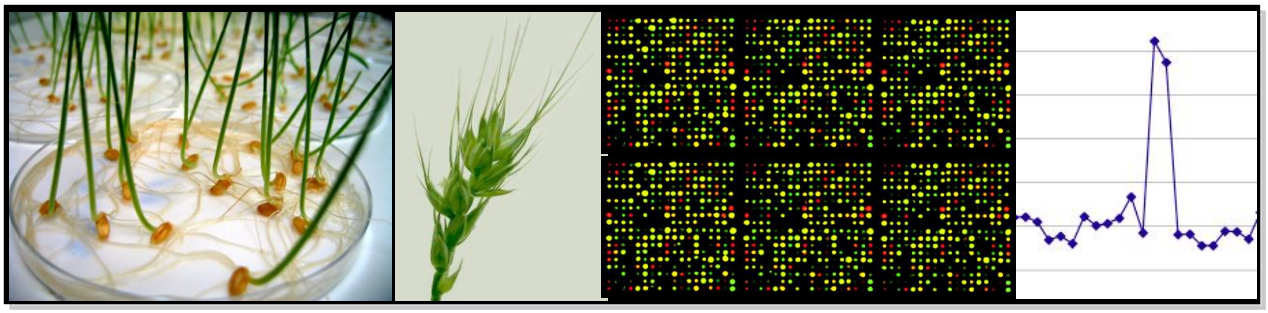


Genetic Diversity in Wheat:

Analysis using Diversity Arrays Technology (DArT) in bread and durum wheats

Brent Robert Thomson

B.Sc (Hons) Monash University

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Faculty of Agriculture, Food and Natural Resources

The University of Sydney

March 2011

Triticarte

THE UNIVERSITY OF SYDNEY

Diversity Arrays Technology Pty Ltd
High-throughput genome profiling

Abstract

With increasing demands on the quality and quantity of food required now and in the future, improvements to current agriculture practices are required. Increased food production requires utilisation of more agricultural land, pushing crops into non-traditional areas. The need for advances in agricultural technologies are not only required for current crop varieties, but for new varieties with increased tolerance to environmental stresses. Technological improvement means better crop yields and reduced land, water, fertilizer and pesticide use.

Diversity Arrays Technology (DArT) was used to study wheat diversity, specifically to identify polymorphic markers between various wheat cultivars for use in marker-assisted breeding programs. The hybridisation based technology was used to analyse various bread and durum wheat cultivars for increased understanding of genomic diversity.

Analysis shows that DArT is able to discriminate between tissue samples from wheat cultivars grown under various environmental stresses with polymorphic markers identified between samples treated with differing salt, light and temperature conditions. Epigenetic diversity was analysed through methylation detection using DArT to identify a list of candidate polymorphic markers. Markers were identified using the methylation sensitive restriction enzyme McrBC to generate control and treated targets. Diversity through cultivar exploration, looking at breeding experiments between cultivars with phenotypic extremes to examine salt tolerance versus in-tolerance using DArT produced a recombinant inbred line genetic linkage map. Bulk segregant analysis was also used to group phenotypic samples.

Candidate markers were identified between cultivars that can be used to genotyping tetraploid and hexaploid wheat cultivars for germplasm identification. In addition, the identification of trait-linked molecular markers, such as salt resistance, plant breeders can genotype individual plants and populations of cultivars to determine the most suitable cultivar to plant that best complements to its local environment. This eliminates the need for multiple planting cycles to optimize crop selections, and gives the plant breeder the highest possible chance for crop success (yield, quality, performance and cost).

Acknowledgments

Firstly I would like to thank my supervisors Dr Andrzej Kilian and Professor Peter Sharp for their help, support, guidance, wisdom and knowledge throughout my PhD studies. I would also like to thank them for the opportunity to study alongside them in a field that they are both so passionate about. I would like to thank the Value Added Wheat Cooperative Research Centre for their continued support and training, especially Bill Rathmel, Peter Vaughan, Clare Johnson and Neil Howes. At Diversity Arrays Technology Pty. Ltd. I would like to thank all staff and visitors over the past few years that have made my laboratory experience enjoyable, I will always treasure the friendships I have made and the immense cultural diversity that I was apart of. Specifically, big thanks to Eric Hutner for his guidance, support, words of wisdom and great sense of humor, Peter Wenzl for his scientific help especially with polymorphism and bulk segregant analysis and Grzegorz Uszynski for all your IT support. Thanks to Vanessa Caig, Jason Carling and Margaret Evers for your laboratory help and guidance but most of all for your friendship and laughter, I couldn't have made it with out you guys! Thanks to Damian Jaccoud, Sujin Patarapuwadol and Alex Wittenberg for your support, as well as being great mates, and to Kasia Heller-Uszynska, Ling Xie and Gosia Aschenbrenner-Kilian for being supportive and in helping provide a friendly environment to work in. Thanks to all our visiting scientists for their cultural inputs and diverse projects that they bring to the laboratory, thanks to Aurelie Bonin, Sophie Bouchet, Ilaria Catizone, Chandrika Perera and Prapit Wongtiem. Also a big thanks to my fellow students at the University of Sydney, Anneliese Rittau, Cindy Soh and Araluen Freeman, your support and friendship made things much easier!

I would like to thank Sue Lyons at the Australian National University Research School of Biological Sciences for your help with cultivating wheat plants in varying environmental conditions. Also thanks to Heidi Mitchel for your help with growing wheat plants at CAMBIA and Mona Akbari for your help in the initial DArT studies in wheat, and of course, for your friendship and countless laughs. Thanks to Rana Munns and Richard James at CSIRO Plant Industry for their salt treatment samples and phenotypic data.

Finally but most importantly, I would like to thank my family and friends for their help and support over the four year that I was away from ‘home’. Thanks to Mum, Dad and Natalie for their support over my many years of schooling and for always believing that I could get through it all! Thanks to my grandparents for your words of wisdom, support and our many phone calls that made me feel as if I was still at home. Thanks to Ryan Fairhall for all your help, support and gentle pressure to get things done, it means the world to me. Thanks to my friends in Melbourne, specifically to Elizabeth and Roger Kopek, Kathryn Richardson, Andrew Paydon, Claire Borg, Elissa Osborne, Kerina Alter and to everyone that made the seven-hour drive to Canberra to visit me.

Thanks all!

Andrzej, Gosia, Margaret, Jason, Ali, Vanessa, Eric, Grzegorz,

Aurelie, Sophie, me, Peter, Shir-Ying, Kasia and Ling

Student Certification

I, Brent Robert Thomson, do hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma of a university or other institute of higher learning, except where due acknowledgement is made in the text.

Brent Robert Thomson

31 / 03 / 2011

Dedicated to my Grandfather

Ronald Charles Furphy

(1917 – 2006)

Abbreviations / Acronyms

Abbreviation/Acronym	Description
AFLP	Amplified fragment length polymorphism
BSA	Bulk segregant analysis
DArT	Diversity Arrays Technology
DNA	Deoxyribonucleic Acid
F1	First filial generation
F2	Second filial generation
PCR	Polymerase Chain Reaction
RAPD	Randomly amplified polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic Acid
mRNA	Messenger RNA

Key Words and Definitions

Key Word	Definition
Allele	Two or more alternative forms of a gene resulting in different gene products and different phenotypes.
Allopolyploid	Composed of chromosome sets from different species.
Amplification	An increase in the number of copies of a specific DNA fragment either by cloning or polymerase chain reaction.
Amplified fragment length polymorphism (AFLP)	A method for detecting polymorphisms in DNA that uses restriction enzymes to digest DNA where a subset of fragments is selected for PCR amplification and visualization.
Autopolyploid	Composed of multiple sets of chromosomes from one species.
Back-cross	A cross between a progeny from a previous cross and either of its parental strains, or a cross of a heterozygote with its homozygous recessive parent, or the cross of a plant of unknown genotype with a homozygous recessive (also called a Testcross).
Bulk Segregant Analysis (BSA)	Plants from a segregating population are grouped according to phenotypic expression of a trait and tested for differences in allelic frequencies between the population bulks.
Call rate	DArTsoft polymorphism analysis value: an expression of reliability of the final scores, representing the number of scored slides against the maximum number of potential scores.
Chromosome	The self-replicating genetic structures of cells containing the cellular DNA that bears in its nucleotide sequence in linear array of genes. In prokaryotes, chromosomal DNA is circular, and the entire genome is carried on one chromosome. Eukaryotic genomes consist of a number of chromosomes.
Cleavage site	A specific nucleotide sequence at which a particular restriction enzyme cuts the DNA.

Clone	A group of genetically identical cells or individuals derived by asexual division from a common ancestor, an individual formed by some asexual process so that it is genetically identical to its parent.
Cloning	The process of asexually producing a group of cells (clones) or individuals, all genetically identical, from a single ancestor. In recombinant DNA technology, the use of DNA manipulation procedures to produce multiple copies of a single gene or DNA segment.
Clustering	DArTsoft polymorphism analysis value: uses a value of 0 or 1 to distribute every point into three groups.
Coleoptile	The pointed protective sheath covering the emerging shoot in monocotyledons.
Complementary DNA (cDNA)	DNA that is synthesised from a messenger RNA template that corresponds to expressed sequences of genomic DNA, DNA that is complementary to a particular DNA sequence.
DArT marker	DNA sequences specific to a species or individual found using the Diversity Arrays Technology platform, by screening a library of several thousand fragments from a genomic representation prepared from a pool of DNA samples that encompass the diversity of the species.
DArTdb	Laboratory information management system style database developed in-house at Diversity Arrays Technology Pty. Ltd.
DArTsoft	Polymorphism analysis software developed in-house at Diversity Arrays Technology Pty. Ltd.
Diploid	A full set of genetic material, consisting of one paired chromosome from each parental set.
Discordance	DArTsoft polymorphism analysis value: a complementary value of the reproducibility, expressing the overall variation of scores within the replicates.
Deoxyribonucleic	The molecule that encodes genetic information, a double-stranded molecule held together by weak bonds between base

acid (DNA)	pairs of the nucleotides adenine (A), guanine (G), cytosine (C), and thymine (T).
DNA clone	A section of DNA that has been inserted into a vector molecule, such as a plasmid or a phage chromosome, replicated then extracted to form many identical copies.
Epiallele	Epigenetic alleles that provide epigenetic variation that can be a used as a source of phenotypic variation
Epigenetic	The study of heritable changes in gene function that occur without a change in the sequence of nuclear DNA, including DNA methylation, associated with the development of an organism including gene regulation phenomena and gene silencing within.
Gamete	Specialised haploid germ cells that combine during fertilisation in organisms that reproduce sexually, in humans, sperm and ovum.
Gene	The fundamental physical and functional unit of heredity, an ordered sequence of nucleotides located in a particular position (locus) on a particular chromosome that encodes a specific product (i.e. protein, RNA molecule).
Gene expression	The process by which a gene's information is converted into the structures and functions of a cell, a multi-step process that begins with transcription, post transcriptional modification (Messenger RNA) and translation, followed by folding, post-translational modification and targeting. The amount of protein that a cell expresses depends on the tissue, the developmental stage of the organism and the metabolic or physiologic state of the cell.
Gene mapping	Determination of the relative positions of genes on a DNA molecule (chromosome or plasmid) and of the distance, in genetic map units (m.u.), between them.
Genetic diversity	A property of a community of organisms of a certain species, in which members of the community have variations in their chromosomes due to a large number of slightly dissimilar ancestors; this property makes the community in general more resistant to diseases or to changing ecological conditions.

Genetic map	A map based on the frequencies of recombination between markers during crossover of homologous chromosomes. The greater the frequency of recombination between two genetic markers, the further apart they are assumed to be.
Genetic map unit (m.u.)	The distance between genes, also called a centiMorgan, is defined as the distance between genes for which one product of meiosis in 100 is recombinant. A recombinant frequency (RF) of 1 percent is equivalent to 1 m.u.
Genotype	The hereditary constitution of an individual, or of particular nuclei within its cells
Haploid	A single set of chromosomes, present in the egg and sperm cells of animals and in the egg and pollen cells of plants.
Hexaploid	To have six sets of chromosomes.
Homologous chromosomes	A pair/group of chromosomes in a cell which have the same structure as each other containing the same genes at the same loci but may contain different alleles. Each homologous chromosome is inherited from a different parent.
In *silico*	Performed on computer or via computer simulation.
In *vitro*	Performed outside a living organism, literally ‘in glass’.
In *vivo*	Performed within a living organism.
Inflorescence	A group or cluster of flowers on a branch of a plant, where the seed is produced in wheat.
Inheritance	Biological inheritance is the process by which an offspring, cell or organism acquires or becomes predisposed to characteristics of its parent cell or organism. Through inheritance, variations exhibited by individuals can accumulate and cause a species to evolve. Genetic or Mendelian inheritance results from DNA replication and cell division. Epigenetic inheritance results from DNA modifications such as DNA methylation. Non-Mendelian inheritance includes inheritance based on cytoplasmic constituents, including mitochondria and chloroplasts.

Linkage	An association in inheritance between characters such that the parental character combinations appear among the progeny more often than the non-parental. The proximity of two or more markers on a chromosome; the closer together the markers are, the lower the probability that they will be separated during DNA repair or replication processes (binary fission in prokaryotes, mitosis or meiosis in eukaryotes), and hence the greater the probability that they will be inherited together.
Linkage map	A linkage map is created by finding the map distances (m.u.) between a number of traits that are present on the same chromosome.
Locus	The position of a gene on a chromosome (plural is loci).
Marker	An identifiable physical location on a chromosome (e.g., restriction enzyme cutting site, gene, minisatellite, microsatellite) whose inheritance can be monitored. Markers can be expressed regions of DNA (genes) or some segment of DNA with no known coding function but whose pattern of inheritance can be determined.
Meiosis	A diploid cell's genome is replicated once and separated twice, producing four sets of haploid cells each containing half of the original cell's chromosomes. These resultant haploid cells will fertilise with other haploid cells of the opposite gender to form a diploid cell again.
Mendel's first law	The two members of a gene pair segregate from each other during meiosis; each gamete has an equal probability of obtaining either member of the gene pair.
Mendel's second law	The law of independent assortment; unlinked or distantly linked segregating gene pairs assort independently at meiosis.
Messenger RNA	Messenger RNA (mRNA) is RNA that encodes and carries information from DNA (via transcription) to sites of protein synthesis (translation).
Methylation	Attachment of methyl groups ($-CH_3$) to DNA most commonly at cytosine residues. May be involved in regulation of gene

	expression and may prevent some restriction enzymes from cutting DNA at their recognition sites.
Microsatellite	Highly polymorphic DNA marker comprised of mononucleotide, dinucleotide, trinucleotide or tetra-nucleotide sequences that are repeated in tandem arrays and distributed throughout the genome.
Minisatellite	Highly polymorphic DNA markers comprised of a variable number of tandem repeats that tend to cluster near the telomeric ends of chromosomes.
Mitosis	The process by which a cell separates its duplicated genome into two identical halves. It is generally followed immediately by cytokinesis which divides the cytoplasm and cell membrane. This results in two identical daughter cells with a roughly equal distribution of organelles and other cellular components.
Monocotyledons	Any of various flowering plants, including grasses, that having a single cotyledon in the seed. Dicotyledons have two embryonic seed leaves that usually appear at germination.
Mutation	An abrupt change of genotype that is inherited. Any permanent and heritable change in DNA sequence. Types of mutations include point mutations, deletions, insertions, and changes in number and structure of chromosomes.
Nucleotide	A subunit of DNA or RNA consisting of a nitrogenous base (purine or pyrimidine), a phosphate molecule and a sugar molecule (deoxyribose in DNA and ribose in RNA).
Oligonucleotide	A short fragment of single-stranded DNA typically 5 to 50 nucleotides.
P	DArTsoft polymorphism analysis value: a measure of variation across individuals performed on one or more dimensions.
P generation	Parental generation of a breeding experiment.
Phenotype	The appearance of an organism with respect to a particular character or group of characters (physical, biochemical, and physiologic), as a result of the interaction of its genotype and its

