and J. I. Schroeder (2001). An mRNA cap binding protein, ABH1, modulates early abscisic acid signal transduction in Arabidopsis. *Cell* **106**: 477-487.

Hurles, M. (2004). Gene duplication: the genomic trade in spare parts. *PLoS Biol* **2**: E206 (online journal).

Islam, A. K. M. R., K. W. Shepherd and D. H. B. Sparrow (1975). Addition of individual barley chromosomes to wheat. Proceedings of the 3rd International Barely Genetics Symposium, Berlin, Verlag Karl Thiemig.

Islam, A. K. M. R., K. W. Shepherd and D. H. B. Sparrow (1981). Isolation and characterization of euplasmic wheat-barley chromosome addition lines. *Heredity* **46**: 161-174.

Islam, A. K. M. R. (1983). Ditelosomic additions of barley chromosomes to wheat. Proceedings of the 6th International Wheat Genetics Symposium, Kyoto, Japan.

Islam, A. K. M. R. and K. W. Shepherd (1990). Incorporation of barley chromosomes in wheat. Wheat (Biotechnology in Agriculture and Forestry). Y. P. S. Bajaj. Berlin, Springer-Verlag. **13:** 128-151.

Islam, A. K. M. R. and K. W. Shepherd (2000). Isolation of a fertile wheat-barley addition line carrying the entire barley chromosome 1H. *Euphytica* **111**: 145-149.

Izawa, T. and K. Shimamoto (1996). Becoming a model plant: the importance of rice to plant science. *Trends in Plant Science* **1**: 95-99.

Jaccoud, D., K. Peng, D. Feinstein and A. Kilian (2001). Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Research* **29**: E25 (online journal).

Jacobsen, S. E., K. A. Binkowski and N. E. Olszewski (1996). SPINDLY, a tetratricopeptide repeat protein involved in gibberellin signal transduction Arabidopsis. *Proceedings of the National Academy of Sciences of the USA* **93**: 9292-9296.

Jacobsen, S. E., N. E. Olszewski and E. M. Meyerowitz (1998). SPINDLY's role in the gibberellin response pathway. *Symposia of the Society for Experimental Biology* **51**: 73-78.

Jansen, R. C. (1993). Interval mapping of multiple quantitative trait loci. *Genetics* **135**: 205-211.

Jiang, C. X., P. W. Chee, X. Draye, P. L. Morrell, C. W. Smith and A. H. Paterson (2000). Multilocus interactions restrict gene introgression in interspecific populations of polyploid *Gossypium* (cotton). *Evolution: International Journal of Organic Evolution* **54**: 798-814.

Jin, W. W., J. R. Melo, K. Nagaki, P. B. Talbert, S. Henikoff, R. K. Dawe and J. M. Jiang (2004). Maize centromeres: organization and functional adaptation in the genetic background of oat. *Plant Cell* **16**: 571-581.

Jones, A. M., K. H. Im, M. A. Savka, M. J. Wu, N. G. DeWitt, R. Shillito and A. N. Binns (1998). Auxin-dependent cell expansion mediated by overexpressed auxin-binding protein 1. *Science* **282**: 1114-1117.

Jones, H. D., N. C. Peters and M. J. Holdsworth (1997). Genotype and environment interact to control dormancy and differential expression of the VIVIPAROUS 1 homologue in embryos of *Avena fatua*. *Plant Journal* **12**: 911-920.

Kao, C. H., Z. B. Zeng and R. D. Teasdale (1999). Multiple interval mapping for quantitative trait loci. *Genetics* **152**: 1203-1216.

Karakousis, A., J. P. Gustafson, K. J. Chalmers, B. A.R. and P. Langridge (2003). A consensus map of barley integrating SSR, RFLP, and AFLP markers. *Australian Journal of Agricultural Research* **54**: 1173-1185.

Karssen, C. M. (1982). Seasonal patterns of dormancy in weed seeds. The physiology and biochemistry of seed development, dormancy and germination. A. A. Khan. New York, Elsevier Biomedical Press**:** 244-270.

Karssen, C. M., D. L. C. Brinkhorst-van der Swan, A. E. Breekland and M. Koornneef (1983). Induction of dormancy during seed development by endogenous abscisic acid: studies on abscisic acid deficient genotypes of *Arabidopsis thaliana* (L.) Heynh. *Planta*: 158-165.

Karssen, C. M., S. Zargorsky, J. Kepczynski and S. P. C. Groot (1989). Key role for endogenous gibberellins in the control of seed germination. *Annals of Botany* **63**: 71-80.

Kato, K., W. Nakamura, T. Tabiki, H. Miura and S. Sawada (2001). Detection of loci controlling seed dormancy on group 4 chromosomes of wheat and comparative mapping with rice and barley genomes. *Theoretical Applied Genetics* **102**: 980-985.

Kawade, R. M., S. D. Ugale and R. B. Patil (1987). Effect of seed size on germination, seedling vigor and test weight of pearl millet. *Seed Research* **15**: 210-213.

Keller, B. and C. Feuillet (2000). Colinearity and gene density in grass genomes. *Trends in Plant Science* **5**: 246-51.

Kieber, J. J., M. Rothenberg, G. Roman, K. A. Feldmann and J. R. Ecker (1993). CTR1, a negative regulator of the ethylene response pathway in Arabidopsis, encodes a member of the RAF family of protein-kinases. *Cell* **72**: 427-441.

Kim, Y. C., M. Nakajima, A. Nakayama and I. Yamaguchi (2005). Contribution of gibberellins to the formation of Arabidopsis seed coat through starch degradation. *Plant and Cell Physiology* **46**: 1317-1325.

King, R. W. and R. A. Richards (1984). Water uptake and preharvest sprouting damage in wheat: ear characteristics. *Australian Journal of Agricultural Research* **35**: 327-336.

King, R. W. (1989). Physiology of sprouting resistance. Pre-harvest sprouting in cereals. N. F. Derera. Boca Raton, Florida, CRC Press**:** 27-60.

King, R. W. and P. Wettstein-Knowles (2000). Epicuticular waxes and regulation of ear wetting and pre-harvest sprouting in barley and wheat. *Euphytica* **112**: 157-166.

Kishimoto, N., H. Higo, K. Abe, S. Arai, A. Saito and K. Higo (1994). Identification of duplicated segments in rice chromosomes 1 and 5 by linkage analysis of cDNA markers of known functions. *Theoretical Applied Genetics* **88**: 722-726.

Klahre, U., T. Noguchi, S. Fujioka, S. Takatsuto, T. Yokota, T. Nomura, S. Yoshida and N. H. Chua (1998). The Arabidopsis DIMINUTO/DWARF1 gene encodes a protein involved in steroid synthesis. *Plant Cell* **10**: 1677-1690.

Kleinhofs, K., A. Kilian and D. Kudrna (1993). The NABGMP Steptoe x Morex mapping progress report. *BGN* **23**: 79-83.

Koornneef, M. and J. H. van der Veen (1980). Induction and analysis of gibberellin-insensitive mutants in *Arabidopsis thaliana* (L.) Heynh. *Theoretical Applied Genetics* **58**: 257-263.

Koornneef, M. and C. M. Karssen (1994). Seed dormancy and germination. Arabidopsis. E. M. Meyerowitz. New York, USA, Cold Spring Harbor Laboratory Press**:** 313-334.

Koornneef, M., L. Bentsink and H. W. M. Hilhorst (2002). Seed dormancy and germination. *Current Opinion in Plant Biology* **5**: 33-36.

Koornneef, M., C. Alonso-Blanco and D. Vreugdenhil (2004). Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annual Review of Plant Biology* **55**: 141-172.

Kosambi, D. D. (1944). The estimation of map distance from recombination values. *Annals of Eugenics* **12**: 172-175.

Ku, H. M., T. Vision, J. Liu and S. D. Tanksley (2000). Comparing sequenced segments of the tomato and Arabidopsis genomes: large-scale duplication followed by selective gene loss creates a network of synteny. *Proceedings of the National Academy of Science USA* **97**: 9121-9126.

Kucera, B., M. A. Cohn and G. Leubner-Metzger (2005). Plant hormone interactions during seed dormancy release. *Seed Science Research* **15**: 281-307.

Kulwal, P. L., R. Singh, H. S. Balyan and P. K. Gupta (2004). Genetic basis of pre-harvest sprouting tolerance using single-locus and two-locus QTL analyses in bread wheat. *Functional Integrated Genomics* **4**: 94-101.

Kumashiro, T. and T. Oinuma (1985). Comparison of genetic variability among anther-derived and ovule-derived doubled haploid lines of tobacco. *Japanese Journal of Breeding* **35**: 301-310.

Kunihisa, M., N. Fukino and S. Matsumoto (2003). Development of cleavage amplified polymorphic sequence (CAPS) markers for identification of strawberry cultivars. *Euphytica* **134**: 209-215.

Kurata, N., G. Moore, Y. Nagumara, T. Foote, M. Yano, Y. Minobe and M. Gale (1994). Conservation of genome structure between rice and wheat. *Biotechnology* **12**: 276-278.

Kurata, N., Y. Nagamura, K. Yamamoto, Y. Harushima, N. Sue, J. Wu, B. A. Antonio, A. Shomura, T. Shimizu, S-Y. Lin, T. Inoue, A. Fukada, T. Shimano, Y. Kuboki, T. Toyama, Y. Miyamoto, T. Kirihara, K. Hayasaka, A. Miyao, L. Monna, H.S. Zhong, Y. Tamura, Z-X. Wang, T. Momma, Y. Umehara, M. Yano, T. Sasaki and Y. Minobe (1994). A 300 kilobase interval genetic map of rice including 883 expressed sequences. *Nature Genetics* **8**: 365-372.