	environment. Often used to define the consequences of a particular mutation.
PIC	DArTsoft polymorphism analysis value: Polymorphism Information Content, a value used to measure the 'informativeness' of a genetic marker for linkage studies.[1]
Plasmid	Typically circular double-stranded DNA molecules that replicate within a cell independently of the chromosomal DNA and usually occur in bacteria and some eukaryotic organisms. Their size varies from 1 to over 400 kilobase pairs and are anywhere from one copy, for large plasmids, to hundreds of copies of the same plasmid present in a single cell.
Polymerase chain reaction (PCR)	A method for amplifying a DNA base sequence using a heat-stable polymerase and two 20-base primers. Successive rounds of primer annealing, strand elongation, and dissociation produce rapid and highly specific amplification of the desired sequence.
Polymorphism	Difference in DNA sequence among individuals. Applied to many situations ranging from genetic traits or disorders in a population to the variation in the sequence of DNA or proteins.
Polyploid	Having multiple sets of chromosomes, triploid (3n), tetraploid (4n), pentaploid (5n), hexaploid (6n) and so on.
Primer	A nucleic acid strand (or related molecule) that serves as a starting point for DNA replication, required for PCR.
Q	DArTsoft polymorphism analysis value: measurement of the fraction of the total variation across all individuals due to bimodality, performed on one dimension.
Quantitative Trait Loci (QTL)	QTL are loci detected and mapped to a position on linkage maps by analyzing the statistical relationships between the quantative trait and marker loci.
Randomly amplified polymorphic DNA (RAPD)	A technique used for amplifying anonymous stretches of DNA using PCR with arbitrary primers.
RatioAvg	DArTdb image extraction output: average intensity based ratio of

	the logarithm of the target against reference using the average intensities of the spot.
RatioCov	DArTdb image extraction output: covariance of pixels based on the ratio of the logarithm of the ratio of target against reference using the covariance of pixels measurement.
RatioMed	DArTdb image extraction output: median intensity based ratio of the logarithm of the ratio of the target against reference using the median intensities of the spot.
RatioPix	DArTdb image extraction output: pixel based ratio of the logarithm of the target against reference calculated at the pixel level.
Reproducibility	DArTsoft polymorphism analysis value: as replicated individuals are supposed to give identical results, replicated points are expected to fall into the same cluster, after binarisation, every point is scored and DArTsoft controls the reproducibility of the experiment.
Restriction enzyme	Any of a group of enzymes that catalyze the cleavage of DNA at specific sites to produce discrete fragments, also called a *restriction endonuclease*.
RNA	Ribonucleic acid, a chemical found in the nucleus and cytoplasm of cells that plays an important role in protein synthesis and other chemical activities of the cell. The structure of RNA is similar to that of DNA. There are several classes of RNA molecules, including messenger RNA, transfer RNA, ribosomal RNA, and other small RNAs, each serving a different purpose.
Saline soil	Soil that contains a high concentration of soluble salts, most commonly the chlorides and sulphates of sodium, calcium and magnesium[2].
Senesce / Senescence	The aging of a plant after it reaches maturity, leading eventually to its death, usually refers to annuals after they set seed. Senescence is the combination of processes leading to deterioration that follows the period of development of an organism, cellular senescence is programmed cell death.

Sequestration	Loss of fluid content into spaces within the organism, so that the circulating volume diminishes, in the case of salt tolerance, the storage of salination ions into vacuoles within the plant cell.
Sodic soil	Soils with high levels of exchangeable sodium (Na) and low levels of total salts
Tetraploid	To have 4 (tetra) copies of chromosomes.
Tiller	A shoot that sprouts from the base of a grass.
Trait	An attribute or character of an individual within a species for which heritable differences can be defined.
Transcription	Transcription is the process through which a DNA sequence is enzymatically copied by an RNA polymerase to produce a messenger RNA. In the case of protein-encoding DNA, transcription is the beginning of the process that ultimately leads to the translation of the genetic code
Translation	Messenger RNA created by transcription is translated on the ribosomes into amino acids (with the help of transfer RNA) that form functional peptide or proteins.
Triticum aestivum	Common wheat, hexaploid, annual grass having erect flower spikes and light brown grains, sometimes cooked whole or cracked as cereal, usually ground into flour.
Triticum durum	The most common durum wheat grown in Australia, tetraploid with a high protein content with hard grains, high in gluten and used for pasta making.
Triticum turgidum	Durum wheat, tetraploid, with hard dark-colored kernels grown in southern Russia, North Africa, and northern central North America.
Variation	Differences in the frequency of genes and traits among individual organisms within a population.
Vernalisation	A requirement of some temperate cereal crops and trees for a period of low winter temperature to initiate or accelerate the flowering process.

Chapter Summary

Table of Contents

List of Figures

List of Tables

Introduction

“There is no happiness except in the realisation that we have accomplished something.”

Henry Ford

Chapter 1

1.0 Introduction

1.1 Thesis Outline

Research was undertaken at Diversity Arrays Technology (DArT) Pty. Ltd. located in Canberra, Australia, through the University of Sydney with supported by the Value Added Wheat Cooperative Research Centre (VAWCRC). The project was supervised by Dr. Andrzej Kilian (DArT), Professor Peter Sharp (University of Sydney) and Clare Johnson (VAWCRC).

Chapter 1 introduces the project; it's aims and potential outcomes as well as giving a general introduction to wheat utilisation and demands in Australia. Wheat evolution, biology and characterisation will be examined as a basic overview of wheats global importance, followed by a brief description of self and cross-pollination. Following is a summary of wheat biotechnology for crop improvements through genetic engineering and molecular breeding programs including the techniques available to perform and evaluate them. Information on epigenetics and methylation mechanisms will be discussed. Diversity Arrays Technology (DArT) will be explained and how it will be used within the project to look at methylation and sequence diversity in wheat.

Chapter 2 gives a detailed description of the methods including DArT protocols. Chapter 3 outlines developmental diversity, with analysis of different tissue samples across cultivars and analysis of seedling and mature leaf samples. Chapter 4 examines developmental diversity, looking at methylation polymorphisms and diversity under environmental stress conditions, including differing light and temperature stress as well as salt treatments. Chapter 5 looks at genetic diversity analysis, analysing different cultivars and breeding experiments. Bulk segregant analysis and recombinant inbred line analysis are used to look at cultivated and wild durum wheats to search for salt tolerant linked molecular markers. Chapters 6 makes final conclusions and examines overall conclusions, limitations and give a detailed discussion from the results obtained. It will report on future requirements and ongoing improvements and directions that are possible. Chapter 7 presents appendix data and references are shown in Chapter 8.

1.1.1 Aims

The aim of the project is to use DArT to discover and evaluate potential wheat DNA molecular markers. Specifically, to develop a high-throughput genotyping system to study wheat genetics and improve plant breeding methods. DArT will be applied to wheat genome studies looking at sequence diversity between wheat cultivars, as well as between tissue types. DArT will also utilize DNA methylation variation as a tool for the analysis of epigenetic phenomena in wheat.

With increasing demands on the quality and quantity of food required now and in the future, improvements to agriculture practices are required. This, as well as the need for larger agricultural land, pushing crops into non-traditional areas, further accentuates the need for advanced agricultural technologies, such as molecular marker technologies, as an increasingly important tool for identifying crop varieties for better performance. Through better plant breeding, comes faster results from the planning stages through to crop harvests.

This project will directly feed into the DArT Wheat mapping project already established to provide a high throughput genotyping services to plant breeders. The project is unique as it involves techniques previously not utilised by DArT to examine methylation diversity and not just DNA sequence diversity in wheat cultivars. Differences in the methylation status of wheat varieties and within a breeding experiment are important as it has been suggested that epigenetic modification of DNA plays an important role in environmental tolerance and adaptation in crops. The project will focus on DNA methylation detected by methylation-sensitive restriction enzymes as well as looking at sequence diversity in tetraploid and hexaploid wheat. Wheat cultivars will be grown under varying conditions and analysed for treatment, tissue and cultivar methylation and sequence polymorphisms.

1.1.2 Outcomes

The direct physical outcome from the project will be a collection of molecular markers that are unique to a set of breeding parameters. These markers will be combined to create a polymorphic diversity array to aid wheat breeders who are looking to improve certain traits within their crops, for example, salt tolerance or light and temperature growth response.

1.2 Wheat Development, Production and Demand

The precise origin of the wheat plant as we know it today is not known[3]. Wheat's great diversity can be attributed partially to its extensive cultivation and conservation in many parts of the world over a long period of time. Records indicate that common bread wheat (*Triticum aestivum*) has been traded in Asia from West to East from as early as 6000BC[4]. By 4000BC, Neolithic farmers were growing and improving wheat across large areas of North Africa and the near East, to South Asia and onwards to China. Farming and breeding techniques further spread into Europe, the Americas and Australia. Within the last century, plant breeding has become an important scientific discipline with many new varieties being developed. In 1970, Norman E. Borlaug, a plant breeder, won the Nobel Peace Prize for being the first to develop a new variety of wheat that yielded much more than previous crops grown in Asia[5]. These new varieties were grown in India, Pakistan and Bangladesh and produced enough grain to prevent millions of people dying during the 1960's famine. Events such as these were the catalyst for research into wheat genomics and the development of advanced molecular techniques for breeding wheat for improved quality and quantity.

Wheat was introduced into Australia in 1788 at the time of European settlement. William Farrer (1845-1906) developed wheat varieties adapted for Australian conditions in the early 20th century, with 'Federation' being the most famous[6]. Early maturity was a key selection criterion that gave his selections disease escape, rather than disease resistance. Federation was grown extensively in Australia from 1910-1925 with a further 29 varieties being developed by 1914[7]. Further wheat varieties were developed using drought resistant Indian wheats and high quality Canadian Fife wheats[8].

Wheat is Australia's most important and valuable crop, having a seasonal gross value approaching $3 billion dollars in 2002[9]. Wheat production is concentrated on mainland Australia in a narrow crescent known as 'the wheat belt' spanning over 13.9 million hectares[10] that produced 21.9 million tons of wheat in 2009/10[11] (Figure 1.1). Australia is one of the top 10 highest wheat producing countries contributing to the 2009/10 production of 677.2 million tones of wheat produced worldwide[12]. Rosegrant et al. (2001) project that between 1997 and 2020 demand for wheat will grow by 45%

(about 266 million tones) as shown in Figure 1.2. High demand for wheat makes it the largest produced grain crop worldwide[13].

1.3 Wheat Classification

Commercially important wheat varieties in Australia and throughout the world are classified into two common groups. The common bread wheats are hexaploid cereal crops grown for their flour producing qualities when their grains are milled into a fine powder. They are considered to be 'soft wheats' and are extensively used for breads, cakes, biscuits and fermented to make beer, vodka and bio fuels. The most common bread wheat is *Triticum aestivum*. Varieties include Janz, Westonia, Kukri and Frame. The second classification are the durum wheats which in contrast are tetraploid and have a high protein coat making them know as 'hard wheats'. They have a high gluten content which gives them great dough strength and are used for pasta and noodle production. Durum wheats are classified *Triticum durum* and their full scientific classification is shown together with common bread wheats in Table 1.1. Australian durum varieties include Kamilaroi, Wollaroi, Yallaroi and EGA Bellaroi, all released from the NSW Agriculture durum breeding program at Tamworth[14]. They have been specially developed to be resistant to stem, leaf and stripe rusts, and have a useful to high level of yellow spot resistance however are very susceptible to crown rot[15].

1.4 Wheat Biology

1.4.1 Wheat Growth and development

Plants of the genus *Triticum* are annuals with winter or spring forms. Australian bread wheat, *Triticum aestivum,* is a cereal of temperate climate and is planted in winter (June-August) or spring (May-June). Winter wheat requires a period below freezing temperature before it can form inflorescence (vernalisation) however spring wheat does not require this period. Once the seed has been planted, germination occurs at an optimum temperature of between 12 and 25°C (3-4°C minimum) and lasts approximately 4 to 10 days. The minimum moisture required for germination is generally in the range of 35 to 45% of the kernel dry weight[16]. During germination, the seminal root extends out of the seed shell first, followed by the coleoptile (embryonic

leaf). Adventitious roots are produced in association with the coleoptile node. When the coleoptile emerges from the soil, it stops growing, and the first true leaf pushes through its tip. The seedling is dependent upon energy and nutrients provided by the endosperm within the seed until the first leaf becomes photosynthetically functional[17]. Root axes are produced at predictable times in relation to shoot development, and the total number of roots formed is associated with the number of leaves on a tiller (flowering stem) and the degree of tillering[18].