La Rota, M. and M. Sorrells (2004). Comparative DNA sequence analysis of mapped wheat ESTs reveals the complexity of genome relationships between rice and wheat. *Functional Integrated Genomics* **4**: 34-46.

Laford, G. P. and R. J. Baker (1986). Effects of temperature, moisture stress and seed size on germination of nin spring wheat cultivars. *Crop Science* **26**: 563-567.

Lander, E. S. and D. Botstein (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185-199.

Langridge, P., E. S. Lagudah, T. A. Holton, R. Appels, P. J. Sharp and K. J. Chalmers (2001). Trends in genetic and genome analyses in wheat: a review. *Australian Journal of Agricultural Research* **52**: 1043-1077.

Larsen, S. U. and C. Andreasen (2004). Light and heavy seeds differ in germination percentage and mean germination thermal time. *Crop Science* **44**: 1710-1720.

Larson, S., G. Bryan, W. Dyer and T. Blake (1996). Evaluating gene effects of a major barley seed dormancy QTL in reciprocal back-cross populations. *Journal of Quantitative Trait Loci*.

Lashermes, P., A. Gaillard and M. Beckert (1988). Gynogenetic haploid plants analysis for agronomic and enzymatic markers in maize (*Zea mays* L.). *Theoretical Applied Genetics* **76**: 570-572.

Lashermes, P., M. C. Combes, N. S. Prakash, P. Trouslot, M. Lorieux and A. Charrier (2001). Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. *Genome* **44**: 589-596.

Lasswell, J., L. E. Rogg, D. C. Nelson, C. Rongey and B. Bartel (2000). Cloning and characterization of IAR1, a gene required for auxin conjugate sensitivity in Arabidopsis. *Plant Cell* **12**: 2395-2408.

Laurie, D. A. and K. M. Devos (2002). Trends in comparative genetics and their potential impacts on wheat and barley research. *Plant Molecular Biology* **48**: 729-740.

LeClere, S., R. Tellez, R. A. Rampey, S. P. T. Matsuda and B. Bartel (2002). Characterization of a family of IAA-amino acid conjugate hydrolases from Arabidopsis. *Journal of Biological Chemistry* **277**: 20446-20452.

Lee, S. C., H. Cheng, K. E. King, W. F. Wang, Y. W. He, A. Hussain, J. Lo, N. P. Harberd and J. R. Peng (2002). Gibberellin regulates Arabidopsis seed

germination via RGL2, a GAI/RGA-like gene whose expression is up-regulated following imbibition. *Genes & Development* **16**: 646-658.

Lehmensiek, A., W. Bovill, P. Wenzl, P. Langridge and R. Appels (2008). Genetic mapping in the Triticeae. Triticeae Genomics. C. Feuillet and G. Muehlbaur. Heidelberg, London, New York, Paris, Springer, *in press*.

Leon-Kloosterziel, K. M., M. A. Gil, G. J. Ruijs, S. E. Jacobsen, N. E. Olszewski, S. H. Schwartz, J. A. D. Zeevaart and M. Koornneef (1996a). Isolation and characterization of abscisic acid-deficient Arabidopsis mutants at two new loci. *Plant Journal* **10**: 655-661.

Leon-Kloosterziel, K. M., G. A. van de Bunt, J. A. Zeevaart and M. Koornneef (1996b). Arabidopsis mutants with a reduced seed dormancy. *Plant Physiology* **110**: 233-240.

Leung, J., S. Merlot and J. Giraudat (1997). The Arabidopsis ABSCISIC ACID-INSENSITIVE2 (ABI2) and ABI1 genes encode homologous protein phosphatases 2C involved in abscisic acid signal transduction. *Plant Cell* **9**: 759-771.

Leung, J. and J. Giraudat (1998). Abscisic acid signal transduction. *Plant Molecular Biology* **49**: 199-222.

Li, B. and M. E. Foley (1997). Genetic and molecular control of seed dormancy. *Trends in Plant Science Reviews* **2**: 384-389.

Li, C. D., A. Tarr, R. C. M. Lance, S. Harasymow, J. Uhlmann, S. Westcot, K. J. Young, C. R. Grime, M. Cakir, S. Broughton and R. Appels (2003a). A major QTL controlling seed dormancy and pre-harvest sprouting/grain -amylase in two-rowed barley (*Hordeum vulgare* L.). *Australian Journal of Agricultural Research* **54**: 1303-1313.

Li, C.D., P. Ni, M. Francki, A. Hunter, Y. Zhang, D. Schibeci, H. Li, A. Tarr, J. Wang, M. Cakir, J. Yu, M. Bellgard, R. Lance and R. Appels (2004). Genes controlling seed dormancy and pre-harvest sprouting in a rice-wheat-barley comparison. *Functional Integrated Genomics* **4**: 84-93.

Li, H. C., K. Chuang, J. T. Henderson, S. D. Rider, Jr., Y. Bai, H. Zhang, M. Fountain, J. Gerber and J. Ogas (2005). PICKLE acts during germination to repress expression of embryonic traits. *Plant Journal* **44**: 1010-1022.

Li, J., K. H. Nam, D. Vafeados and J. Chory (2001). BIN2, a new brassinosteroid-insensitive locus in Arabidopsis. *Plant Physiology* **127**: 14-22.

Lin, S. Y., T. Sasaki and M. Yano (1998). Mapping quantitative trait loci controlling seed dormancy and heading date in rice, *Oryza sativa* L., using backcross inbred lines. *Theoretical and Applied Genetics* **96**: 997-1003.

Lin, R., R.D. Horsley, N.L.V. Lapitan, Z. Ma and P.B. Schwarz (2008). QTL mapping of dormancy in barley using the Harrington/Morex and the Chevron/Stander mapping populations. *Crop Science*, in press.

Liscum, E. and J. W. Reed (2002). Genetics of Aux/IAA and ARF action in plant growth and development. *Plant Molecular Biology* **49**: 387-400.

Liu, H., R. Sachidanandam and L. Stein (2001). Comparative genomics between rice and Arabidopsis shows scant collinearity in gene order. *Genome Research* **11**: 2020-2026.

Lopez-Molina, L., S. Mongrand and N. H. Chua (2001). A postgermination developmental arrest checkpoint is mediated by abscisic acid and requires the AB15 transcription factor in Arabidopsis. *Proceedings of the National Academy of Sciences of the USA* **98**: 4782-4787.

Lorieux, M., B. Goffinet, X. Perrier, D. G. Deleon and C. Lanaud (1995a). Maximum-likelihood models for mapping genetic-markers showing segregation distortion .1. Backcross populations. *Theoretical and Applied Genetics* **90**: 73-80.

Lorieux, M., X. Perrier, B. Goffinet, C. Lanaud and D. G. Deleon (1995b). Maximum-likelihood models for mapping genetic-markers showing segregation distortion .2. F2 populations. *Theoretical and Applied Genetics* **90**: 81-89.

Lotan, T., M. Ohto, K. M. Yee, M. A. L. West, R. Lo, R. W. Kwong, K. Yamagishi, R. L. Fischer, R. B. Goldberg and J. J. Harada (1998). Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* **93**: 1195-1205.

Lu, C. and N. Fedoroff (2000). A mutation in the Arabidopsis HYL1 gene encoding a dsRNA binding protein affects responses to abscisic acid, auxin, and cytokinin. *Plant Cell* **12**: 2351-2365.

Lu, H., J. Romero-Severson and R. Bernardo (2002). Chromosomal regions associated with segregation distortion in maize. *Theoretical Applied Genetics* **105**: 622-628.

Lynch, M. and B. Walsh (1998). Genetics and Analysis of Quantitative Traits. MA, USA, Sinauer Associates Incorporated.

Lynch, M. and J. S. Conery (2000). The evolutionary fate and consequences of duplicate genes. *Science* **290**: 1151-1155.

Maddison, W. P. and D. R. Maddison (1992). MacClade version 3. Sunderland, Massachusetts, Sinauer Associates.

Major, D. J. (1980). Photoperiod response characteristics controlling flowering of nine crop species. *Canadian Journal of Plant Science* **60**: 777-784.

Manninen, O. M. (2000). Associations between anther-culture response and molecular markers on chromosomes 2H, 3H and 4H of barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* **100**: 57-62.

Mares, D. (1999). The seed coat and dormancy in wheat grains. Eight International Symposium on Preharvest Sprouting in Cereals 1998, Detmold, Germany,

Association of Cereal Research Federal Centre for Cereal Potato and Lipid Research.

Mares, D., K. Mrva, J. Cheong, K. Williams, B. Watson, E. Storlie, M. Sutherland and Y. Zou (2005). A QTL located on chromosome 4A associated with dormancy in white- and red-grained wheats of diverse origin. *Theoretical Applied Genetics* **111**: 1357-1364.

Mares, D. J. (1989). Preharvest sprouting damage and sprouting tolerance: assay methods and instrumentation. Preharvest field sprouting in cereals. N. F. Derera. Boca Raton, Florida, CRC Press**:** 130-166.

Mares, D. J. (1993a). Genetic studies of sprouting tolerance in red and white wheats. Preharvest sprouting in cereals 1992. M. K. Walker-Simmons and J. Reid. St Paul, NM, AACC**:** 21-29.

Mares, D. J. (1993b). Preharvest sprouting in wheat. 1. Influence of cultivar, rainfall and temperature during grain ripening. *Australian Journal of Agricultural Research* **44**: 1259-1272.

Mares, D. J. and K. Mrva (2001). Mapping quantitative trait loci associated with variation in grain dormancy in Australian wheat. *Australian Journal of Agricultural Research* **52**: 1257-1265.