Table 1.1: Scientific Classification of Bread and Durum wheats

Classification	**Bread Wheat**	**Durum Wheat**
Kingdom	Plantae	Plantae
Division	Magnoliophyta	Magnoliophyta
Class	Liliopsida	Liliopsida
Order	Poales	Poales
Family	Poaceae	Poaceae
Genus	*Triticum*	*Triticum*
Species	*T. aestivum*	*T. durum*
Common varieties	Janz, Westonia, Kukri, Frame, Angus, Grebe	Wollaroi, EGA Bellaroi, Yallaroi, Kamilaroi

Source: New South Wales Department of Primary Industries, Australia[19].

http://www.dpi.nsw.gov.au/agriculture

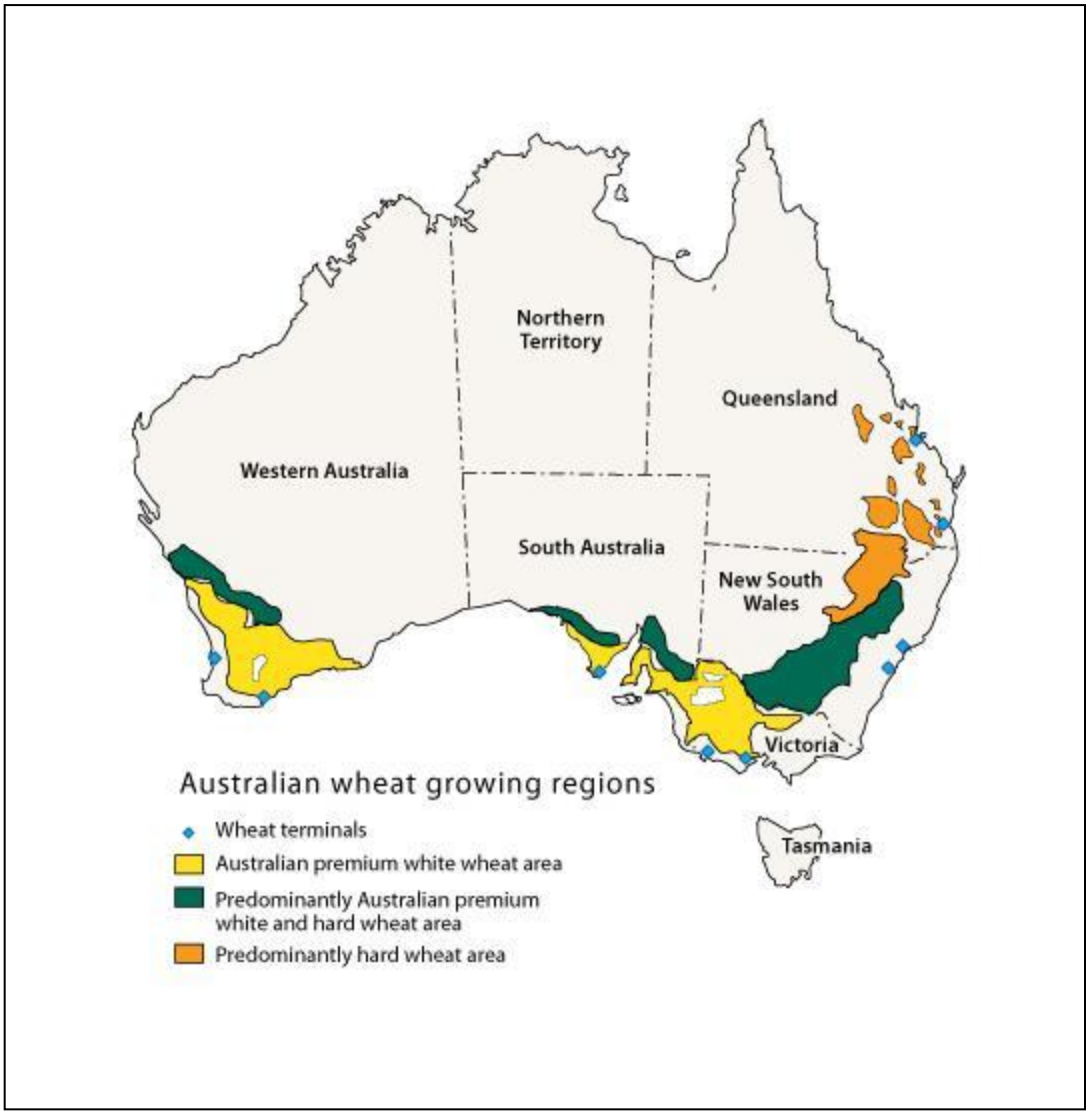


Figure 1.1: Australia Wheat Growing Regions (2009)

Source: Australian Bureau of Agricultural and Resource Economics (ABARE) Australian Crop Report, 2009, Number 151[20].

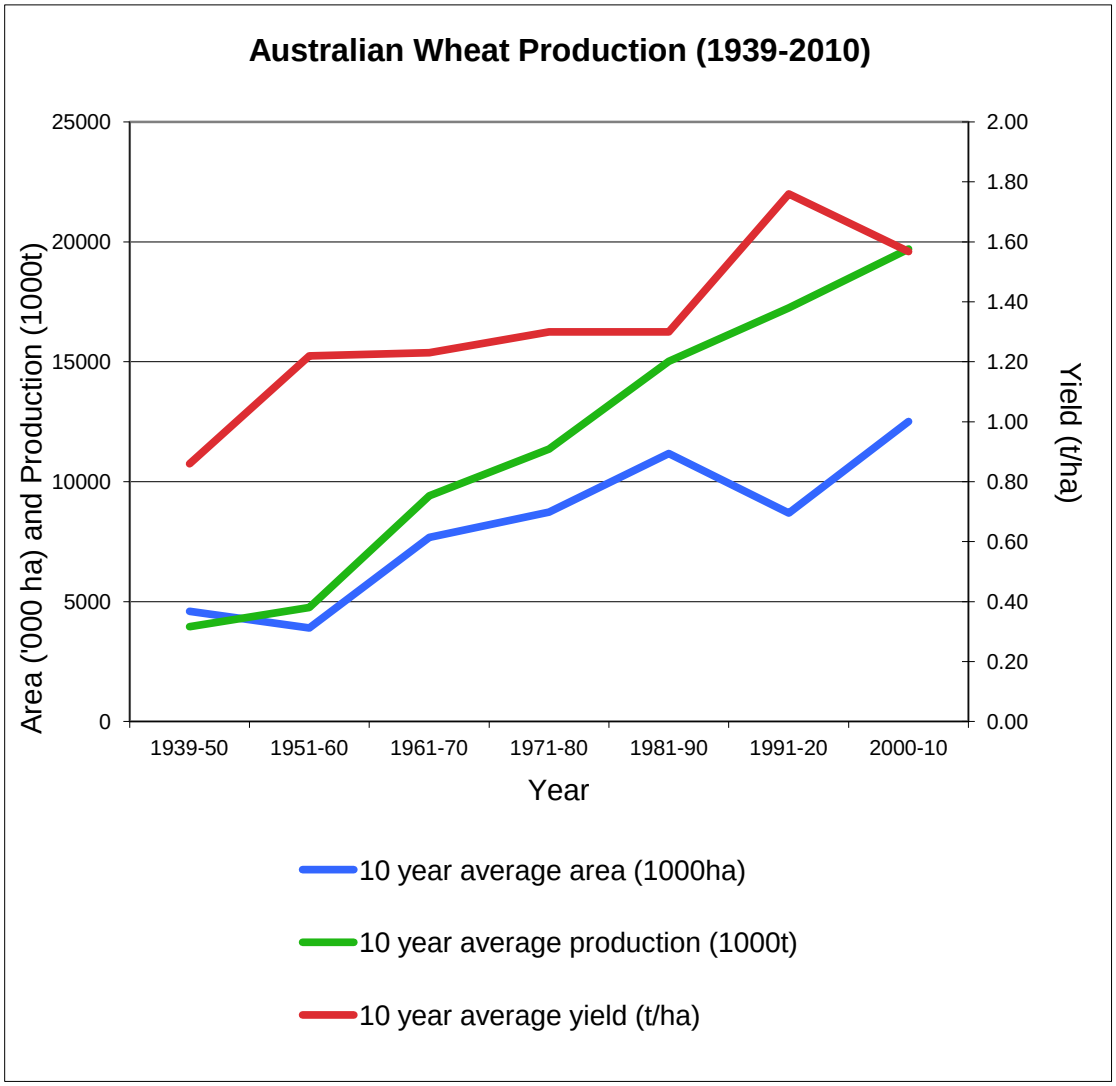


Figure 1.2: Australian Wheat Production (1939-2010)

Source: Adapted from Australian Wheat Board (AWB) (2004)[21] and The Australian Bureau of Agricultural and Resource Economics (ABARE) (2010)[22].

After germination, the vegetative shoot apex initiates and additional leaves are grown. The number of leaves is affected by genotype, light intensity, nutritional status of the plant and temperature. Kirby *et al (*1983) has shown that temperature has a major influence on leaf appearance and extension[23]. It has been shown that the minimum temperature for leaf extension was 0°C, with the optimum being 28°C with a maximum greater than 38°C[24]. Stem elongation coincides with the growth of leaves, tillers, roots and inflorescence[25]. Wheat plants are typically 30-150 cm high, and their differences are determined by genotype and growing conditions.[26] Vegetative growth for winter wheat is on average 280-350 days and a shorter 120-145 days for spring wheat. The shorter vegetative period for spring wheats is due to warmer temperatures that promote tiller formation and growth, thus maturing the plant faster than winter wheats[27]. The main shoot and early formed tillers complete development and form grains while tillers formed in later stages usually senesce prematurely[28]. This process is summarised in Figure 1.3, showing the emergence and growth stages 1, 2 and 3 [29].

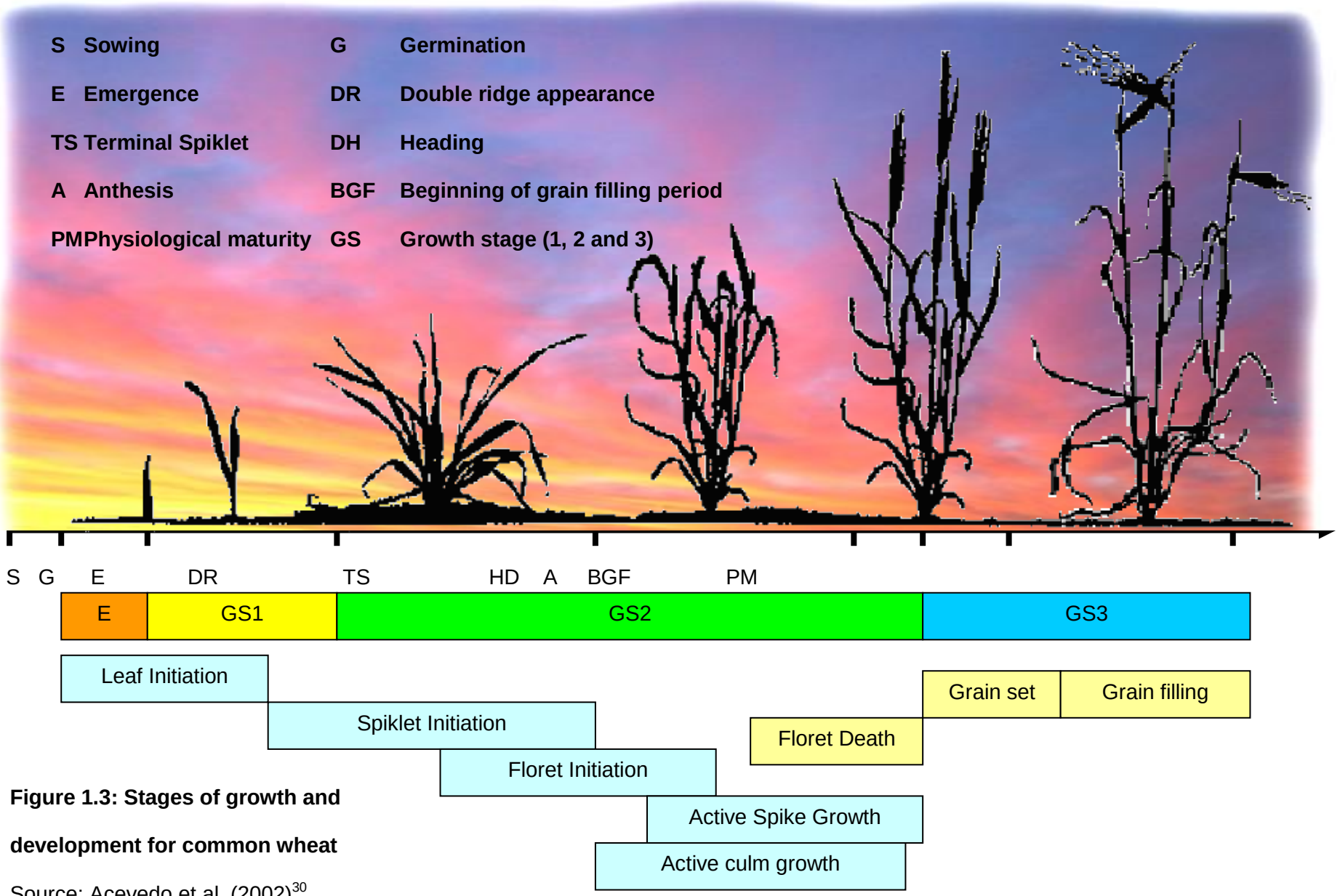


Figure 1.3: Stages of growth and development for common wheat

Source: Acevedo et al. (2002)[30].

1.5 Wheat Genetics

Wheat has been extensively studied for a wide range of agronomic traits located across its genome. There are over a thousand natural and induced variants that are well characterised for morphological, developmental, biochemical and disease resistant phenotypes[31]. The wheat genome is large, estimated at 16,000 Mb compared with that of Human (3,000 Mb), Rice (400 Mb), and Arabidopsis thaliana (130-140 Mb) as shown in Figure 1.4[32]. Bread wheat varieties are hexaploid and comprise three closely related genomes, designated A, B and D derived from three progenitor species. Gene redundancy is thus common with at least a triplicate homoeoallelic set for most genes. Durum wheat is tetraploid, containing only the A and B genomes. It is believed that hexaploid wheat evolved from a crossing of tetraploid wheat and the species containing the D genome. Cytogenetic studies and sequencing of bacterial artificial chromosome (BAC) clones containing portions of the progenitor A and D genomes indicate the presence of high density gene regions giving rise to exponential growth of expressed sequence tagged (EST) databases.

Functional genomic studies combined with the phenotypic variants at a given locus play an important part in wheat breeding. Hexaploid wheat contains 42 chromosomes in the 2N state and is not homogenous, in that three different genomes each contribute 7 chromosomes to give a hexaploid compliment of genetic material. Thus the 7 chromosomes from each of the 3 genomes in duplicate give rise to the 42 chromosomes (Figure 1.5). The A and D genomes are the most similar while genome B is more diverged. During Meiosis, the 42 chromosomes line up in pairs so that chromosome 1(A) pairs with 1(A), 1(B) with 1(B) and 1(D) with 1(D). This occurs across all wheat chromosomes. This was demonstrated by Martinez-Perez *et al* (1999) where the homologous paring (Ph1) chromosomal locus, responsible for the regulation of intra-genome pairing of chromosomes in wheat, prevents chromosome 1(A) from paring with 1(B) or 1(D)[33]. It was shown that the absence of the Ph1 locus interrupted this regulation, so that chromosome 1(A) paired with 1(A), 1(B) or 1(D)[34]. However, even with the disruption of the Ph1 locus, chromosome 1 will always pair with chromosome 1, indicating that the role of this locus is to distinguish one genome from another, and not one chromosome from another[35].

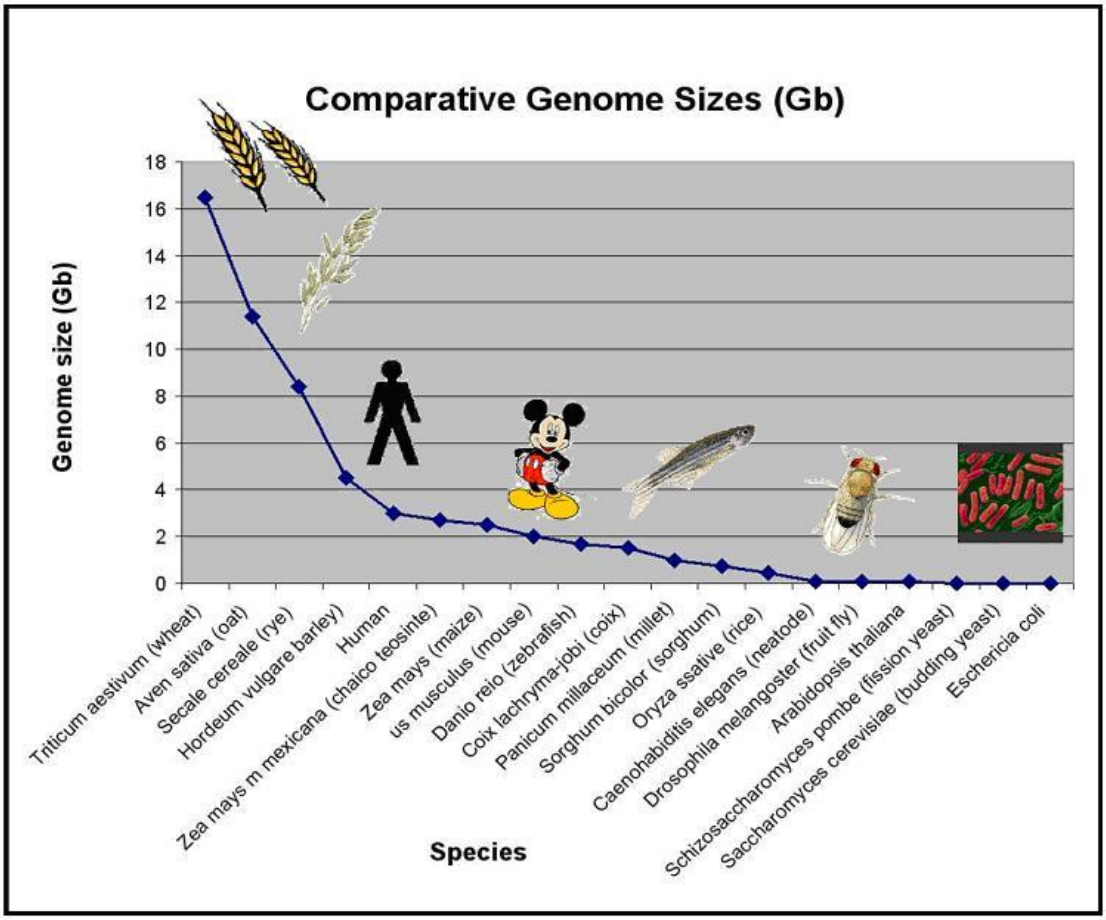


Figure 1.4: Genome size comparisons

Source: Adapted from the Human Genome Project[36] and other sources.