Marquez-Cedillo, L. A., P. M. Hayes, B. L. Jones, A. Kleinhofs, W. G. Legge, B. G. Rossnagel, K. Sato, S. E. Ullrich and D. M. Wesenberg (2000). QTL analysis of malting quality in barley based on the doubled-haploid progeny of two elite North American varieties representing different germplasm groups. *Theoretical Applied Genetics* **101**: 173-184.

Martinez, O. and R. N. Curnow (1992). Estimating the locations and the sizes of the effects of Quantitative Trait Loci using flanking markers. *Theoretical and Applied Genetics* **85**(4): 480-488.

Massingham, T., L. J. Davies and P. Lio (2001). Analysing gene function after duplication. *Bioessays* **23**: 873-876.

Masuzaki, S., S. Yaguchi, N. Yamauchi and M. Shigyo (2007). Morphological characterisation of multiple alien addition lines of Allium reveals the chromosomal locations of gene(s) related to bulb formation in *Allium cepa* L. *Journal of Horticultural Science and Biotechnology* **82**: 393-396.

Mather, D. E., N. A. Tinker, D. E. LaBerge, M. Edney, B. L. Jones, B. G. Rossnagel, W. G. Legge, K. G. Briggs, R. B. Irvine, D. E. Falk and K. J. Kasha (1997). Regions of genome that affect grain and malt quality in a North American two-row barley cross. *Crop Science* **37**: 544-544.

Mathur, P. N., N. C. Sinha and R. P. Singh (1982). Effect of seed size on germination and seed vigour in oat (*Avena sativa* L.). *Seed Research* **10**: 109-113.

Mauricio, R. (2001). Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. *Nature Reviews Genetics* **2**: 370-381.

Mayer, U., R. Torres Ruiz, T. Berleth, S. Misera and G. Jürgens (1991). Mutations affecting body organization in the *Arabidopsis* embryo. *Nature* **353**: 402-407.

Mazzella, M. A., M. V. Arana, R. J. Staneloni, S. Perelman, M. J. R. Batiller, J. Muschietti, P. D. Cerdan, K. H. Chen, R. A. Sanchez, T. Zhu, J. Chory and J. J. Casal (2005). Phytochrome control of the Arabidopsis transcriptome anticipates seedling exposure to light. *Plant Cell* **17**: 2507-2516.

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.

McCarty, D. R. (1995). Genetic control and integration of maturation and germination pathways in seed development. *Annual Review of Plant Physiology and Plant Molecular Biology* **46**: 71-93.

McKibbin, R. S., M. D. Wilkinson, P. C. Bailey, J. E. Flintham, L. M. Andrew, P. A. Lazzeri, M. D. Gale, J. R. Lenton and M. J. Holdsworth (2002). Transcripts of Vp1 homeologues are misspliced in modern wheat and ancestral species. *Proceedings of the National Academy of Science USA* **99**: 10203-10208.

Meinke, D. W., L. H. Franzmann, T. C. Nickle and E. C. Yeung (1994). Leafy cotyledon mutants of Arabidopsis. *Plant Cell* **6**: 1049-1064.

Melzer, J. M., A. Kleinhofs, D. A. Kudrna, R. L. Warner and T. K. Blake (1988). Genetic-mapping of the barley nitrate reductase-deficient Nar-1 and Nar-2 Loci. *Theoretical and Applied Genetics* **75**: 767-771.

Merlot, S., F. Gosti, D. Guerrier, A. Vavasseur and J. Giraudat (2001). The ABI1 and ABI2 protein phosphatases 2C act in a negative feedback regulatory loop of the abscisic acid signalling pathway. *Plant Journal* **25**: 295-303.

Miura, H., N. Sato, K. Sato and Y. Amano (2002a). Detection of chromosomes carrying genes for seed dormancy of wheat using backcross reciprocal monosomic method. *Plant Breeding* **121**: 394-399.

Miura, K., Y. Lin, M. Yano and T. Nagamine (2002b). Mapping quantitative trait loci controlling seed longevity in rice (*Oryza sativa* L.). *Theoretical Applied Genetics* **104**: 981-986.

Miura, Y., M. Hirata and M. Fujimori (2007). Mapping of EST-derived CAPS markers in Italian ryegrass (*Lolium multiflorum* Lam.). *Plant Breeding* **126**: 353-360.

Miyamoto, T. and E. H. Everson (1958). Biochemical and physiological studies of wheat grain pigmentation. *Agronomy Journal* **50**: 733-734.

Moore, G., K. M. Devos, Z. Wang and M. D. Gale (1995). Cereal genome evolution. Grasses, line up and form a circle. *Current Biology* **5**: 737-739.

Mora-Garcia, S., G. Vert, Y. H. Yin, A. Cano-Delgado, H. Cheong and J. Chory (2004). Nuclear protein phosphatases with Kelch-repeat domains modulate the response to brassinosteroids in Arabidopsis. *Genes & Development* **18**: 448-460.

Mori, M., N. Uchino, M. Chono, K. Kato and H. Miura (2005). Mapping QTLs for grain dormancy on wheat chromosome 3A and the group 4 chromosomes, and their combined effect. *Theoretical Applied Genetics* **110**: 1315-1323.

Muehlbauer, G. J., O. Riera-Lizarazu, R. G. Kynast, D. Martin, R. L. Phillips and H. W. Rines (2000). A maize chromosome 3 addition line of oat exhibits expression of the maize homeobox gene liguleless3 and alteration of cell fates. *Genome* **43**: 1055-1064.

Muir, C. E. and R. A. Nilan (1973). Registration of Steptoe barley. *Crop Science* **13**: 770.

Mulligan, R. M., J. Chory and J. R. Ecker (1997). Signaling in plants. *Proceedings of the National Academy of Science USA* **94**: 2793-2795.

Nagpal, P., L. M. Walker, J. C. Young, A. Sonawala, C. Timpte, M. Estelle and J. W. Reed (2000). AXR2 encodes a member of the Aux/IAA protein family. *Plant Physiology* **123**: 563-573.

Nakamura, S. and T. Toyama (2001). Isolation of a VP1 homologue from wheat and analysis of its expression in embryos of dormant and non-dormant cultivars. *Journal of Experimental Botany* **52**: 875-876.

Nakamura, S., T. Komatsuda and H. Miura (2007). Mapping diploid wheat homologues of Arabidopsis seed ABA signaling genes and QTLs for seed dormancy. *Theoretical Applied Genetics* **114**: 1129-1139.

Nambara, E., S. Naito and P. Mccourt (1992). A mutant of Arabidopsis which is defective in seed development and storage protein accumulation is a new ABI3 allele. *Plant Journal* **2**: 435-441.

Nambara, E., K. Keith, P. McCourt and S. Naito (1994). Isolation of an internal deletion mutant of the *Arabidopsis thaliana* ABI3 gene. *Plant Cell Physiology* **35**: 509-513.

Nambara, E., K. Keith, P. Mccourt and S. Naito (1995). A regulatory role for the ABI3 gene in the establishment of embryo maturation in *Arabidopsis thaliana*. *Development* **121**: 629-636.

Nambara, E., R. Hayama, Y. Tsuchiya, M. Nishimura, H. Kawaide, Y. Kamiya and S. Naito (2000). The role of ABI3 and FUS3 loci in *Arabidopsis thaliana* on phase transition from late embryo development to germination. *Developmental Biology* **220**: 412-423.

Nambara, E. and A. Marion-Poll (2003). ABA action and interactions in seeds. *Trends in Plant Science* **8**: 195-244.

Nambara, E. and A. Marion-Poll (2005). Abscisic acid biosynthesis and catabolism. *Annual Review of Plant Biology* **56**: 165-185.

Nielson, M. T., A. J. McCrate, E. G. Heyne and G. M. Paulsen (1984). Effect of weather variables on preharvest sprouting of hard white winter wheat. *Crop Science* **24**: 779-782.

Nikolaeva, M. G. (1977). Factors controlling the seed dormancy pattern. The physiology and biochemistry of seed dormancy and germination. A. A. Khan. Amsterdam**:** 51-74.

Nikolaeva, M. G. (2004). On criteria to use in studies of seed evolution. *Seed Science Research* **14**: 315-320.

Noda, K., T. Matsuura, M. Maekawa and S. Teketa (2002). Chromosomes responsible for sensitivity of embryo to abscisic acid and dormancy in wheat. *Euphytica* **123**: 203-209.

Nyachiro, J. M., F. R. Clarke, R. M. DePauw, R. E. Knox and K. C. Armstrong (2002). Temperature effects on seed germination and expression of seed dormancy in wheat. *Euphytica* **126**: 123-127.

O'Malley, R. C., F. I. Rodriguez, J. J. Esch, B. M. Binder, P. O'Donnell, H. J. Klee and A. B. Bleecker (2005). Ethylene-binding activity, gene expression levels, and receptor system output for ethylene receptor family members from Arabidopsis and tomato. *Plant Journal* **41**: 651-659.

Oberthur, L., W. Dyer, S. E. Ullrich and T. K. Blake (1995). Genetic analysis of seed dormancy in barley (*Hordeum vulgare* L.). *Journal of Quantitative Trait Loci*.

Ogas, J., J. C. Cheng, Z. R. Sung and C. Somerville (1997). Cellular differentiation regulated by gibberellin in the *Arabidopsis thaliana* pickle mutant. *Science* **277**: 91-94.

Ogawa, M., A. Hanada, Y. Yamauchi, A. Kuwalhara, Y. Kamiya and S. Yamaguchi (2003). Gibberellin biosynthesis and response during Arabidopsis seed germination. *Plant Cell* **15**: 1591-1604.