1.6 Wheat Evolution

At present it is understood that hexaploid wheat is the product of two unique hybridisation events. In the first hybridisation event, the A genome progenitor combined with the B genome progenitor to form tetraploid wheat, commonly known today as durum or pasta wheat (2n=4x=28, AABB). This hybrid occurred in the cytoplasm of the B genome. The second event involved hybridisation between the tetraploid (AABB) form and the D genome progenitor[37] to form the basic hexaploid configuration, AABBDD, again in the B genome cytoplasm[38]. Figure 1.6 shows the chromosomal arrangement of hexaploid wheat. McFadden and Sears (1946) identified the D genome progenitor as *Triticum tauschii* (Coss.) Schmal. (formerly *Aegilops squarossa*)[39]. The A genome progenitor has been identified as *Triticum boeoticum* L. Synonyms for this species are *T. urartu, T. monococcum* and *T. thaoudar*[40]. Differences between the C-banding patterns of chromosome 4A of *T. monococcum* and chromosome 4A of *T. turgidum* in figure 1.6 are attributed to structural rearrangements that occurred in the tetraploid form[41].

The specific identity of the B genome donor remains unclear. Sarkar & Stebbins (1956) originally proposed that the B genome donor was based upon *T. speltoides* Tausch[42] however Feldman (1979) concluded that *T. longissimum* and *T. searsii* were candidates for the B genome progenitor. Nath *et al.* (1983) later concluded that *T. searsii* was the possible source of the B genome after studying several likely progenitors with DNA hybridisations.

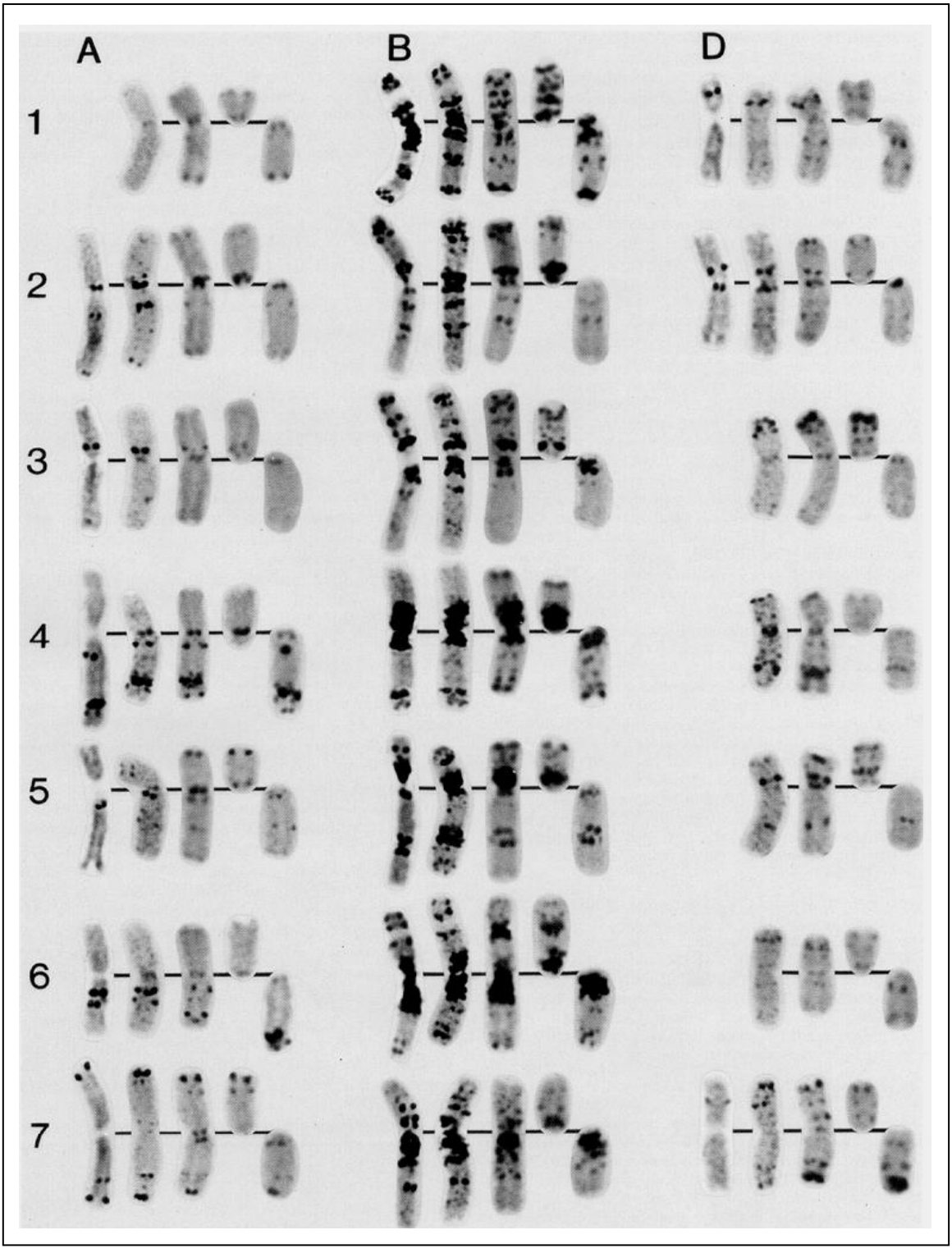


Figure 1.5: Hexaploid Wheat Chromosomes stained (from left to right) by N-banding, modified C-banding, and C-banding (whole and telosomic chromosomes). Chromosomes 1A and 3D to 6D do not show any N-bands and are not shown.

Source: Gill et al, 1991[43].

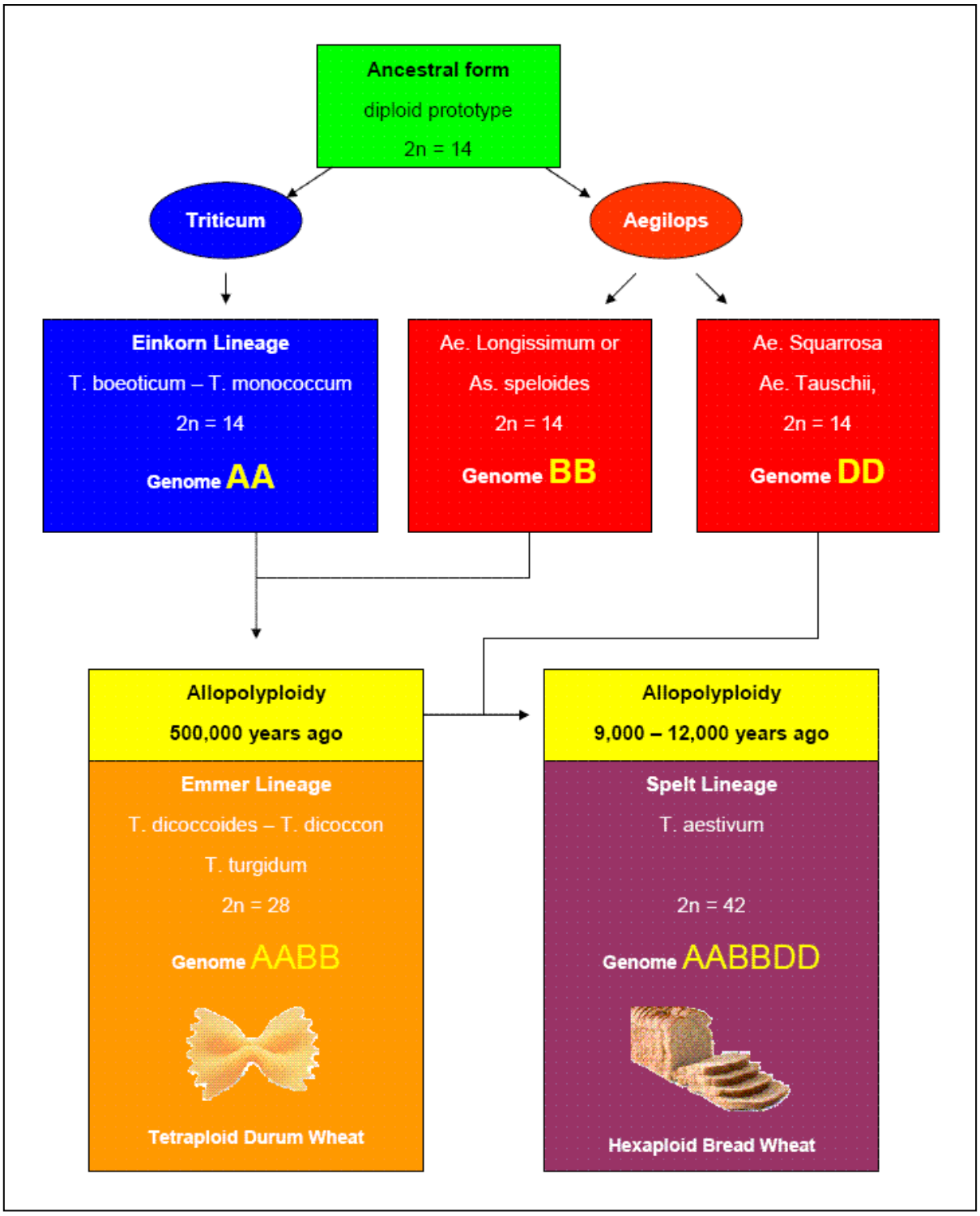


Figure 1.6: Hexaploid and Tetraploid wheat evolution

Source: adapted from "The biology and ecology of bread wheat (Triticum aestivum L. Em Thell) in Australia", Department of Health and Ageing, Office of the Gene Technology Regulator, Australian Government, April 2005. [44]

1.7 Environmental Stress

Environmental stress is an important consideration in wheat breeding, as plants need to be grown in a variety of locations and environmental conditions. Light intensities, temperature fluctuations, water availability, salt concentrations and soil composition and structure are all important considerations in crop production. They can influence crop growth, yield and quality of the desired product as well as contribute to the cost of production and to the use of pesticides and fertilizers.

1.7.1 Light and Temperature Stress

Environmental stress plays an important role in the growth and development of a plant. In wheat, it has been shown that temperatures above 35°C during the grain filling growth stage are directly associated with weaker dough properties[45]. High temperature is a major determinant of wheat development and growth, with decreasing yields by 3 to 5% per 1°C increase above 15°C observed by Gibson and Paulsen (1999)[46].

1.7.2 Salt Stress

With the demand for food increasing world wide, wheat and other crops are being grown in a wider range of environments where both saline and sodic soils are commonly encountered. In Victoria and South Australia, it has been reported that over half of the agricultural soil is sodic[47]. Durum cultivars are relatively intolerant of saline and sodic soils compared to bread wheats and other hexaploid wheats, resulting in significant yields reductions[48].

Greenway and Munns (1980) state that the two main mechanisms for salt tolerance in plants are low rates of salt transport to shoots and the tolerance of high leaf salt concentrations by efficient sequestration within cell vacuoles[49]. In wheat, Shah *et al* (1987) showed that salt tolerance is associated with low rates of transport of Na+ to shoots with high selectivity for K+ over Na+, but not Cl- transport.[50] The same study showed that hexaploid wheats (A, D and D genomes) have a low rate of Na+ accumulation and enhanced K+/Na+ discrimination. Gorham *et al* (1987) showed that the character for this is found on chromosome 4D[51] and Dubcovsky *et al* (1996) showed it to be controlled by the single locus *Kna*1[52]. As one would expect, tetraploid

wheats that lack the D genome (A and B genomes only), have been shown to have high rates of Na+ accumulation and poor K+/Na+ discrimination. Homologous recombination experiments with the wheat 4D chromosome by Dvorak *et al* (1994) have created a novel tetraploid germplasm with low accumulation of Na+ and enhanced K+/Na+ discrimination, thus improving the salt tolerance of the durum wheat.

1.7.2.1 Screening methods for salt tolerance

Screening for salt tolerance can be performed in the field based on growth or yield, however this can be difficult due to spatial heterogeneity of soil chemical and physical properties and seasonal fluctuations in rainfall[53]. Srivastava and Jana (1984) reported a field study where the International Center for Agricultural Research in the Dry Areas (ICARDA) advanced durum breeding lines indicated that significant genetic variation for salt tolerance may exist, but the confounding presence of drought stress made it difficult to identify genotypes with salt tolerance.[54] Thus the majority of screening methods are performed under controlled environments, including measurements of growth (root, leaf, biomass, yield), measurements of injury (leakage, chlorophyll content or fluorescence), for specific traits (Na^+ exclusion, K^+/Na^+ discrimination, Cl^- exclusion) or by germination or survival rates. Trials under controlled environments often correlate poorly with their performance in field trials, limiting the successful application of selected plant varieties. Studies by Francois *et al* (1986)[55] and Gorham *et al* (1987)[56] have shown that genetic differences for Na^+ exclusion correlate highly with differences in salinity tolerance between hexaploid and tetraploid wheat. Field trials using molecular market technology will thus play an important role in plant breeding as seedlings can be genotyped for certain traits without the need for large scale plantations, lengthy growth periods or phenotypic data collection/testing.

1.8 Techniques for Crop Improvements

The success of wheat breeding has largely come from the application of new technologies to breeding and selection. Biotechnology offers two new means for improving wheat, firstly through genetic engineering and secondly through the development and application of molecular markers technology.

1.8.1 Genetic Engineering

Genetically engineered crops are readily available in most countries including Australia. These include wheat, canola, corn, cotton, flax, papaya, potatoes, soybeans, squash, sugar beet and tomatoes. The prime targets so far considered are the engineering of resistance to herbicides, resistance to viral and some fungal pathogens and the modification of the quality characteristics of grain, particularly starch and protein composition. Anderson et al. (2003) illustrates this with their development of aphid resistant wheat in Colorado, USA[57]. The study shows that genetic resistance in wheat is the most effective and economical means to control the damage caused by the aphid. The group used the *Dn7* rye gene located on chromosome 1RS that confers resistance to the Russian wheat aphid and transferred it from rye into a wheat background via a 1RS/1BL translocation. This new variety of wheat allows for reduced pesticide usage and increases crop yields for farmers.