Ogbonnaya, F.C., M. Imtiaz, G. Ye, P.R. Hearnden, E. Hernandez, R.F. Eastwood, M. van Ginkel, S.C. Shorter and J.M. Winchester (2008). Genetic and QTL analyses of seed dormancy and preharvest sprouting resistance in the wheat germplasm CN10955. *Theoretical and Applied Genetics* **116**: 891-902.

Ohta, M., Y. Guo, U. Halfter and J. K. Zhu (2003). A novel domain in the protein kinase SOS2 mediates interaction with the protein phosphatase 2C ABI2. *Proceedings of the National Academy of Science USA* **100**: 11771-11776.

Osa, M., K. Kato, M. Mori, C. Shindo, A. Torada and H. Miura (2003). Mapping QTLs for seed dormancy and the Vp1 homologue on chromosome 3A in wheat. *Theoretical Applied Genetics* **106**: 1491-1496.

Osakabe, Y., K. Maruyama, M. Seki, M. Satou, K. Shinozaki and K. Yamaguchi-Shinozaki (2005). Leucine-rich repeat receptor-like kinase1 is a key membrane-bound regulator of abscisic acid early signaling in Arabidopsis. *Plant Cell* **17**: 1105-1119.

Osakabe, Y., S. Mizuno, K. Maruyama, K. Osakabe, K. Shinozaki and K. Shinozaki (2007). Functional analysis of a leucine-rich repeat receptor like kinase, RPK1, involved in ABA signal transduction of arabidopsis. *Plant and Cell Physiology* **48**: S158-S158.

Ouellet, F., P. J. Overvoorde and A. Theologis (2001). IAA17/AXR3: Biochemical insight into an auxin mutant phenotype. *Plant Cell* **13**: 829-841.

Papi, M., S. Sabatini, D. Bouchez, C. Camilleri, P. Costantino and P. Vittorioso (2000). Identification and disruption of an Arabidopsis zinc finger gene controlling seed germination. *Genes and Development* **14**: 28-33.

Papi, M., S. Sabatini, M. M. Altamura, L. Hennig, E. Schafer, P. Costantino and P. Vittorioso (2002). Inactivation of the phloem-specific dof zinc finger gene DAG1 affects response to light and integrity of the testa of Arabidopsis seeds. *Plant Physiology* **128**: 411-417.

Parcy, F., C. Valon, A. Kohara, S. Misera and J. Giraudat (1997). The ABSCISIC ACID-INSENSITIVE3, FUSCA3, and LEAFY COTYLEDON1 loci act in concert to control multiple aspects of Arabidopsis seed development. *Plant Cell* **9**: 1265-1277.

Paterson, A. H., J. E. Bowers and B. A. Chapman (2004). Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proceedings of the National Academy of Sciences of the USA* **101**: 9903-9908.

Penfield, S., E. M. Josse, R. Kannangara, A. D. Gilday, K. J. Halliday and I. A. Graham (2005). Cold and light control seed germination through the bHLH transcription factor SPATULA. *Current Biology* **15**: 1998-2006.

Peng, J., D. E. Richards, N. M. Hartley, G. P. Murphy, K. M. Devos, J. Flintham, J. Beales, L. J. Fish, A. J. Worland, F. Pelica, D. Sudhakar, P. Christou, J. W. Snape, M. D. Gale and N. P. Harberd (1999). 'Green Revolution' genes encode mutant giberellin response modulators. *Nature* **400**: 256-261.

Peng, J. R., D. E. Richards, T. Moritz, H. Ezura, P. Carol and N. P. Harberd (2002). Molecular and physiological characterization of Arabidopsis GAI alleles obtained in targeted Ds-tagging experiments. *Planta* **214**: 591-596.

Pickett, F. B., A. K. Wilson and M. Estelle (1990). The AUX1 mutation of Arabidopsis confers both auxin and ethylene resistance. *Plant Physiology* **94**: 1462-1466.

Pillen, K., A. Zacharias and J. Leon (2003). Advanced backcross QTL analysis in barley (*Hordeum vulgare* L.). *Theoretical Applied Genetics* **107**: 340-352.

Potokina, E., M. Prasad, L. Malysheva, M. S. Roder and A. Graner (2006). Expression genetics and haplotype analysis reveal cis regulation of serine carboxypeptidase I (Cxp1), a candidate gene for malting quality in barley (*Hordeum vulgare* L.). *Functional Integrated Genomics* **6**: 25-35.

Prada, D., S. E. Ullrich, J. L. Molina-Cano, L. Cistue, J. A. Clancy and I. Romagosa (2004). Genetic control of dormancy in a Triumph/Morex cross in barley. *Theoretical Applied Genetics* **109**: 62-70.

Ramaih, S., M. Guedira and G. M. Paulsen (2003). Relationship of indoleacetic acid and tryptophan to dormancy and preharvest sprouting of wheat. *Functional Plant Biology* **30**: 939-945.

Rasmusson, D. C. and R. D. Wilcoxson (1979). Registration of Morex barley. *Crop Science* **19**: 293.

Riefler, M., O. Novak, M. Strnad and T. Schmulling (2006). Arabidopsis cytokinin receptor mutants reveal functions in shoot growth, leaf senescence, seed size, germination, root development, and cytokinin metabolism. *Plant Cell* **18**: 40-54.

Ringlund, K. (1980). Starch quality in wheat and barley at different maturity stages in relation to seed dormancy. *Cereal Research Communications* **8**: 193.

Rock, C. (2000). Pathways to abscisic acid-regulated gene expression. *New Phytologist* **148**: 357-396.

Rodriguez, F. I., J. J. Esch, A. E. Hall, B. M. Binder, G. E. Schaller and A. B. Bleecker (1999). A copper cofactor for the ethylene receptor ETR1 from Arabidopsis. *Science* **283**: 996-998.

Rodriguez, P. L. (1998). Protein phosphatase 2C (PP2C) function in higher plants. *Plant Molecular Biology* **38**: 919-27.

Rogowsky, P. M., F. L. Y. Gtuidet, P. Langridge, K. W. Shepherd and R. M. D. Koebner (1991). Isolation and characterization of wheat-rye recombinants involving chromosome arm 1DS of wheat. *Theoretical Applied Genetics* **82**: 537-544.

Rolston, M. P. (1978). Water impermeable seed dormancy. *Botanical Review* **44**: 365-396.

Romagosa, I., F. Han, J. A. Clancy and S. E. Ullrich (1999). Individual locus effects on dormancy during seed development and after ripening in barley. *Crop Science* **39**: 74-79.

Rouse, D., P. Mackay, P. Stirnberg, M. Estelle and O. Leyser (1998). Changes in auxin response from mutations in an AUX/IAA gene. *Science* **279**: 1371-1373.

Roy, J. K., M. Prasad, R. K. Varshney, H. S. Balyan, T. K. Blake, H. S. Dhaliwal, H. Singh, K. J. Edwards and P. K. Gupta (1999). Identification of a microsatellite on chromosome 6B and a STS on 7D of bread wheat showing an association with preharvest sprouting tolerance. *Theoretical Applied Genetics* **99**: 336-340.

Roy, S. K. S., A. Hamid, M. Giashuddin Miah and A. Hashem (1996). Seed size variation and its effects on germination and seedling vigour in rice. *Journal of Agronomy Crop Science* **176**: 79-82.

Russell, L., V. Larner, S. Kurup, S. Bougourd and M. Holdsworth (2000). The Arabidopsis COMATOSE locus regulates germination potential. *Development* **127**: 3759-3767.

Sachidanandam, R., D. Weissman, S. C. Schmidt, J. M. Kakol, L. D. Stein, G. Marth, S. Sherry, J. C. Mullikin, B. J. Mortimore, D. L. Willey, S. E. Hunt, C. G. Cole, P. C. Coggill, C. M. Rice, Z. Ning, J. Rogers, D. R. Bentley, P. Y. Kwok, E. R. Mardis, R. T. Yeh, B. Schultz, L. Cook, R. Davenport, M. Dante, L. Fulton, L. Hillier, R. H. Waterston, J. D. McPherson, B. Gilman, S. Schaffner, W. J. Van Etten, D. Reich, J. Higgins, M. J. Daly, B. Blumenstiel, J. Baldwin, N. Stange-Thomann, M. C. Zody, L. Linton, E. S. Lander and D. Altshuler (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**: 928-933.

Sakai, H., J. Hua, Q. G. Chen, C. Chang, L. J. Medrano, A. B. Bleecker and E. M. Meyerowitz (1998). ETR2 is an ETR1-like gene involved in ethylene signaling in Arabidopsis. *Proceedings of the National Academy of Science USA* **95**: 5812-5817.

Sasaki, T. and R. R. Sederoff (2003). Genome studies and molecular genetics - The rice genome and comparative genomics of higher plants - Editorial overview. *Current Opinion in Plant Biology* **6**: 97-100.

Sayed, H., H. Kayyal, L. Ramsey, S. Ceccarelli and M. Baum (2002). Segregation distortion in doubled haploid lines of barley (*Hordeum vulgare* L.) detected by simple sequence repeat (SSR) markers. *Euphytica* **125**: 265-272.

Schmidt, R. (2002). Plant genome evolution: lessons from comparative genomics at the DNA level. *Plant Molecular Biology* **48**: 21-37.

Schumacher, K. and J. Chory (2000). Brassinosteroid signal transduction: still casting the actors. *Current Opinion in Plant Biology* **3**: 79-84.

Schwartz, S. H., K. M. LeonKloosterziel, M. Koornneef and J. A. D. Zeevaart (1997). Biochemical characterization of the ABA2 and ABA3 mutants in *Arabidopsis thaliana*. *Plant Physiology* **114**: 161-166.

Scott, R. W., M. Appleyard, G. Fellowes and E. J. M. Kirby (1983). Effect of genotype and position in the ear on carpel and grain growth and mature grain weight of spring barley. *Journal of Agricultural Science, Cambridge* **100**: 383-391.

Seo, M., A. Hanada, A. Kuwahara, A. Endo, M. Okamoto, Y. Yamauchi, H. North, A. Marion-Poll, T. P. Sun, T. Koshiba, Y. Kamiya, S. Yamaguchi and E. Nambara (2006). Regulation of hormone metabolism in Arabidopsis seeds: phytochrome regulation of abscisic acid metabolism and abscisic acid regulation of gibberellin metabolism. *Plant Journal* **48**: 354-366.

Seshu, D. V. and M. E. Sorrells (1986). Genetic studies on seed dormancy in rice. Rice Genetics. Manila, The Philippines, IRRI**:** 269-382.

Sherman, J. D., A. L. Fenwick, D. M. Namuth and V. Lapitan (1995). A barley RFLP map: alignment of three barley maps and comparisons to Gramineae species. *Theoretical Applied Genetics* **91**: 681-690.

Simillion, C., K. Vandepoele, M. C. Van Montagu, M. Zabeau and Y. Van de Peer (2002). The hidden duplication past of *Arabidopsis thaliana*. *Proceedings of the National Academy of Science USA* **99**(21): 13627-13632.

Smilde, W. D., J. Haluskova, T. Sasaki and A. Graner (2001). New evidence for the synteny of rice chromosome 1 and barley chromosome 3H from rice expressed sequence tags. *Genome* **44**: 361-367.

Solano, R. and J. R. Ecker (1998). Ethylene gas: perception, signaling and response. *Current Opinion in Plant Biology* **1**: 393-398.

Soller, M., T. Brody and A. Genizi (1979). The expected distribution of marker-linked quantitative effects in crosses between inbred lines. *Heredity* **43**: 179-180.

Spielmeyer, W., M. Ellis, M. Robertson, S. Ali, J. R. Lenton and P. M. Chandler (2004). Isolation of gibberellin metabolic pathway genes from barley and comparative mapping in barley, wheat and rice. *Theoretical Applied Genetics* **109**: 847-855.

Steber, C. M. and P. McCourt (2001). A role for brassinosteroids in germination in Arabidopsis. *Plant Physiology* **125**: 763-769.

Stoy, V. and K. Sundin (1976). Effects of growth regulating substances in cereal grain germination. *Cereal Research Communications* **4**: 157-163.

Sun, T. P. and Y. Kamiya (1994). The Arabidopsis GA1 locus encodes the cyclase ent-kaurene synthetase A of gibberellin biosynthesis. *Plant Cell* **6**: 1509-1518.

Takeda, K. (1996). Varietal variation and inheritance of seed dormancy in barley. Pre-harvest Sprouting in Cereals 1995. K. Noda and D. Mares. Osaka, Japan, Centre for Academic Societies**:** 205-212.

Taketa, S., M. Choda, R. Ohashi, M. Ichii and K. Takeda (2002). Molecular and physical mapping of a barley gene on chromosome arm 1HL that causes sterility in hybrids with wheat. *Genome* **45**: 617-625.

Tanaka, A. (1999). Mutation induction by ion beams in *Arabidopsis*. *Gamma Field Symposium* **38**: 19-28.

Tanksley, S. D. (1993). Mapping polygenes. *Annual Review of Genetics* **27**: 205-233.

Teichmann, S. A. and M. M. Babu (2004). Gene regulatory network growth by duplication. *Nature Genetics* **36**: 492-496.

The Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana.* *Nature* **408**: 796-815.

Thomas, W. T. B., W. Powell, J. S. Swanston, R. P. Ellis, K. J. Chalmers, U. M. Barua, P. Jack, V. Lea, B. P. Forster, R. Waugh and D. B. Smith (1996). Quantitative trait loci for germination and malting quality characters in a spring barley cross. *Crop Science* **36**: 265-273.

Thomas, W. T. B., B. Gertson and B. P. Forster (2003). Doubled haploids in breeding. Doubled haploid production in crop plants. M. Maluszynski, K. J. Kasha, B. P. Forster and I. Szarejko. Dordrecht, Kluwer Academic Publishing**:** 337-350.

Thornton, T. M., S. M. Swain and N. E. Olszewski (1999). Gibberellin signal transduction presents ellipsis the SPY who O-GlcNAc'd me. *Trends Plant Science* **4**: 424-428.

Tikhonov, A. P., P. J. SanMiguel, Y. Nakajima, N. M. Gorenstein, J. L. Bennetzen and Z. Avramova (1999). Colinearity and its exceptions in orthologous adh regions of maize and sorghum. *Proceedings of the National Academy of Science USA* **96**: 7409-7414.

Torada, A., S. Ikeguchi and M. Koike (2005). Mapping and validation of PCR-based markers associated with a major QTL for seed dormancy in wheat. *Euphytica* **143**: 251-255.

Trethowan, R. M. (1995). Evaluation and selection of bread wheat (*Triticum aestivum* L.) for preharvest sprouting tolerance. *Australian Journal of Agricultural Research* **46**: 463-474.

Trewavas, A. J. and R. Malho (1997). Signal perception and transduction: the origin of the phenotype. *Plant Cell* **9**: 1181-1195.

Ullrich, S. E., P. M. Hayes, W. E. Dyer, T. K. Blake and J. A. Clancy (1993). Quantitative Trait Locus analysis of seed dormancy in 'Steptoe' barley. Pre-harvest sprouting in cereals 1992. M. K. Walker-Simmons and J. L. Ried. St Paul, Minnesota, USA, AACC**:** 136-145.

Ullrich, S. E., J. A. Clancy, I. A. del Blanco, H. Lee, V. A. Jitkov, F. Han, A. Kleinhofs and K. Matsui (2008). Genetic analysis of preharvest sprouting in a six-row barley cross. *Molecular Breeding* **21**: 249-259.

Urao, T., K. Yamaguchi-Shinozaki and K. Shinozaki (2000). Two-component systems in plant signal transduction. *Trends in Plant Science* **5**: 67-74.

van Der Schaar, W., C. Alonso-Blanco, K. M. Leon-Kloosterziel, R. C. Jansen, J. W. van Ooijen and M. Koornneef (1997). QTL analysis of seed dormancy in Arabidopsis using recombinant inbred lines and MQM mapping. *Heredity* **79**: 190-200.

Van Deynze, A. E., J. Dubcovsky, K. S. Gill, J. C. Nelson, M. E. Sorrells, J. Dvorak, B. S. Gill, E. S. Lagudah, S. R. McCouch and R. Appels (1995). Molecular-genetic maps for chromosome 1 in Triticeae species and their relation to chromosomes in rice and oats. *Genome* **38**: 45-59.

van Dodeweerd, A. M., C. R. Hall, E. G. Bent, S. J. Johnson, M. W. Bevan and I. Bancroft (1999). Identification and analysis of homoeologous segments of the genomes of rice and *Arabidopsis thaliana*. *Genome* **42**: 887-892.

Van Os, H., P. Stam, R. G. F. Visser and H. J. Van Eck (2005). RECORD: a novel method for ordering loci on a genetic linkage map. *Theoretical Applied Genetics* **112**: 30-40.

Vandepoele, K., Y. Saeys, C. Simillion, J. Raes and Y. Van De Peer (2002). The automatic detection of homologous regions (ADHoRe) and its application to microcolinearity between Arabidopsis and rice. *Genome Research* **12**: 1792-1801.

Vandepoele, K., C. Simillion and Y. Van De Peer (2003). Evidence that rice and other cereals are ancient aneuploids. *Plant Cell* **15**: 2192-2202.

Varshney, R. K., T. C. Marcel, L. Ramsay, J. Russell, M. S. Röder, N. Stein, R. Waugh, P. Langridge, R. E. Niks and A. Graner (2007). A high density barley microsatellite consensus map with 775 SSR loci. *Theoretical Applied Genetics* **114**: 1091-1103.

Verbyla, A. P., B. R. Cullis and R. Thompson (2007). The analysis of QTL by simultaneous use of the full linkage map. *Theoretical Applied Genetics* **116**: 95-111.

Verslues, P. E. and E. A. Bray (2006). Role of abscisic acid (ABA) and *Arabidopsis thaliana* ABA-insensitive loci in low water potential-induced ABA and proline accumulation. *Journal of Experimental Botany* **57**: 201-212.

Vision, T. J., D. G. Brown and S. D. Tanksley (2000). The origins of genomic duplications in Arabidopsis. *Science* **290**: 2114-2117.

Vleeshouwers, I. M., H. J. Bouwmeester and C. M. Karssen (1995). Redefining seed dormancy: an attempt to integrate physiology and ecology. *Journal of Ecology* **15**: 1031-1037.

von Bothmer, R., K. Sato, T. Komatsuda, S. Yasuda and G. Fischbeck (2003). The domestication of cultivated barley. Diversity in Barley (*Hordeum vulgare*). R. von Bothmer, T. van Hintum, H. Knüpffer and K. Sato. Amsterdam, Elsevier**:** 9-27.

Wan, J., T. Nakazaki, K. Kawaura and H. Ikehashi (1997). Identification of marker loci for seed dormancy in rice (*Oryza sativa* L.). *Crop Science* **37**: 1759-1763.

Wan, J., Y. J. Cao, C. M. Wang and H. Ikehashi (2005). Quantitative Trait Loci associated with seed dormancy in rice. *Crop Science* **45**: 712-716.

Wang, H., L. Yu, F. Lai, L. Liu and J. Wang (2005). Molecular evidence for asymmetric evolution of sister duplicated blocks after cereal polyploidy. *Plant Molecular Biology* **59**: 63-74.

Warner, R. L., D. A. Kurdna, S. C. Spaeth and S. S. Jones (2000). Dormancy in white-grained mutants of Chinese Spring wheat (*Triticum aestivum* L.). *Seed Science Research* **10**: 51-60.