1.8.2 Molecular Markers

Molecular marker technology offers a wide range of novel approaches to improving the efficiency of selection strategies. The techniques are based on the detection of sequence variation between varieties. Where the sequence variant sits in a region of the genome closely linked to a trait of interest, such as a disease resistance locus, the variant can be used to predict the presence or absence of the resistance allele. The strength of the prediction will depend upon the closeness of the genetic linkage between the sequence variant and the target locus. Markers used in plant breeding programs fall into three broad categories: morphological, biochemical and DNA-based

1.8.2.1 Morphological markers

Morphological markers produce phenotypes which can be readily identified, but which are not usually of direct economic importance. Their value is due to close linkage with economically important traits. Brown *et al.* (1993) showed that the phenotype ‘pseudo-black chaff’ or ‘high-temperature-induced seedling chorosis’ was a visible marker used to identify the presence of the linked gene *Sr2* that conferred resistance from stem rust[58]. A PCR-based DNA marker for the detection of Sr2 is now available and is currently being used by breeding programs Australia wide.

1.8.2.2 Biochemical markers

Biochemical markers produce an enzyme or storage protein that can be identified by biochemical assay that is linked with a trait of economical importance. For example, glutenins are a major component of the storage protein in wheat in which there are two types, high molecular weight and low molecular weight subunits. Genes encoding for high molecular weight subunits of glutenin are located at the *Glu-A1*, *Glu-B1*, and *Glu-D1* loci on the long arm of chromosomes 1A, 1B and 1D while the low molecular weight glutenins, *Glu-A3*, *Glu-B3* and *Glu-D3*, are on the short arm of the same chromosomes. Proteins produced by these genes are polymorphic and readily detected by polyacrylamide gel electrophoresis, making these genes useful markers for linked genes on Group 1 chromosomes. As glutenins have a major influence on dough strength, a careful selection of combinations of different alleles at homologous Group 1 loci often predicts dough strength with reasonable accuracy. Thus plant lines are selected or discarded for grain quality based on this glutenin allele characterisation[59].

1.8.2.3 DNA-markers

DNA-markers identify molecular differences among genes determining traits of interest, or of DNA segments linked to genes determining the traits of interest. Many types of DNA-markers have been developed including restriction fragment length polymorphisms (RFLP's), allele-specific polymerase chain reaction markers (AS-PCR's) and microsatellite or simple sequence repeats (SSP's). These markers have all been used for marker-assisted selection of wheat in Australia. Table 1.2 shows DNA markers used for cultivar development of wheat for abiotic, biotic and quality traits.

Dholakia *et al.* (2001) used PCR-based DNA markers to identify markers linked to the grain protein concentration (GPC) of hexaploid wheat[60]. The GPC is an important factor that determines the end-product quality as well as playing a pivotal role in human nutrition. The study used 106 recombinant inbred lines (RILs) from a cross between two wheat cultivars PH132 and WL711, which differ significantly in GPC. The RILs were phenotyped for GPC at two diverse agroclimatic locations as described by Pune and Ludhiana. The parents were screened with 85 inter simple sequence repeat (ISSR) primers and 350 random primers. The selective genotyping

and whole population analysis revealed nine DNA markers associated with the trait. Three markers were observed to be associated with the trait in both locations, two markers were found to be specific to Pune, and four markers were specific to Ludhiana. This study clearly demonstrates the applicability of DNA-markers in finding regions on chromosomes associated with quantitative characters in wheat such as GPC.

1.8.2.3.1 Single-nucleotide polymorphisms (SNP's)

Another group of DNA markers are single-nucleotide polymorphisms (SNP's), the most frequent variations in the genome of any organism. SNP discovery approaches such as re-sequencing or data mining enable the identification of insertion deletion (InDel) polymorphisms. Bhattramakki *et al.* (2002) used 655 InDels that had been identified by resequencing 502 maize (*Zea mays*) loci across 8 maize inbreds (selected for their high allelic variation)[61]. Of these 502 loci, 433 were polymorphic, with InDels identified in 215 loci. Of the 655 InDels identified, single-nucleotide InDels accounted for more than half (54.8%) followed by two- and three-nucleotide InDels. A high frequency of 6-base (3.4%) and 8-base (2.3%) InDels were also observed. The value of indels as genetic markers was demonstrated in the same study by using InDels polymorphisms to map 22 loci in a B73 x Mo17 recombinant inbred population. This study clearly demonstrates that the discovery and mapping of InDels markers will position corresponding expressed genes on a genetic map. It also shows that insertion-deletion polymorphisms occur frequently and can be used as highly informative genetic markers.

1.8.2.3.2 Disease-linked resistance markers

Genes linked to a desired marker, such as disease resistance, can be used in plant breeding to predict the presence of the desired trait in progeny plants. Spielmeyer and Lagudah (2003) have showed that homologous group 1 chromosomes of wheat contain important genes that confer resistance to leaf, stem and stripe rusts, powdery mildew and Russian wheat aphid[62]. Marker-assisted breeding is being used in US rice breeding programs to enhance the development of rice cultivars for improved cooking quality and genetic resistance to rice blast disease[63]. The study shows that highly disease resistant cultivars possess undesirable cooking traits for US market classes. Using molecular marker technology to select those cultivars with both

improved blast disease resistance and good cooking quality can accelerate the breeding process by increasing selection efficiency[64]. Genotyping the germplasm collection gives the breeders additional information regarding the genetic background and diversity of the parental material and assists in variety verification and assessment of seed purity.

1.9 Molecular Marker Analysis

There has been numerous DNA-based genetic marker analysis methods developed over the past 25 years since their discovery. These include methods devised to look at restriction length fragment polymorphisms (RFLP), simple sequence repeats (SSR), random amplified polymorphic DNA detection (RAPD), amplified fragment length polymorphisms (AFLP), nucleic acid indexing and the restriction enzyme amplification display system. Table 1.3 outlines the most widely used marker detection systems, showing their advantages and disadvantages with an explanation of each in detail in the following sections.

1.9.1 Restriction Fragment Length Polymorphisms (RFLP)

Restriction Fragment Length Polymorphism (RFLP) detection refers to inherited differences in sites for restriction enzymes. These can include base changes for a target site that result in differences in the lengths of the fragments produced by cleavage with the relevant restriction enzyme. RFLP detection is routinely used for genetic mapping to link the genome directly to a conventional genetic marker. Botstein *et al.* (1980) showed that random single-copy DNA probes were capable of detecting DNA sequence polymorphisms when hybridized to restriction digests of an individual's DNA[65].

Table 1.2: DNA markers used for cultivar development of wheat in Australia.

Stress Tolerance		
Trait	**Market Information**	**Reference**
Flour colour Psy-A1e p,r,s Epsilon-cyclase	Flour colour Psy-A1e p,r,s Epsilon-cyclase	Huang *et al* 2007, Plant Physiology 142:1718-1727
Salt tolerance Nax2	SSR marker (csLinkNax2) within 5 cM of Nax2. Perfect marker derived from gene Nax2 (CsNax2), dominant.	Byrt et al 2007, Plant Physiology 143:1918-1928.
Aluminium tolerance ALMT-1	Markers based on ALMT1: Promoter coding region (CAPS) Introns. All very tightly linked and in some instances perfect (promoter).	Reman et al 2008, Theoretical and Applied Genetics 116:343-354 Sasaki et al 2006, Plant Cell Physiology 47:1343-1354.
Physiological Traits		
Rht-B1b/Rht-D1b Dwarfing genes	SNP marker derived from gene and promoter, perfect marker but difficult to assay.	Ellis et al 2002, Theoretical and Applied Genetics 105:1038-1042.
Tiller inhibition gene (tin)	SSR marker (gwm136), within 1cM, co-dominant.	Spielmeyer and Richards 2004, Theoretical and Applied Genetics 109:1303-1310.
Rust Resistance		
Lr37/Sr38/Yr17	NBS-LRR derived STS from alien segment, perfect marker, dominant.	Seah et al 2001, Theoretical and Applied Genetics 102:600-605.
Lr34/Yr18/Pm38	Perfect marker.	Lagudah et al 2009, Theoretical and Applied Genetics 119:889-898.
Lr46/Yr29	EST derived, tightly linked marker requiring restriction with enzyme.	
Sr46, Sr2	EST derived tightly linked marker	
SrR	AFLP derived tightly linked STS marker, dominant, amplifies from shortened chromosome (not sticky).	Mago et al 2002, Theoretical and Applied Genetics 104:1317-1324.
Sr31	RFLP derived tightly linked co-dominant marker, amplifies from	Mago et al 2002, Theoretical and Applied Genetics

	shortened chromosome (not sticky).	104:1317-1324.
Sr24/Lr24	AFLP derived tightly linked STS markers, dominant.	Mago et al 2005, Theoretical and Applied Genetics 111:496-504.
Sr26	AFLP derived tightly linked STS marker, dominant, amplifies from shortened chromosome.	Mago et al 2005, Theoretical and Applied Genetics 111:496-504.
Sr39	AFLP derived tightly linked STS markers (separate markers for R and S), can be used as co-dominant.	Mago et al 2009, Theoretical and Applied Genetics 119 (8): 1441-1450.
Sr22	STS marker, tightly linked to Sr22.	

Cereal Cyst Nematode		
Cre1 / Cre3	NBS-LRR derived tightly linked marker.	

Barley Yellow Dwarf Virus Resistance

Trait	**Market Information**	**Reference / Contact**
BYDV2 / BDV3	PCR SCAR for the resistance-carrying translocation.	Stoutjesdijk et al 2001 Aust J.Ag Res 52: 1383 – 1388.

Quality		
ɣ-gliadin	SNP perfect markers derived from genes.	Zhang et al 2003, Theoretical and Applied Genetics 107:130-139.
GluA3 alleles a-g	SNP perfect markers derived from genes.	Zhang et al 2004, Theoretical and Applied Genetics 108:1409-1419.
Glu-1Bx7OE	Co-dominant, perfect marker.	Butow et al 2003, Theoretical and Applied Genetics 107:1524-1532.
Flour colour Psy-A1e p,r,s Epsilon-cyclase	SNP based co-dominant perfect markers, multiplexed CAPS Marker, CAPS Marker.	Howitt et al Funct Integr Genomics (2009) 9:363–376.

Source: CSIRO Wheat Markers (2011)[66]

Each of these probes defined a single locus and loci can be arranged into linkage groups to form a true genetic map of "DNA marker loci." Pedigrees in which inherited traits are known to be segregating can then be analysed, making possible the mapping of the gene(s) responsible for the trait with respect to the DNA marker loci, without requiring direct access to a specified gene's DNA[67].

1.9.2 Simple Sequence Repeat (SSR)

SSR analysis was used by Weber *et al.* (1989) to study interspersed DNA elements of the form (dC-dA)n.(dG-dT)n[68]. These repeats constitute one of the most abundant human repetitive DNA families. The study reported that specific human (dC-dA)n.(dG-dT)n blocks are polymorphic in length among individuals and therefore represent a vast pool of potential genetic markers[69]. Comparison of sequences from the literature for (dC-dA)n.(dG-dT)n blocks cloned two or more times revealed length polymorphisms in seven of eight cases[70]. Variations in the lengths of 10 (dC-dA)n.(dG-dT)n blocks were directly demonstrated by amplifying the DNA within and immediately flanking the repeat blocks by using PCR and then resolving the amplified DNA on polyacrylamide DNA sequencing gels[71]. SSR's are also referred to as Variable Number Tandem Repeats (VNTR) that includes sets of tandemly repeated pairs flanked by conserved restriction enzyme sites[72]. SSR's are also refereed to as minisatellites[73], microsatellites[74], di- and tri-nucelotide repeats and Short Tandem Repeats (STR)[75].

1.9.3 Random Amplified Polymorphic DNA (RAPD)

Similar to RFLP detection, RAPD detection uses random DNA segments with single primers of arbitrary nucleotide sequence. These polymorphisms, simply detected as DNA segments that amplify from one parent but not the other, are inherited in a Mendelian fashion and can be used to construct genetic maps in a variety of species[76]. Hahn *et al.* (2003) has shown that RAPD has been successfully used to detect genetic variations among isolates of *Paracoccidioides brasiliensis*[77]. The group investigated the applicability of RAPD in revealing important intrinsic and extrinsic features of the fungus associated with geographical origin, time of isolation, source of clinical specimen, clinical forms of human disease and also *in vitro* and *in vivo* susceptibility to antimicrobial and antifungal drugs[78]. The RAPD patterns allowed them to distinguish all of the analysed strains, which included 26 clinical isolates, 2

animal isolates and 1 environmental isolate of *P. brasiliensis* obtained from different geographic regions, confirming the strong discriminating power of RAPD technology[79].

1.9.4 Amplified Fragment Length Polymorphism (AFLP®)

The Amplified Fragment Length Polymorphism (AFLP®)* technique is based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA. Described by Vos *et al.* (1995) the technique involves three steps: (i) restriction of the DNA and ligation of oligonucleotide adapters, (ii) selective amplification of sets of restriction fragments and (iii) gel analysis of the amplified fragments[80]. PCR amplification of restriction fragments is achieved by using the adapter and restriction site sequence as target sites for primer annealing. The selective amplification is achieved by the use of primers that extend into the restriction fragments, amplifying only those fragments in which the primer extensions match the nucleotides flanking the restriction sites. Using this method, sets of restriction fragments may be visualised by PCR without knowledge of nucleotide sequence. The method allows the specific co-amplification of high numbers of restriction fragments. The number of fragments that can be analysed simultaneously, however, is dependent on the resolution of the detection system[81]. Typically 50-100 restriction fragments are amplified and detected on denaturing polyacrylamide gels[82]. The AFLP® technique provides a novel and very powerful DNA fingerprinting technique for DNA of any origin or complexity[83].

AFLP is a registered trademark of Keygene N.V. (www.keygene.com).

1.9.5 Nucleic Acid Indexing

Unrau *et al.* (1994) describes a highly systematic, non-cloning method of distinguishing and isolating every fragment in a class-IIS or interrupted palindrome restriction[84]. These enzymes produce informative, non-identical cohesive ends that can be selectively modified by ligation to individual synthetic oligonucleotides with the corresponding complementary ends[85]. In this way, PCR and sequencing primer sites and labels can be introduced specifically into a single fragment in a total genomic digest. The Unrau *et al.* (1994) study used known and unknown fragments from *Escherichia coli* and isolated fragments directly in sequencable form without the necessity of synthesising unique primers[86]. The group isolated Human DNA in this

way avoiding the problems intrinsic to cloning (selective fragment loss, mutation and sequence rearrangement)[87]. Systematic characterisation of DNA fragments by their cohesive ends and length provides tremendous power and flexibility for analysis of any DNA molecule without specific clones, probes or libraries[88]. Further applications include DNA sequence tagged site and restriction mapping, sequencing, RFLP analysis and DNA diagnostics[89].