Wen, C. K. and C. Chang (2002). Arabidopsis RGL1 encodes a negative regulator of gibberellin responses. *Plant Cell* **14**: 87-100.

Wendel, J. F. (2000). Genome evolution in polyploids. *Plant Molecular Biology* **42**: 225-249.

Wenzl, P., J. Carling, D. Kudrna, D. Jaccoud, E. Huttner, A. Kleinhofs and A. Kilian (2004). Diversity Arrays Technology (DArT) for whole-genome profiling of barley. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 9915-9920.

Wilkinson, M. D., R. S. McKibbin, P. C. Bailey, J. E. Flintham, M. D. Gale, J. R. Lenton and M. J. Holdsworth (2002). Use of comparative molecular genetics to study pre harvest sprouting in wheat. *Euphytica* **126**: 27-33.

Willenborg, C. J., J. C. Wilderman, A. K. Miller, B. G. Rossnagel and S. J. Shirtliffe (2005). Oat germination characteristices differ among genotypes, seed sizes and osmotic potentials. *Crop Science* **45**: 2023-2029.

Winzeler, H., J. Schmidt and P. M. Fried (1987). Field performance of androgenetic doubled haploid spring wheat lines in comparison with lines selected by the Pedigree system. *Plant Breeding* **99**: 41-48.

Wolfe, K. H., M. Gouy, Y. W. Yang, P. M. Sharp and W. H. Li (1989). Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. *Proceedings of the National Academy of Science USA* **86**: 6201-6205.

Xiong, L. M., M. Ishitani and J. K. Zhu (1999). Interaction of osmotic stress, temperature, and abscisic acid in the regulation of gene expression in arabidopsis. *Plant Physiology* **119**: 205-211.

Xiong, L. M., H. J. Lee, M. Ishitani and J. K. Zhu (2002). Regulation of osmotic stress-responsive gene expression by the LOS6/ABA1 locus in Arabidopsis. *Journal of Biological Chemistry* **277**: 8588-8596.

Xue, G. P., C. L. McIntyre, S. Chapman, N. I. Bower, H. Way, A. Reverter, B. Clarke and R. Shorter (2006). Differential gene expression of wheat progeny with contrasting levels of transpiration efficiency. *Plant Molecular Biology* **61**: 863-881.

Yamaguchi, S., M. W. Smith, R. G. S. Brown, Y. Kamiya and T. P. Sun (1998a). Phytochrome regulation and differential expression of gibberellin 3 beta-hydroxylase genes in germinating Arabidopsis seeds. *Plant Cell* **10**: 2115-2126.

Yamaguchi, S., T. P. Sun, H. Kawaide and Y. Kamiya (1998b). The GA2 locus of Arabidopsis thaliana encodes ent-kaurene synthase of gibberellin biosynthesis. *Plant Physiology* **116**: 1271-1278.

Yamaguchi, S., Y. Kamiya and T. Sun (2001). Distinct cell-specific expression patterns of early and late gibberellin biosynthetic genes during Arabidopsis seed germination. *Plant Journal* **28**: 443-453.

Yan, G. P. and X. M. Chen (2006). Molecular mapping of a recessive gene for resistance to stripe rust in barley. *Theoretical Applied Genetics* **113**: 529-537.

Yang, J., C. Hu, H. Hu, R. Yu, Z. Xia, X. Ye and J. Zhu (2008). QTLNetwork: mapping and visualizing genetic architecture of complex traits in experimental populations. *Bioinformatics* **24**: 721-723.

Yin, Y. H., H. Cheong, D. Friedrichsen, Y. D. Zhao, J. P. Hu, S. Mora-Garcia and J. Chory (2002). A crucial role for the putative Arabidopsis topoisomerase VI in plant growth and development. *Proceedings of the National Academy of Sciences of the USA* **99**: 10191-10196.

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

Zanetti, S., M. Winzeler, M. Keller, B. Keller and M. Messmer (2000). Genetic analysis of pre-harvest sprouting resistance in a wheat x spelt cross. *Crop Science* **40**: 1406-1417.

Zeng, Z. B. (1993). Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proceedings of the National Academy of Science USA* **90**: 10972-10976.

Zeng, Z. B. (1994). Precision mapping of quantitative trait loci. *Genetics* **136**: 1457-68.

Zeng, Z. B., C. H. Kao and C. J. Basten (1999). Estimating the genetic architecture of quantitative traits. *Genetic Research* **74**: 279-289.

Zhang, F., G. Chen, Q. Huang, O. Orion, T. Krugman, T. Fahima, A. B. Korol, E. Nevo and Y. Gutterman (2005). Genetic basis of barley caryopsis dormancy and seedling desiccation tolerance at the germination stage. *Theoretical Applied Genetics* **110**: 445-453.

Zhang, F. L. and P. J. Casey (1996). Protein prenylation: molecular mechanisms and functional consequences. *Annual Review of Biochemistry* **65**: 241-269.

Zhou, L., J. C. Jang, T. L. Jones and J. Sheen (1998). Glucose and ethylene signal transduction crosstalk revealed by an Arabidopsis glucose-insensitive mutant. *Proceedings of the National Academy of Science USA* **95**: 10294-10299.

Zhu, C. S., C. M. Wang and Y. M. Zhang (2007). Modeling segregation distortion for viability selection I. Reconstruction of linkage maps with distorted markers. *Theoretical and Applied Genetics* **114**: 295-305.

Zivy, M., P. Devaux, J. Blaisonneau, R. Jean and H. Thiellement (1992). Segregation distortion and linkage studies in microspore-derived double haploid lines of *Hordeum vulgare* L. *Theoretical Applied Genetics* **83**: 919-924.

Zou, J., S. Zhang, W. Zhang, G. Li, Z. Chen, W. Zhai, X. Zhao, X. Pan, Q. Xie and L. Zhu (2006). The rice HIGH-TILLERING DWARF1 encoding an ortholog of Arabidopsis MAX3 is required for negative regulation of the outgrowth of axillary buds. *Plant Journal* **48**: 687-698.

APPENDIX A: Arabidopsis Candidate Gene List

Arabidopsis genes with a function associated to seed dormancy and/or germination. Description and evidence of gene function are given.

Gene	**Gene Name**	**NCBI CDS accession**	**Description of function pertaining to seed dormancy/germination**	**Determination of gene function**	**Reference**
ABA1	ABA deficient	NM_126103	Involved in 1st step of ABA biosynthesis.	Mutagenesis	Xiong et al. (2002)
ABA2	ABA deficient	NM_104113	Involved in the conversion of xanthoxin to ABA-aldehyde during ABA biosynthesis. Mutants insensitive to sucrose and glucose.	Mutagenesis	Schwartz et al. (1997), Seo et al. (2006)
ABA3	ABA deficient	NM_101519	Involved in the conversion of ABA-aldehyde to ABA, last step of ABA biosynthesis.	Mutagenesis	Leno-Kloosterziel et al. (1996)
ABH1	ABA hyper-sensitive	NM_126934	Involved in ABA signalling. Exhibit hypersensitive response to ABA in germination inhibition.	Mutagenesis	Hugouvieux et al. (2001)
ABI1	ABA insensitive	NM_118741	Involved in ABA signal transduction. Facilitates the activation of slow anion channels and stomatal closure mediated by ABA.	Mutagenesis	Finkelstein (1993)
ABI2	ABA insensitive	NM_125087	Encodes a protein phosphatase 2C and is involved in ABA signal transduction.	Mutagenesis	Merlot et al. (2001), Verslues & Bray (2006).
ABI3	ABA insensitive	NM_113376	Homologous to the maize transcription factor *Vp1*. Regulator of the transition between embryo maturation and early seedling development. Putative seed-specific transcriptional activator.	Mutagenesis	Nambara et al. (1994)
ABI4	ABA insensitive	NM_129580	Encodes a member of the *DREB* subfamily A-3 of *ERF*/*AP2* transcription factor family. Involved in ABA signal transduction.	Mutagenesis	Arenas-Huertero et al. (2000), Finkelstein et al. (1998)
ABI5	ABA insensitive	NM_129185	Encodes a member of the basic leucine zipper transcription factor family.	Mutagenesis	Lopez-Molina et al. (2001)
ABP1	Auxin binding protein	NM_116532	Auxin binding protein involved in cell elongation and cell division.	Mutagenesis	Jones et al. (1998)
ATPER1	1-cys peroxiredoxin	NM_128298	Encodes a protein similar to the 1-cys peroxiredoxin family of antioxidants. Expression is limited to seed (aleurone and embryo). Not induced by ABA or drought.	Transgenics	Dubreucq et al. (2000)
ATRAB28	*Arabidopsis thaliana* response to abscisic acid 28	ATRAB28	*Atrab28* plays a role in the ion cell balance during late embryogenesis and germination.	mRNA expression	Borrell et al. (2002)

AUX1	Auxin-resistant	NM_129368	Encodes an auxin influx transporter. Involved in response to auxin and ethylene stimulus. Auxin polar transport.	Non-traceable	Pickett et al. (1990)
AXR1	Auxin resistant	NM_100396	*AXR1*-similar to the ubiquitin-activating enzyme E1. Involved in auxin action.	mutagenesis	Abel et al 1995
AXR2	Auxin resistant	NM_113205	Plants carrying the dominant *IAA7*/*AXR2* mutation are severe dwarfs and display defects in growth orientation of both the shoot and root (gravitropic growth?). Roots resistant to ethylene, aba and auxin.	Mutatgenesis	Liscum et al. (2002), Nagpal et al. (2000)
AXR3	Auxin-resistant	NM_100306	Auxin-inducible *AUX*/*IAA* gene. Involved in auxin signalling.	Mutagenesis	Rouse et al. (1998), Ouellet et al. (2001), Liscum et al. (2002)
AXR6	Auxin resistant	NM_116491	Encodes a cullin that is a component of SCF ubiquitin ligase complexes involved in mediating responses to auxin and jasmonic acid. Homozygous auxin-resistant mutants arrest growth soon after germination. Heterozygotes display phenotypes consistent with impaired auxin response.	Mutagenesis	Hobbie et al. (2000)
BIN2	Brassinosteroid insensitive	NM_117987	*ATSK* (shaggy-like kinase) family member that encodes an AtSK protein involved in the cross-talk between auxin and brassinosteroid signalling pathways.	Mutagenesis	Li et al. (2001)
BIN3	Brassinosteroid insensitive	NM_112969	Response to brassinosteroid stimulus.	Mutgenesis	Yin et al. (2002)
BRI1	Brassinosteroid insensitive	NM_120100	Encodes a putative leucine-rich repeat receptor kinase involved in brassinosteroids signal transduction	Inferred from genetic reactions	Mora-Garcia et al. (2004)
CRE1	Cytokinin response	NM_126244	WOL-histidine kinase: cytokinin-binding receptor that transduces cytokinin signals across the plasma membrane. Splice variant – AT2G01830 histidine kinase. Regulation of seed germination.	Mutagenesis	Riefler et al. (2005)
CTR1	Constitutive triple response1	NM_120454	Mutant is enhancer of the aba insensitive mutant abi1-1. ctr1 monogenic mutants slightly aba resistant. Negative regulation in the ethylene signal transduction pathway. Interacts with *ETR1* and *ERS*.	Mutagenesis	Kieber et al. (1993), Gibson et al. (2001)
DAG2	Dof Affecting Germination	NM_130224	Encodes a protein containing Dof zinc finger motifs expression is limited to vascular system of the mother plant. Mutants are defective in seed germination, more dependent on light and cold treatment and less sensitive to gibberellin during seed germination.	Mutagenesis	Gualberti et al. (2002)
DAG1	Dof Affecting Germination	NM_116050	Zinc finger transcription factor of the *Dof* family involved in the control of seed germination.	Mutagenesis	Papi et al. (2002)
DWF1	Dwarf, auxin resistant	NM_112872	Replaces *DW1*. Involved in the conversion of the early brassinosteroid precursor 24-methylenecholesterol to campesterol. Dwarf phenotype.	Mutagenesis	Klahre et al. (1998), Choe et al. (1999a)

DWF4	Dwarf, auxin resistant	NM_114926	Encodes a 22 alpha hydroxylase whole reaction is a rate-limiting step in brassinosteroid biosynthetic pathway.	Mutagenesis	Choe et al. (1998)
DWF5	Dwarf, auxin resistant	NM_103926	Involved in the brassinosteriod biosynthetic process.	Mutagenesis	Choe et al. (2000)
DWF7	Dwarf, auxin resistant	NM_111126	STE1-Brassinosteroid biosynthetic enzyme, catalyses delta7sterol c-5 desaturation step. Mutant has dwarf phenotype.	Mutagenesis	Choe et al. (1999b)
EIN2	Ethylene insensitive	NM_120406	Involved in ethylene signal transduction. Acts downstream of *CTR1*. Involved in regulation of abscisic acid mediated signalling.	Mutagenesis	Ghassemian et al. (2000), Alonso et al. (1999)
EIN3	Ethylene insensitive	NM_112968	Involved in ethylene signal transduction.	Mutagenesis	Chao et al. (1997)
EBF1	EIN3-binding F-box protein	NM_128106	Encodes an F-box protein involved in the ubiquitin/proteasome-dependent proteolysis of *EIN3* (see *EIN3*)	Mutagenesis	Guo & Ecker (2003)
ERA1	Enhanced response to ABA	NM_123392	Encodes a beta subunit of farnesyl-trans-transferase, involved in ABA mediated signal transduction pathway.	Mutagenesis	Cutler et al. (1996)
ETR1	Ethylene response	NM_105305	Similar to prokaryote sensory transduction proteins. Binds ethylene, putative ethylene receptor. Response to ABA, GA and auxin stimulus.	Mutants	Rodriguez et al. (1999), Chiwocha et al. (2005) O'Malley et al. (2005)
ETR2	Ethylene response	NM_113216	Involved in ethylene perception.	In vitro binding assay	O'Malley et al. (2005)
FUS3	Fusca3	NM_113591	Specifies cotyledon identity. Regulator of gene expression during late embryogenesis.	Mutagenesis	Baumlein et al. (1994) Parcy et al. (1997) Gazzarrini et al. (2004)
GA1	GA requiring	NM_116512	Catalyzes the conversion of geranylgeranyl pyrophosphate (GGPP) to copalyl pyrophosphate (CPP) of gibberellin biosynthesis.	Enzyme assays	Sun & Kamiya (1994)
GA2	GA requiring	NM_106594	Catalyzes the second step in the cyclization of GGPP to ent-kaurene in the gibberellins biosynthetic pathway.	Mutagenesis, enzyme assays	Yamaguchi et al. (1998b)
GA3	GA requiring	NM_122491	Encodes a member of the *CYP701A* cytochrome p450 family that is involved in later steps of the gibberellin biosynthetic pathway.	Mutagenesis, enzyme assays	Helliwell et al. (1999)
GA4	GA requiring	NM_101424	Involved in later steps of the gibberellic acid biosynthetic pathway.	Expression pattern (RT-PCR), mutagenesis, in situ hybridisation assay.	Yamaguchi et al. (1998a), Hisamatsu et al. (2005), Kim et al. (2005)

GA5	GA requiring	NM_118674	Encodes *gibberellin 20-oxidase* that is involved in the later steps of the gibberellin biosynthetic pathway.	Mutagenesis	Hisamatsu et al. (2005)
GAI	GA insensitive	NM_101361	Similar to a putative transcription factor and transcriptional coactivators. Repressor of GA responses.	Mutagenesis	Peng et al. (2002)
GCR1	G-protein-coupled receptor 1	U95143	Encodes a protein similar to G-coupled receptor with 7 transmembrane regions. Overexpression studies suggest this gene is involved in dormancy and flowering. Reduction of expression results in decreased sensitivity to cytokinin.	Mutagenesis	Colucci et al. (2002)
HYL1	Hyponastic leaves	NM_100842	Encodes a dsRNA binding protein. Among many other things, mutants exhibit less sensitivity to auxin and cytokinin.	Mutagenesis	Lu & Federoff 2000
IAR3	IAA-alanine resistant	NM_104055	Encodes a member of the six Arabidopsis IAA-amino acid conjugate hydrolase subfamily and conjugates and IAA-Ala in vitro. Gene is expressed most strongly in roots, stems, flowers.	Mutagenesis	Davies et al. (1999)
IAR1	IAA-alanine resistant	NM_105480	Member of IAA-alanine resistance protein 1.	Mutagenesis	Lasswell et al. (2000)
ILR1	IAA-leucine resistant	NM_111156	Hydrolyzes amino acid conjugates of the plant growth regulator indole-3-acetic acid (IAA), including IAA-Leu and IAA-Phe.	Mutagenesis	Bartel et al. (1995), LeClere et al. (2002)
IMB1	Imbibition-inducible 1	AY180100	Encodes a member of the *BET* subgroup of bromodomain proteins, a novel class of putative transcription factors. Its expression is induced during seed imbibition and downregulated during germination.	Mutagenesis, expression assays	Duque et al. (2003)
LEC1	Leafy cotyledon	NM_102046	Transcriptional activator of genes required for both embryo maturation and cellular differentiation. Mutants are dessication intolerant, exhibit precocious meristem activation.	Mutagenesis	Lotan et al. (1998) Baumbusch et al. (2004)
LEC2	Leafy cotyledon	NM_102595	B3 domain transcription factor that induces embryo development	Mutagenesis	Meinke et al. (1994)
MARD1	Mediator of ABA regulated dormancy	NM_116186	Encodes a novel zinc-finger protein with a proline-rich N-terminus, identical to senescence-associated protein SAG102.	Annotation	He & Gan (2004)
PKL	Pickle	NM_128074	Cell proliferation, response to gibberellic acid stimulus, root development. Root defective.	Mutagenesis	Ogas et al. (1997)
PXA1/CTS	Peroxisome defective 3	AB070615	Encodes a peroxisomal protein of the ATP binding cassette (ABC) transporter class (PMP subfamily). The gene product promotes germination and represses embryo dormancy. *ABI3*, *ABA1*, *FUS3* and *LEC1* are epistatic to this gene.	Mutagenesis	Russell et al. (2000) Footitt et al. (2002)
RGL1	*RGA*-like	NM_105306	Negative regulator of GA responses, member of *GRAS* family of transcription factors.	Mutagenesis	Wen et al. (2002)

RPK1	Receptor-like protein kinase 1	U55875	*RPK1* is a leucine-rich receptor-like kinase located in the plasma membrane which is upregulated by abscisic acid, dehydration, high salt, low temperature, but not by other plant hormones. *RPK1* knock-out and antisense plants show an ABA-insensitive phenotype. *RPK1* plays a role in ABA-controlled cell proliferation and is a regulator of the ABA signal transduction pathway.	Mutagenesis	Osakabe et al. (2005), Hong et al. (1997)
SPY	Spindly	NM_111987	Contains a tetratricopeptide repeat region, and a novel carboxy-terminal region. Mutations in both domains increase GA signal transduction. Splice variant – gibberellin signal transduction protein.	Mutagenesis	Jacobsen et al. (1996), Jacobsen et al. (1998), Greenboim-Wainberg et al. (2005)

APPENDIX B: Splicing junctions of candidate genes

Amino acid residues at the splice junctions of Arabidopsis and rice orthologs for the candidate genes *ERA1, ABI2, ETR1,* and *HYL1*. Identical amino acid residues are indicated by grey shading. Full length cDNA accession numbers obtained from Genbank are given.