1.9.6 Restriction Enzyme Amplification Display System (READS)

Described by Prashar *et al.* (1996) READS was used to study changes in gene expression by selective PCR amplification and display of 3' end restriction fragments of double-stranded cDNAs[90]. This method produces highly consistent and reproducible patterns, can detect almost all mRNAs in a sample and can resolve hidden differences such as bands that differ in their sequence but co-migrate on a gel[91]. Bands corresponding to known cDNAs move to predictable positions on the gel, making this a powerful approach to correlate gel patterns with cDNA databases. Applying this method, Prashar *et al.* (1996) examined differences in gene expression patterns during T-cell activation[92]. Of a total of 700 bands that were evaluated in this study, as many as 3-4% represented mRNAs that are up-regulated, while 2% were down-regulated within 4 hr of activation of Jurkat T cells[93]. These and other results suggest that this approach is suitable for the systematic, expeditious and nearly exhaustive elucidation of subtle changes in the patterns of gene expression in cells with altered physiologic states[94].

Table 1.3. Comparison of the most widely used marker systems for wheat

Marker System	Loci detected per assay	DNA quantity	Advantages	Disadvantages
RFLP	3	5 µg	Highly reliable, co-dominant, serve as reference to other cereal maps, large number available, target specific regions	Technically difficult, slow, requires large amount of DNA, detect low levels of polymorphisms
SSR	1	0.2 µg	Reliable, co-dominant, often genome-specific, target specific regions, amenable to automation	High development cost
AFLP	50	0.2 µg	Reliable, detect large number of loci simultaneously, amenable to automation	Random, dominant
RAPD	10	0.2µg	Cheap, technically simple	Unreliable, dominant

Source: Langridge *et al*., (2001)[95].

1.10 Bulk Segregant Analysis (BSA)

Plants from segregating populations can be grouped (bulked) according to phenotypic expression of a trait and tested for differences in allele frequency between the population bulks[96]. Probes, such as molecular markers, can be used to show polymorphisms between parents of a segregation or between extreme phenotypes in a group of populations. BSA is a way of finding marker-trait associations, both for quantitative and qualitative characters for any specific gene or genomic region[97]. Michelmore *et al* (1991) first reported the BSA method to identify markers linked to disease resistance genes based on the principle of marker allele frequency differences at the extremes of phenotypic distribution as a result of linkage to the trait[98]. The study examined two bulked DNA samples that were generated from a segregating population from a single cross. Each bulk contained individuals that were identical for a particular trait or genomic region, but arbitrary at all unlinked regions[99]. The two bulks are therefore genetically dissimilar in the selected region but seemingly heterozygous at all other regions[100]. The bulks were screened for differences using RFLP probes or RAPD primers. The study identified three RAPD markers in lettuce linked to a gene for resistance to downy mildew[101]. The efficiency of marker identification through BSA has been shown to be lower than through Nearly Isogenic Line (NIL) Analysis. However, BSA has the advantage that genetic walking is possible to identify markers in specific regions of the genome using multiple rounds of BSA, where each new pair of bulks will differ at a locus identified in the previous round of analysis. BSA is most useful for Quantitative Trait Loci (QTL) analysis that is aimed primarily at a single quantitative trait, though results of BSA can be strongly affected by any dominance relationship at the trait locus[102].

1.11 Recombinant Inbred Line Analysis

Recombinant inbred lines (RILs) are created by crossing 2 inbred strains followed by repeated selfing to create a new inbred line, whose genome is a mosaic of the parental genomes[103]. As each RIL is an inbred strain, it can be propagated eternally and can be used for genetic mapping. Tiwari et al (2009) studied wheat germplasm with high grain iron (Fe) and zinc (Zn) concentrations to understanding the genetic basis of their accumulation[104]. The manipulation of these micronutrients in food crops is a good approach for alleviating the micronutrient deficiencies in hunman diets.

One accession of *Triticum boeoticum* (pau5088) that had relatively higher grain Fe and Zn was crossed with *Triticum monococcum* (pau14087), and a RIL population generated[105]. The grains of the RIL population were evaluated for Fe and Zn concentration and a linkage map available for the population was used for mapping quantitative trait loci (QTL) for grain Fe and Zn accumulation. The QTL analysis led to the identification of 2 QTL for grain Fe on chromosomes and 1 QTL for grain that were used in molecular breeding programs[106].

1.12 Diversity Arrays Technology

Numerous DNA-based genetic marker analysis methods have been developed over the last two decades as described in section 1.9. While these genotyping methods have contributed greatly to our current understanding of genome organisation and genetic variation, they are constrained by their dependence on gel electrophoresis, resulting in low throughput. Some of these methods, SSR for example, require pre-identification of a polymorphism or a potential site before analysis of other individuals is possible. Furthermore, all methods based on size separation of multiple DNA fragments suffer from difficulties in precisely correlating bands on gels with allelic variants.

To overcome many of these restrictions, Jaccoud *et al.* (2001) developed a hybridisation-based method using nucleic acids immobilised on solid-state surface[107]. DNA chips or microarrays, have been developed to analyse genotypes for single nucleotide polymorphisms (SNPs). These minor, but abundant differences in DNA sequence among genotypes are identified through an expensive and laborious DNA sequencing process. SNPs promise to revolutionise biomedicine, but the technology depends on intensive genomic sequencing and the high cost of analysis that cannot be matched in agriculture or basic research. Jaccoud *et al.* (2001) reports the development of a solid-state, open-platform method for DNA polymorphism analysis called Diversity Array Technology (DArT).[108] Genetic marker analysis through DArT offers a low-cost, high-throughput, robust system with minimal DNA sample requirement capable of providing comprehensive genome coverage even in organisms without any DNA sequence information.

To test the potential of the DArT high-throughput genome analysis method, Jaccoud *et al.* (2001) tested the application of the microarray technology platform with the analysis of DNA polymorphisms. Using the rice genome as a model, the group assayed for the presence (or amount) of a specific DNA fragment in a representation derived from the total genomic DNA of an organism or a population of organisms. Two different approaches were presented; the first involves contrasting two representations on a single array while the second involves contrasting a representation with a reference DNA fragment common to all elements of the array. The Diversity Panels created using this method allow genetic fingerprinting of any organism or group of organisms belonging to the gene pool from which the panel was developed. DArT enabled the rapid and economical application of a highly parallel, solid-state genotyping technology to any genome or complex genomic mixtures.

1.12.1 The DArT procedure

Genomic DNA was extracted from the organism in question, in the case of Jaccoud *et al.* (2001) young seedlings from rice. Approximately 5 ng of DNA from each cultivar was bulked and digested using a pre determined restriction enzyme (Figure 1.7). After digestion, an enzyme-specific adapter was ligated to the DNA fragments. The mixture was diluted and used as template in a PCR reaction. The amplicons generated from the PCR reaction were ligated into the PCR2.1-TOPO vector and transformed into heat-shock competent *Escherichia coli* cells. Transformants were selected on medium containing ampicillin and X-gal. Individual white colonies (containing recombinant plasmids) were transferred into 10% glycerol. From each glycerol sample, an aliquot was transferred to a PCR mix containing forward and reverse primers. After amplification, the PCR products were precipitated and the DNA was resuspended in printing buffer. The products were then arrayed onto glass slides using a microarrayer.

Genomic representation targets were made using the same steps as above, except instead of cloning into the TOPT-vector, representations were precipitated and labeled with fluorescent dye. The labeled representations were mixed and dissolved in hybridisation solution. The target solution was pipetted directly onto the microarray surface and covered with a glass cover slip. Slides were placed into a humidification chamber and incubated overnight. After hybridisation, the cover slips were removed and slides were washed and dried. Slides were scanned using a fluorescent

microarray scanner and spot signal intensities and background intensity recorded. Raw data was analysed using in-house software and commercial microarray software.

From the Jaccoud *et al.* (2001) study, it was shown that the DNA microarray platform was successfully adapted to DNA polymorphism analysis. DArT, which is not reliant on DNA sequence information, has the potential to include applications such as germplasm characterisation, genetic mapping and gene tagging, molecular marker-assisted breeding and tracking genome methylation changes. By using composite diversity panels to resolve the complex genomic samples into respective components, DArT offers genotyping in parallel with pathogen or endosymbiont detection and characterisation.

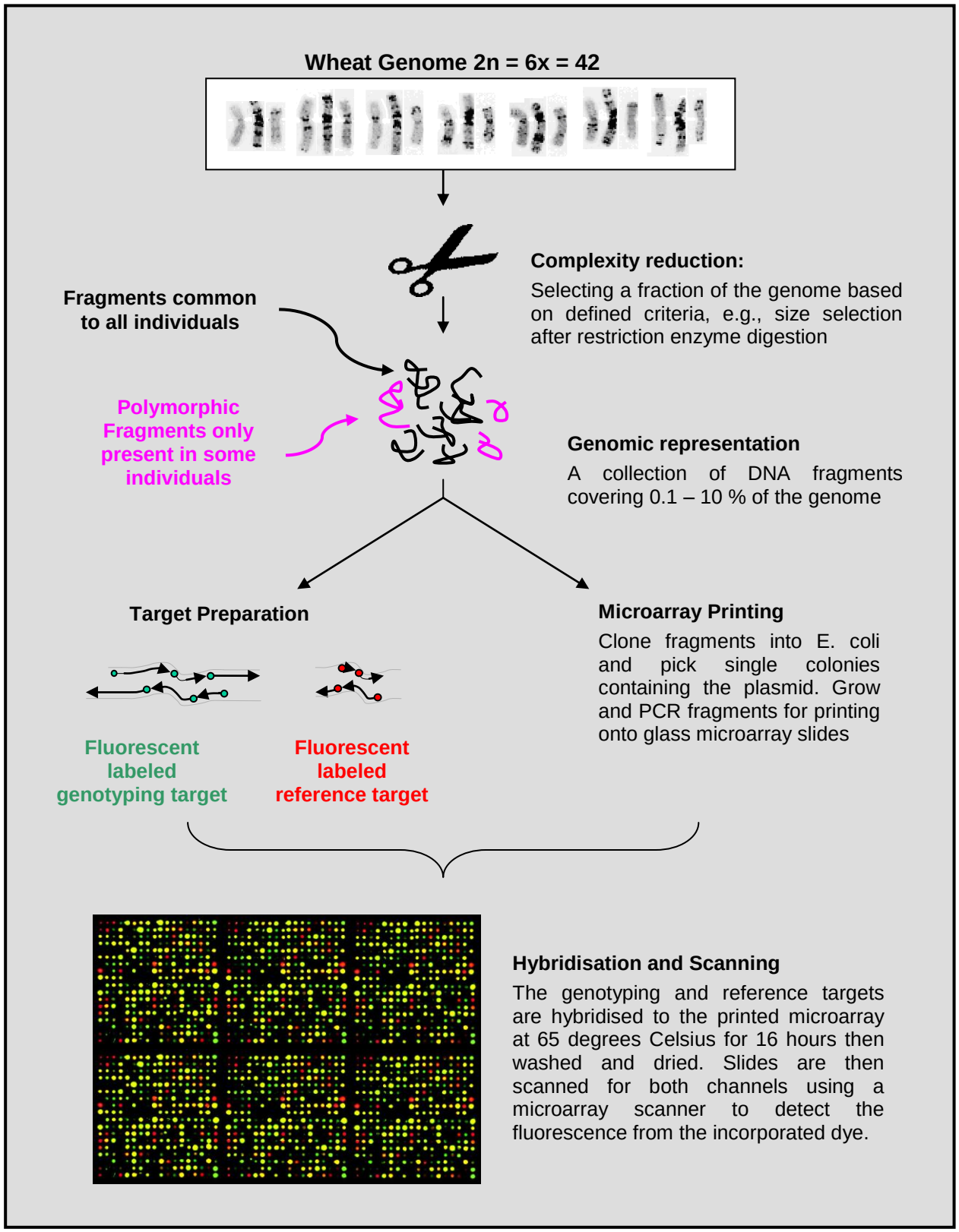


Figure 1.7: DArT Flow Chart: From genome to data.

Source and copyright: Diversity Arrays Technology Pty. Ltd.

1.12.2 DArT and Methylation

Methylation, as described in the following chapters, is the addition of methyl groups to the DNA molecule. The methyl group, CH_4, can be detected using DArT in the complexity reduction stage, where restriction enzymes are chosen that will either cut at a specific site or will be blocked at a specific site. Enzymes, such as McrBC, are blocked by the presence of a methyl site, thus by comparing a genomic representation that has been treated with McrBC versus one that has not, it is possible to observe the difference in DNA fragments within both representations.

1.13 Epigenetics and DNA Methylation

Several mechanisms have been described in the scientific literature relating to the inheritance, evolution and molecular biology of epigenetics and DNA methylation.[109] [110] A primary such epigenetic mechanism involves post-replicative covalent modification of DNA by methylation of cytosine bases. Typically, the modification in plants is methylation of cytosine bases in the dinucleotide CpG and the trinucleotide CpNpG (where N could be any of the four nucleotide bases). Because the methylated sequence is palindromic, both strands of DNA can be methylated. The modification is inherited epigenetically because of the existence of a system that recognises hemimethylated sequences (with one strand modified) and converts them to the fully methylated state (with both strands modified). Systems also exist to reverse the methyl group by removing the methyl group. It is a paradox of conventional genetics that two alleles can have the same genetic sequence but show different states of inheritance. A methylated sequence is frequently not expressed while the same sequence is expressed when unmethylated.

In the past, genetic variation was considered to be due to allelic and epistatic combinations that owed their existence to alterations in the primary nucleotide sequence of the respective genes. Nucleotides modified by methylation result from post-replicative events and were usually not considered to be a part of the primary nucleotide sequence of an individual. However, since DNA methylation occurs at defined target sequences (mainly CpG and CpNpG) and not all target sites are methylated, it represents a potentially important form of polymorphism. In this way epigenetic information systems, like DNA methylation, could produce alleles, called

epialleles that could generate epigenetic variation that had never been considered as cause of phenotypic variation.

1.13.1 Epigenetics in Plant Breeding

Epigenetic phenomena, specifically DNA methylation, is associated with all the important steps of a conventional plant breeding program, e.g., the creation of favorable genetic variation that will form the basis for subsequent selection schemes or used directly as F1 hybrid seed in hybrid breeding; the selection of superior genotypes through their phenotypes in the field; the multi-site multi-year testing of putative new cultivars and estimation of their adaptation and the preservation and stability or even further improvement of pure line cultivars. This is also true for breeding programs using more modern methods, such as cell and tissue culture and plant genetic engineering. Variation observed in some clones, the unexpected silencing through methylation of certain transgenes inserted into plants and the silencing through methylation of endogenous plant sequences homologous to the transgene, have raised serious problems for those wishing to exploit transgenic plants. However, research on these areas has also helped to further the understanding of epigenetic phenomena involved in the regulation of gene action, allelic and epistatic interactions of genes, variation in plant somatic cells and plant virology.

1.13.2 Epigenetic occurrence

Plant genomes are generally more methylated compared to other eukaryotic genomes. More than 30% of cytosine bases in some plant genomes are methylated in certain tissues and/or certain developmental stages,[111] while in most vertebrates, less than 10% of all cytosines are methylated[112]. In addition to the methylation of cytosine in some CpG dinucleotides, the only methylation sites of animals and other eukaryotes, plant genomes contain 5mC methylation in the trinucleotide sequence CpNpG, where N is any of the four DNA bases[113]. The high proportion of methylated C residues in plants compared to animals could be due to angiosperm genomes contain a higher proportion of CpG dinucleotides and due to differential degrees of CpG depletion. According to the deamination theory, 5mC residues can undergo deamination to thymine, which leads to point mutations and to depletion of CpG dinucleotides and subsequently to an increase of TpG and CpA dinucleotides[114].