ERA1

Exon	Species	Nucleotides	Amino acids
EXON 1	ARABIDOPSIS AF214106	A T G C C A G T A G T A A C C 234 A C T C A A A A A T A C A T G	M P V V T T Q K Y M
	RICE AK073299	A T G G A C C C C C C C T C G 195 G C C A A T T C C C T C A T G	M D P P S A N S L M
EXON 2	ARABIDOPSIS AF214106	A T G G A G A T T C A G C G A 87 T C C T T A G A T G C T A A T	M E I Q R S L D A N
	RICE AK073299	T T A G A G C T G T G G C G T 87 G T G C T C G A T G C C A A T	L E L W R V L D A N
EXON 3	ARABIDOPSIS AF214106	C G A C C T T G G C T T T G T 105 C T T G G A C G C T G C C A G	R P W L C L G R C Q
	RICE AK073299	C G A C C T T G G C T G T G C 102 T T A T C T C G A T G T C A G	R P W L C L S R C Q
EXON 4	ARABIDOPSIS AF214106	G G C T C T G A A G G T G G A 36 G G T G G T C C T G G C C A A	G S E G G G G P G Q
	RICE AK073299	G A C A A A G A T G G T G G T 36 G G A G G A C C T G G A C A G	D K D G G G G P G Q
EXON 5	ARABIDOPSIS AF214106	C T T C C A C A T C T T G C A 81 C T T T C T T C A A T T A A T	L P H L A L S S I N
	RICE AK073299	T T G C C T C A T C T C G C T 84 C T A T C A T C G G T A A A C	L P H L A L S S V N
EXON 6	ARABIDOPSIS AF214106	A G A G A A A A A A T G T C T 54 A C A A G T G G A G G T T T C	R E K M S T S G G F
	RICE AK073299	A G G G A C A A C C T G T A C 54 A C A T C G G G A G C T T T C	R D N L Y T S G A F
EXON 7	ARABIDOPSIS AF214106	A G G A T G C A T G A T A T G 54 T A C A C T G C A A T T T C G	R M H D M Y T A I S
	RICE AK073299	A G A A T G C A T G A T G G T 54 T A T A C T G C A A T A T C G	R M H D G Y T A I S
EXON 8	ARABIDOPSIS AF214106	G T T G C A A G C A T C C T A 63 G G A G A T T A C A T C T T G	V A S I L G D Y I L
	RICE AK073299	G T T G C C A G C C T T G T G 63 G G A A A T T A C A T A A C A	V A S L V G N Y I T
EXON 9	ARABIDOPSIS AF214106	A G T T G C C A A A C T T A T 60 G A A G C T C A C G G T G G G	S C Q T Y E A H G G
	RICE AK073299	A G G T G T C A A A C C T A T 60 G A A G C T C A T G G T G G G	R C Q T Y E A H G G
EXON 10	ARABIDOPSIS AF214106	T A T A C C T A C T G T G G T 72 T T G G A T T C A T T A A T G	Y T Y C G L D S L M
	RICE AK073299	T A C A C T T T T T G T G G G 72 T T G G C T A G C T T G A T T	Y T F C G L A S L I
EXON 11	ARABIDOPSIS AF214106	A A T T G G G C T G T A C A T 87 T A C A C A T T T T G G C A G	N W A V H Y T F W Q
	RICE AK073299	G G C T G G G T G G C A T T T 87 T A C T C C T T T T G G C A G	G W V A F Y S F W Q
EXON 12	ARABIDOPSIS AF214106	G C A G C C C C T T G T G T T 165 G A G G A C A A C G A T G A A	A A P C V E D N D E
	RICE AK073299	G G A G C T G C T C T T G C T 117 A C G A G C T C T T C T A C T	G A A L A T S S S T
EXON 13	ARABIDOPSIS AF214106	G A T T C A G T G A A T G G T 111 C T C T T G T G C T C T A A G	D S V N G L L C S K
	RICE AK073299	G A A G C A G C A T A T T A T 111 C T G C T T T G C G C A C A G	E A A Y Y L L C A Q
EXON 14	ARABIDOPSIS AF214106	A T C C C T G A C G G T G G A 240 T T T A A A G C A G C A T G A	I P D G G F K A A STOP
	RICE AK073299	G T G C T G G A T G G A G G G 234 T T C T T T T C A A G C T A G	V L D G G F F S S STOP

ERA1

ABI2

	EXON 1			EXON 2			EXON 3		
ARABIDOPSIS Y08965	A T G G A C G A A G T T T C T	513	C A T G G C G G T T C T C A G	G T A G C G A A T T A T T G T	297	T C G G T T G A T C A C A A A	C C G G A T A G G G A T G A T	105	A T G T C A A G A T C C A T T
	M D E V S		H G G S Q	V A N Y C		S V D H K	P D R D D		M S R S I
RICE AK242616	A T G G A G G A C G T G G C G	633	C A C G G T G G C G C G C A G	G T T G C A A A T T A C T G T	342	T C A G T G G A T C A T A A A	C C T A A C A G G G A G G A T	105	A T G T C G C G A T C A A T A
	M E D V A		H G G A Q	V A N Y C		S V D H K	P N R E D		M S R S I

	EXON 4		
ARABIDOPSIS Y08965	G G C G A T A G A T A C C T T	357	A A A T C C T T G A A T T G A
	G D R Y L		K S L N STOP
RICE AK242616	G G T G A C A G A T A T C T C	342	A A G A G C A A A A G C T A A
	G D R Y L		K S K S STOP

ETR1

	EXON 1			EXON 2			EXON 3		
ARABIDOPSIS L24119	A T G G A A G T C T G C A A T	906	G T C G T C G C T G A T C A G	G T G G C T G T A G C T C T C	369	A C A T T A T T T A G A G A G	G T C C T C A A T C T G A T A	255	T A C T T G A G A G T G A A G
	M E V C N		V V A D Q	V A V A L		T L F R E	V L N L I		Y L R V K
RICE AK067813	A T G G A T G G A T G T G A C	906	G T C G T G G C T G A T C A G	G T T G C A G T T G C A C T T	369	G C T G T T T T C A A G G A G	G T A A T G A G T T T C A T C	267	T A T T T G A A A G T T C A G
	M D G C D		V V A D Q	V A V A L		A V F K E	V M S F I		Y L K V Q
RICE AK111696	A T G G A T G G A T C A T G T	909	G T C G T T G C T G A T C A G	G T A G C G G T C G C A T T G	369	T C C A C C T T T A C A G A T	G T G G T T A A T T T G A T T	267	T A C C T G G C T G T T C A G
	M D G S C		V V A D Q	V A V A L		S T F T D	V V N L I		Y L A V Q

	EXON 4			EXON 5			EXON 6		
ARABIDOPSIS L24119	G T A A A A G A C T C T G G A	128	G C C A T C T C C A A G A G G	T T T G T G A A T C T G A T G	198	A T G G A T G A G A A C G G G	G T A A G T A G A A T G G T G	361	T A C G A G G G C A T G T A A
	V K D S G		A I S K R	F V N L M		M D E N G	V S R M V		Y E G M STOP
RICE AK067813	A T T A A A G A T A C G G G C	129	G C C A T T T G C A A G A G G	T T T G T T A C T T T G A T G	240	C A G A G A A G T A T A T G A			
	I K D T G		A I C K R	F V T L M		Q R S I STOP			
RICE AK111696	G T A A A A G A C A C A G G C	126	G C C C T T T C C A G A A G A	T T T G T C A G T C T A A T G	237	C A A T C A A G C G T A T G A			
	V K D T G		A L S R R	F V S L M		Q S S V STOP			

HYL1

	EXON 1			EXON 2			EXON 3		
ARABIDOPSIS AF276440	A T G A C C T C C A C T G A T	25	G T T T C C T C T	G G T G T T T C C A A T T G C	273	G T T T C A C A A C C T G T T	C A C G A A A C G G G A T T A	972	G C A A G C C A C G C A T A A
	M T S T D		V S S	G V S N C		V S Q P V	H E T G L		A S H A STOP
RICE DQ009988	A T G G A C A T G C C G C C C	78	G C C A C C G C C G G A A T C	C G G G T T G A G A A C T G C	261	A A T A T C C C A G C T G T T	C A A G A G A C T G G G C T G	216	C T C G C A A T C C A A G G T
	M D M P P		A T A G I	R V E N C		N I P A V	Q E T G L		L A I Q G
RICE AB036988	A T G G A C A T G C C G C C C	78	G C C A C C G C C G G A A T C	C G G G T T G A G A A C T G C	261	A A T A T C C C A G C T G T T	C A A G A G A C T G G G C T G	219	C T G G C A A T C C A A G G T
	M D M P P		A T A G I	R V E N C		N I P A V	Q E T G L		L A I Q G

	EXON 4		
ARABIDOPSIS AF276440			
RICE DQ009988	C A A T C A G A G G G T T C G	677	G C T G G T G A C A T G T G A
	Q S E G S		A G D M STOP
RICE AB036988	C A A T C A G A G G G T T C G	720	G T C G T A A A G G A G T A G
	Q S E G S		V V K E STOP

APPENDIX C: Pilot map of Stirling/Harrington double haploid (DH) population using SSR and DArT markers

Genetic map of Stirling/Harrington DH population showing marker order and distances between SSR and DArT markers. DArT markers were screened on 90 lines of the population and are indicated by the prefix ‘bPb’.