Depletion levels are determined by comparing the numbers of observed CpG dinucleotides with the level expected from the base composition. In animals this ratio varies between 0.15% and 0.35%[115] [116], while dicot and monocot genomes vary between 0.68% and 0.79% respectively[117].

The degree of DNA methylation varies among individual plant species. Arabidopsis thaliana contains only 6.3% cytosine methylation, one-fourth of the methylation level of most other angiosperms[118]. In tomatoes, 85% of the CpNpG sites were methylated[119]. With the degree of DNA methylation varying among plant species, DNA methylation also varies depending on tissues or developmental state. For example in carrot, a differing content of 5mC was observed among different tissues[120]. In tomato, the level of methylation of mature tissues was significantly higher than that of immature ones and protoplasts[121]. Moreover, a reversible variation in the methylation pattern was observed during the process of carrot somatic embryogenesis. Thus, as in mammals, a cycle of demethylation and de novo methylation appears to take place during plant development and differentiation[122].

1.13.3 DNA Methylation in Plants through evolution

CpG and CpNpG methylation was surveyed by Belanger *et al.* (1990) in a range of non-vascular and vascular plants to determine when CpNpG methylation evolved and whether the two methylation systems found in higher plants were likely to be under common or separate control. They discovered that although both systems exist in many vascular plant taxa, the nonvascular plant taxa appear to contain only CpNpG methylation and this in only limited amounts. The data suggest that both systems may have evolved at the same time and that speciation involved the loss of one or the other methylation systems, or involved the evolution of a stage-specific control system operating during differentiation. Thus the extra methylated CpNpG sequence found only in plants is not a recent acquisition of the plant kingdom[123].

1.13.4 DNA Methylation classification: Dam, Dcm and CpG

DNA methyltransferases transfer a methyl group from S-adenosylmethionine to either adenine or cytosine residues and are found in a wide variety of prokaryotes and eukaryotes. There are various classes of DNA methylation determined by the base

that is methylated and at what nitrogen (N) or carbon (C) atom it is methylated. Studies of laboratory strains of *Escherichia coli* have found three site-specific DNA methylases. The methylase encoded by the *dam* gene (Dam methylase) transfers a methyl group from S-adenosylmethionine to the N6 position of the adenine residues in the sequence GATC[124] [125]. The Dcm methylase is encoded by the *dcm* gene and methylates the internal cytosine residues in the sequences CCAGG and CCTGG [126] [127] at the C5 position. The *EcoKI* methylase, M. EcoK I, modifies adenine residues in the sequences AAC(N6)GTGC and GCAC(N6)GTT. *EcoKI* sites are found approximately once every 8 kb where as Dam sites occur approximately once per 256 bp and Dcm sites once every 512 bp in DNA of random sequence (GC=AT). CpG methyltransferases, found in higher eukaryotes, transfers a methyl group to the C5 position of cytosine residues. Patterns of CpG methylation are heritable, tissue specific, and correlate with gene expression.

1.13.4 DNA Methylation detection

The modified base 5mC was initially detected using chromatographic techniques, however a number of other methods have been developed. For gross comparative quantitation of genome methylation, isoschisomeric restriction enzymes that recognise and cut similar nucleotide sequences of DNA, which only differ in cytosine methylation, have been used[128]. The isoschisomer pairs MspI/HpaII and EcoRII/BstNI, which recognize 4bp and 5bp nucleotide stretches, are routinely used. The restriction enzymes MspI and HpaII have the same recognition site C*CGG, where the * indicated the cleavage site. Both MspI and HpaII cannot cleave the sequence if the external C in the sequence is methylated, however MspI can cleave the sequence when the internal C residue is methylated[129]. Any difference in fragment sizes generated by these two enzymes should thus be due to differences in methylation at the CpG site. The restriction enzyme McrBC cleaves DNA containing methylated cytosine on one or both strands in the presence of GTP[130]. McrBC recognises the half-sites (G/A)mC, where these half-sites can be separated by up to 3 kb, with the optimal being 55-103 bps[131]. Analysis of genomic digests treated and non-treated with McrBC in addition to a frequent cutter(s) allows for discrimination of DNA methylation. In this way, MspI/HpaII and separately, McrBC restriction enzymes, will be used during the complexity reduction methods of DArT for methylation polymorphism detection. Polymorphic fragments will be identified and

used as a source of epigenetic variation to study differences in wheat cultivars, tissues types, development stages and growth conditions.

The development of PCR has allowed for other techniques to identify DNA methylation and to map methylation polymorphisms. One such technique is based upon the coupled restriction enzyme digestion and random amplification (CRED-RA) of genomic DNA[132]. Random amplification of genomic DNA by PCR with arbitrary 10-mer oligonucleotide primers is widely used to generate random amplification polymorphic DNA (RAPD) markers for fingerprinting or genetic mapping[133]. The CRED-RA technique is based on the following hypothesis: a DNA fragment cannot be amplified if it contains a specific restriction site in the region between two primer binding sites and that site is cut by restriction enzyme digestion prior to PCR. If DNA methylation of the restriction site prevents digestion within the genomic fragment, the fragment can be amplified. However the amplified product will then be susceptible to cleavage because the restriction site will not be methylated during DNA amplification. Thus, DNA methylation can be identified by comparing the banding patterns of template DNA amplified without restriction, template DNA amplified after restriction and product DNA restricted after amplification. The technique has been used by Bedford and Van Helden (1990) to detect allelic differences in methylation[134] and by Tsaftaris et al.(1997, 1998) to study the variation in patterns of DNA methylation among maize inbreds, and between maize inbreds/hybrids from plants grown under different conditions[135].

A further methylation detection method combines PCR with sequencing and bisulfite treatment to modify cytosine to uracil residues in the DNA. All cytosines are converted to uracil, except those that are methylated, which are resistant to modification and remain as methyl cytosine[136]. Each altered DNA sample must then be amplified, cloned and sequenced. The main disadvantage of this technique is that it is technically difficult and labor intensive, but for a single structural gene and its upstream few thousands bp of regulatory sequences, it can provide a complete map of methylated sites in different tissues and developmental stages. A recent modification of this procedure takes advantage of the bisulfite-mediated chemical conversion of cytosine to uracil, followed by PCR using primers designed to distinguish methylated from unmethylated DNA. The main advantage is that it avoids

the use of restriction enzymes and resulting problems associated with incomplete digestion.

For fast analysis of the methylation state of thousands of genes simultaneously, a powerful new technique called Restriction Landmark Genomic Scanning (RLGS) has been recently developed[137]. RLGS is a multiplex method that allows simultaneous analysis of more that 3,000 loci. It employs the NotI restriction enzyme because its restriction site makes good landmarks for genetic analysis. NotI cuts neither GCGG5mCCGC nor GCGGC5mCGC, but cuts GCGGCC5mC and it is blocked by CpG methylation. The technique uses high-resolution, 2-D electrophoresis to visualize radioactive DNA fragments produced by restriction digestion.

Southern blot analysis of DNA digested with isoschisomeric restriction enzymes that have different sensitivities to recognition site methylation has been used to localize methylation in the genome. This technique allows cytosine methylation associated with specific genes or specific regions of DNA such as repeat sequences to be identified. This procedure is laborious, requires specific probes and does not always identify DNA methylation mutants or polymorphisms because it cannot always discriminate between cytosine methylation at a restriction site and loss of the site due to nucleotide mutation.

For more accurate quantitative determinations of DNA methylation, different types of High Performance Liquid Chromatography (HPLC) as well as mass spectroscopy[138] [139] have been employed for amounts as small as 5-10 µg DNA. Vilpo et al. (1986) employed immunological techniques, generating specific antibodies against 5mC, to measure DNA methylation[140]. Unfortunately, none of the above techniques can provide information about the location of methylated nucleotides in the genome.

1.13.5 DNA Methylation and Gene Expression

Several lines of evidence suggest that DNA methylation in eukaryotes plays a role in gene expression. Studies of numerous tissue-specific genes using different techniques have shown a clear correlation between the methylation status of active and inactive genes. Thus, findings suggest that most genes are undermethylated in

tissues in which they are expressed, while they are heavily methylated in non-expressing tissues[141]. This data suggest that changes in the methylation pattern during differentiation may modulate gene activity. In plants, induction of several endogenous genes in certain tissues has been linked to loss of cytosine methylation[142 143 144], correlating hypomethylation of these genes with transcription in the respective tissues. Although such examples suggest a correlation between gene repression and DNA methylation, other reports can be found in the literature that do not detect any changes in DNA methylation patterns although gene activity is altered. Certainly not all changes in gene activity are based on regulation of gene transcription by DNA methylation. The literature suggests that it is most likely that DNA methylation is mainly involved in the regulation of promoter activities, but not in post-transcriptional regulation. If changes in gene activity are due to post-transcriptional regulation, promoter activity would probably not be impaired and no significant changes in DNA methylation should be detectable. Even for transcriptional control it is difficult to exclude the involvement of DNA methylation in changing promoter activity, because most DNA methylation studies have limited accuracy as they frequently use isoschisomers. The state of DNA methylation at a restriction site might not always correspond to the degree of methylation of a neighboring sequence that is involved in promoter regulation. For example, genomic sequencing analysis provides a precise tool, as the methylation state of every C residue can be analysed. Genomic sequencing of a 900-bp region upstream from the translation start codon of the maize alcohol dehydrogenase gene did not reveal any cytosine methylation although the gene was silenced[145]. Apparently the Adh1 gene provides an example where gene activity is not regulated by DNA methylation. It cannot be excluded, however, that changes in DNA methylation further upstream of the promoter region may have an influence on repression of the gene. For example, the cell-specific transcription of the PEPCase gene of a C4 plant like maize corresponds to demethylation of a region located 3.3 kb upstream of the gene[146]. *In vitro* DNA methylation of a few specific gene sequences inhibited the activity of these genes when inserted into animal cells *in vivo* [147]. In plants Weber et al. (1990) showed that *in vitro* hemimethylation of the CaMV 35S promoter inhibited transient gene expression of reporter genes after transfection into protoplasts, the methylated state was maintained and inherited during regeneration of plants and correlated with inhibition of transgene expression[148].

Evidence for the involvement of DNA methylation in preventing the expression of certain plant genes has been provided by studying the expression of genes within plant transposable elements. The pioneering work of Barbara McClintock[149 150] and Peter A. Peterson[151 152] with maize revealed that transposons undergo reversible heritable inactivation. Early in the study of the suppresor-mutator (Spm) transposable element, McClintock recognised that certain isolates of the element either cycled between inactive and active phases during development or underwent an inactivation event of longer duration and sufficient stability to be heritable, but which was nonetheless occasionally reversed [64 65]. In subsequent studies, she developed a deeper understanding of the ways in which the Spm element alternated between active and inactive phases[153]. She later reported that the Activator element was also subject to a similar type of reversible inactivation, although the Activator element's inactivation mechanism was not analysed in detail[154 155].

Finally, strong evidence for the role of DNA methylation in modulating plant gene expression has been more recently obtained from studies of transgenic plants. By introducing extra copies of a specific gene, one might expect in many cases to overproduce the corresponding mRNA and protein products. Conversely, attempts at silencing genes have often employed an antisense strategy of expressing single-stranded RNA from the noncoding strand of a gene to bind to the mRNA, thereby preventing accumulation of the corresponding protein. Although these techniques have been successful in numerous applications, a body of literature is emerging that documents cases with unexpected outcomes in organisms as diverse as nematodes and plants. These observations encompass transgene silencing, i.e., failure to express certain transgenes. In some cases not only the transgenes introduced at ectopic positions in plant genomes can be unpredictably silenced, but also if the ectopic sequences are homologous to endogenous plant genes, silencing of the endogenous gene can frequently occur[156 157 158]. Transgene epigenetic inactivation has provided clear-cut evidence for the involvement of DNA methylation in gene action. The involvement of epigenetic phenomena in unpredictable transgene inactivation in transgenic plants has also attracted the attention of the scientific community to epigenetics for practical reasons. Plant cultivars bred for specific characteristics, e.g., herbicide tolerance, in the laboratory, may lose this character when cultivated in farmer's fields. In addition, analysis and understanding of specific cases of allelic gene-gene interactions and inactivation of the transgene through DNA

methylation helped our understanding of similar types of naturally occurring phenomena such as paramutation and viral resistance of plants.

Research on structural plant genes, transposon genes, and transgenes all point to a significant role of DNA methylation in gene transcription. The most direct mechanism by which DNA methylation could interfere with transcription would be to prevent binding of the basal transcriptional machinery to promoters. This is not a generally applicable mechanism because some promoters are transcribed effectively as naked DNA templates independent of DNA methylation. Certain transcription factors bind less well to methylated recognition elements, however the reduction in affinity is often insufficient to account for the inactivity of promoters *in vivo.* It seems unlikely that DNA methylation would function to repress transcription globally by modifying the majority of CpGs in a chromosome, if the only sites of action are to be a limited set of recognition elements for individual transcription factors. The second possibility is that specific transcriptional repressors exist, that recognise methyl-CpG and either independently or together with other components of chromatin, turn off transcription. This mechanism would have the advantage of being substantially independent of DNA sequence itself, thereby offering a simple means of global transcriptional control. It would be especially attractive if the methylation-dependent repressors work in a chromatic context because then DNA could maintain the nucleosomal and chromatin fiber architecture necessary to compact DNA. Moreover, because chromatin assembly also represses transcription, methylation dependent repression mechanisms would add to those already in place.

1.13.6 Implications of epigenetics for plant breeding

Recognition that the concept of heredity has to be extended to incorporate epigenetic inheritance systems (EIS) is likely to have major impacts on plant breeding. The theory of selection is based on the existence of heritable variation that affects performance. A theory of variation is therefore a fundamental part of a theory of selection and will determine its efficiency, its limits, and the end result. The present theory is based largely on the assumption that heritable variation is random and involves changes in DNA sequences. If some variation is not based on sequence change but rather is epigenetic, which in addition, is affected by the environmental conditions of plant growth, this must modify and complement breeding theory. By inclusion of inherited epigenetic information as a source of variation, the

interpretation of some breeding results could be different, and often simpler. Conventional breeding is a Mendelian approach in the sense that the genotypic merit is assessed from the phenotypic expression. In other words, genotypic values have to be assessed from phenotypic values, and this requires knowledge of the conditions that ensure the best correspondence between genotype and phenotype.

1.13.7 DNA Methylation and Genetic Variation.

DNA methylation, in addition to being a cause of epigenetic variation, could be the cause of mutation and generation of genetic variation. Methylated cytosines are hot spots for mutations since 5mC frequently deaminates to T[159] [160]. The heavy mutational burden induced by methylation of C that could be seen as either a mutagenesis system not requiring the use of exogenous mutagens and occurring in non replicating DNA, or as the price that just be paid for employing a 5mC epigenetic system. The mutability of 5mC was first demonstrated in E. coli[161]. Cytosine bases that were methylated in the E. coli *lacI* gene were found to be hot spots for spontaneous base substitution mutations and the hot spots disappeared when the same sites were unmethylated[162]. It was speculated that the reason for this increase was that, whereas C deaminates to uracil (U), 5mC deaminates to T, which is a normal DNA base and therefore inherently more difficult to repair[163] [164]. In vertebrates, the presence of high levels of CpG methylation was associated with significant deamination of 5mC to T, a change that was incompletely or inefficiently repaired[165] [166]. Thus, where a 5mCpG dinucleotide pair was initially present in a gene, the deamination process would convert this into a TG/CA dinucleotide pair. Presently, mutations at CpG sites continue to play a significant role in the formation of new germ-line mutations contributing to genetic disease. Cooper and Krawczak (1990), in a survey of a wide variety of genetic diseases, found 44 of 139 (32%) point mutations were C to T or G to A transitions occurring at CpG dinucleotides[167]. The isolation of tumor suppressor genes and the detection of mutations within them in somatic cells, has led to the realisation that 5mC is a frequent contributor to mutations relevant to human carcinogenesis[168].

Chapter 2

Materials and Methods

“Play for more than you can afford to lose

and you will learn the game”

Winston Churchill

2.0 General Methodology

2.1 Introduction

The general methodology described in this section will include all plant growth techniques and DArT procedures. Specific divergence, changes or improvements to these general methods will be described in the results chapters under 'specific methodology', relating to experiment-specific methods.

2.2 Justification for methodology

DArT technology has focused extensively on sequence-based approaches for studying the diversity between cultivars/breeds/cultivars of many organisms. For wheat, hundreds of experiments have been performed to generate a targeted wheat microarray where clones have been chosen for their consistency, reliability and discriminating ability for certain traits and cultivars of wheat. The research performed and presented in this thesis, will not only look at DNA sequence differences, but will be expand to include methylation differences between cultivars, tissue types, developmental stages and stress responses, as to determine whether this source of diversity can complement existing molecular marker technologies. Apart from DNA sequence differences that affect certain crop traits, such as yield, growth patterns, disease resistance, salinity tolerance etc, there are differences among individual communities of organisms that have evolved to cope with changing environments, including reduced water availability and saline soils. This diversity will play an important role in modern agriculture, due to worldwide increases in the demand for food, requiring more land for cultivation and spreading agricultural land to regions where traditional crops are less suited to the local environment. The data and conclusions presented in this thesis will use DArT methods to develop markers to aid in breeding better adapted crops that are produced through molecular-based breeding programs, as apposed to genetic modification, to address the future agricultural needs and further research into diversity in wheat.

2.3 Plant growth methodology

It has been suggested that induced stress can lead to changes in methylation patterns in plants, leading to better adaptation to changing environments through natural selection. We have devised a set of experiments where plants have been grown at different conditions, as described below under 'light and temperature stress' and 'salt stress'.

2.3.1 Light and Temperature Stress

Experiments were developed to grow hexaploid bread wheats at three temperatures and at two lighting levels. All plants were germinated at room temperature and allowed to grow for one week before being transferred to soil and grown at one of six conditions until they matured. Plants were grown in growth rooms and cabinets located at the Research School of Biological Sciences (Australia's National University) and in the Centre for Agricultural and Molecular Biology to International Agriculture (CAMBIA) building at the Crown Scientific, Industrial and Research Organisation (CSIRO) Black Mountain, Canberra, Australia.

2.3.1.2 Temperature Conditions

Plants were grown in temperature controlled rooms/cabinets at 10°, 20° and 30°Celcius with a relative humidity of approximately 50%. The cold stress phenotype was represented by 10°C, the average growing temperature by 20°C and the heat stress at 30°C.

2.3.1.3 Light Conditions

Plants were grown in full light conditions of approximately 250 lumens and shaded light conditions of around 75 lumens. Day cycles of 16 hours were used, with darkness of 8 hours. Figure 2.1 shows how the variation in light conditions was achieved using shade cloth to filter out 70% of the light. Figure 2.2 shows an example of the data produced using a HOBO® (Onset Computer Corporation, MA, USA) temperature, humidity and light intensity monitor for the 10°C room at high light intensity.

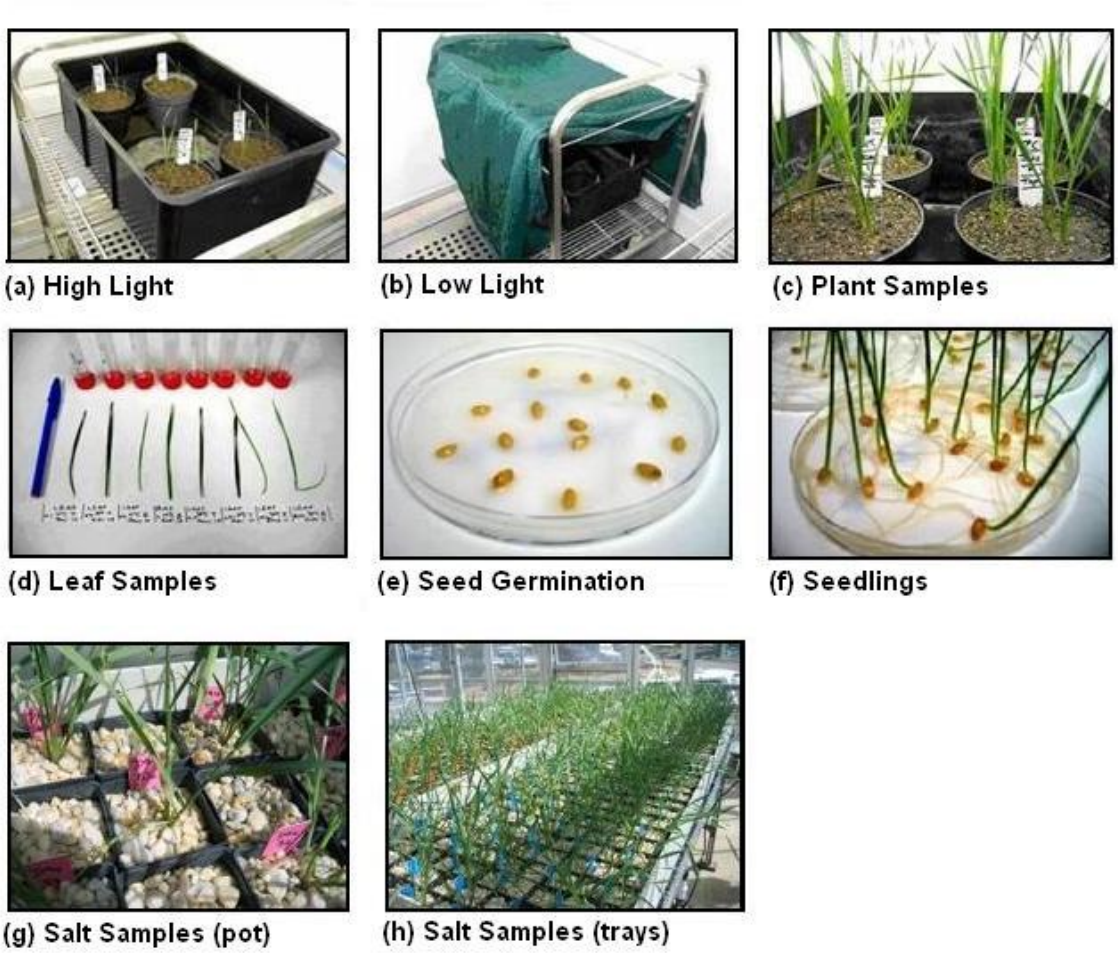


Figure 2.1: Plant growth conditions

(a) High light conditions, (b) Low light conditions showing shade cloth over plants to block out light, (c) Plant samples in pots, (d) Leaf samples after harvest, (e) Wheat seed before germination, (f) Seedlings after germination, (g) Salt samples in pots and (h) Salt samples in trays.

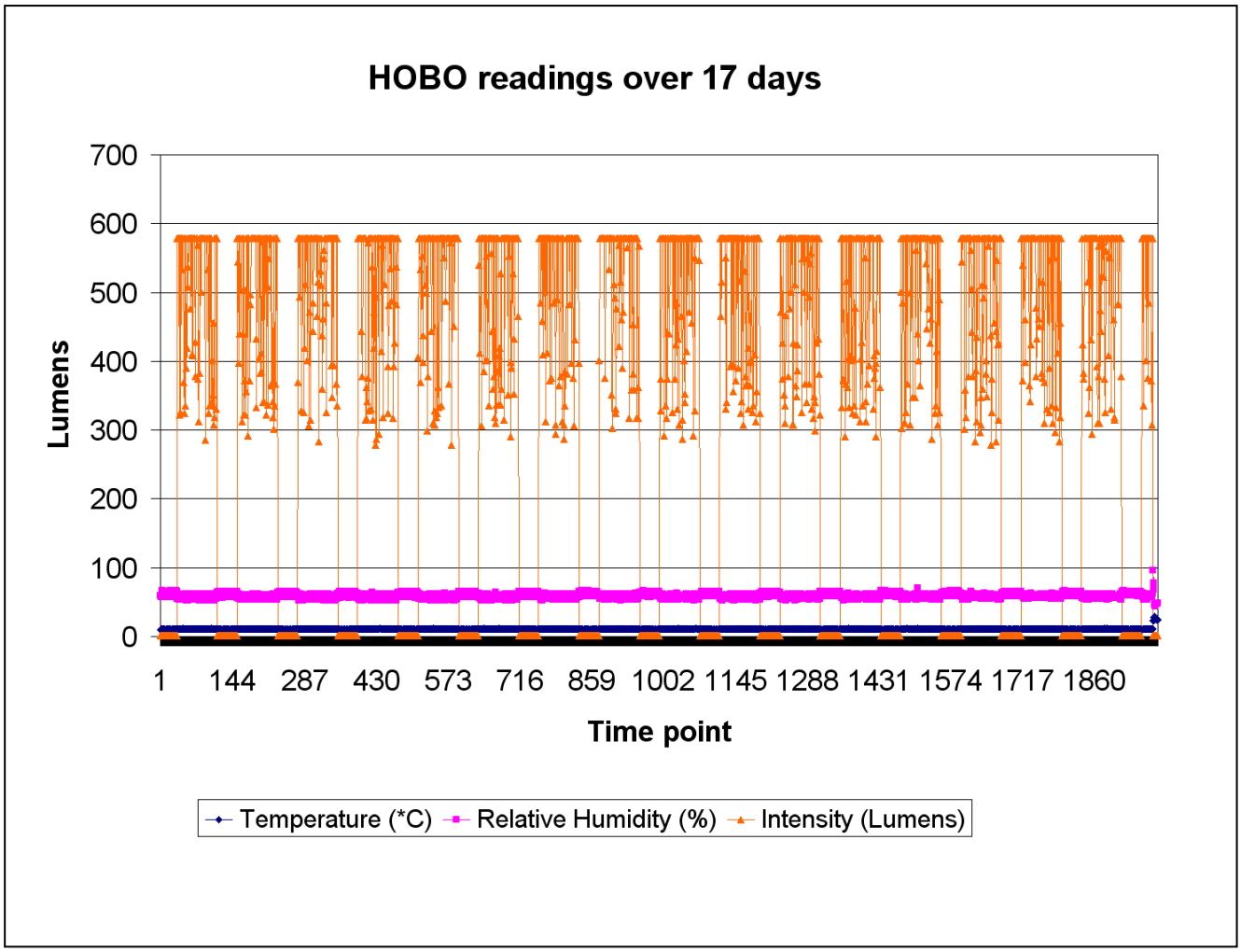


Figure 2.2: HOBO readings from the 10°C room at high light intensity

Data points collected by HOBO™ over 17 days showing an average temperature of 10 degrees, relative humidity levels of around 50-60% and an average light intensity of 510 lumens.

2.3.2 Salt Stress

Tetraploid Durum wheat plants were grown at CSIRO Plant Industries (Canberra) under controlled conditions. Plants were grown in quadruplicate under control, incrementally increased salt and all-at-once shock salt conditions (figure 2.1). Seeds were selected by weight, surface sterilised with 1% hypochlorite for 15 min and germinated in Petrie dishes for 3 days. Germinated seeds were planted into pots containing quartz gravel, one plant per alternate pot, in moulded trays containing 144 pots. Trays were sub-irrigated with either saline or non-saline nutrient solution, as described in Munns *et al*. (1995)[169]. The nutrient solution at full strength was Hoagland and Arnon solution No 2, containing 4 m*M* Ca2+ and 1 m*M* P. Seedlings were watered initially with tap water, then half strength nutrient solution was introduced 2 days after emergence (DAE) and increased to full strength at 3 DAE. Commencing at 4–10 DAE, 25 m*M* NaCl was added to the 'incremental' plants in the irrigation solution twice daily over 3 days to a final concentration of 150 m*M*. Supplemental Ca2+ was added (as $CaCl_2$) to bring the total concentration of Ca2+ to 10 m*M*, and the molar ratio of Na+:Ca2+ to 15:1. Control treatments always had 1 m*M* NaCl added to the nutrient solution. Salt 'shock' plants had 200 mM of NaCl added for several hours before samples taken. The pH was measured twice weekly and adjusted as needed to pH 6.0 with HCl. Root temperature was controlled using condensers in the solution reservoirs and monitored every 5 min using thermocouples. All experiments were conducted in a glasshouse with natural light and controlled air temperature. Daily glasshouse air temperature ranged from between 23 ◦C (day) and 18 ◦C (night).

2.3.2.1 Tetraploid Durum seedlings

Seed was provided my Rana Munns and Richard James (CSIRO) from 8 tetraploid durum wheats. Seeds were germinated and leaf DNA extracted and used for library construction to represent all 8 durum genotypes.[170] The plant descriptions are shown in table 2.1 with the addition of 3 durum cultivars, Kukri, Janz and Westonia.

Table 2.1: Tetraploid durum wheat samples.

Species	Genotype
T. turgidum L. ssp. Carthlicum	14
T. turgidum L. ssp. Durum 39	39
T. turgidum L. ssp. Durum 149	149
T. turgidum L. ssp. Polonicum	55
T. turgidum L. ssp. Turanicum	28
T. turgidum L. ssp. Turgidum	62
T. turgidum L. ssp. Durum Tamaroi	T
T. turgidum L. ssp. Durum Wollaroi	W
Additional samples	
T. turgidum L. ssp. Durum Janz	J
T. turgidum L. ssp. Durum Westonia	W
T. turgidum L. ssp. Durum Kukri	K

2.3.3 Plant growth

2.3.3.1 Seed germination

All seeds were pre-treated before germination by placing them into a container of absorbent silica at 4°C for 24 hours to break any possible dormancy that some seeds can have and to equilibrate the seeds so that they germinate simultaneously[171]. The seeds were then removed and soaked in a Petri dish containing 4-6 layers of filter paper and milli-Q water. After 4 hours, the excess water was drained and the seeds arranged evenly on the wet filter paper. Petri dishes were sealed and seeds incubated in the dark at 37° C for 24 hours. Seeds were then incubated uncovered at room temperature near a window for up to 7 days. The filter paper was constantly

kept moist, with approximately 2-3 ml's of milli-Q water added daily. Once the seeds had germinated and the leaf was approximately 5 cm in length, the seedling was removed from the Petri dish and either potted in soil or processed for DNA extraction.

2.3.3.2 Seedling plantation

Seedlings from the seed germination stage were placed in a 10 cm radius pot containing coarse quartz gravel (light and temperature experiments) and in hydroponic pots (salt experiments). The coarse gravel was wet and 2-4 seedlings were placed 1-3cm under the surface of the soil. The seedlings were positioned so that all the roots were covered and the leaf was extended upwards. Plants were kept moist and grown in temperature and light specific growth rooms and cabinets.

2.3.4 Tissue collection

Tissue from leaf was collected at two time points, during the initial growth stages as a seedling and after several weeks of growth. These time points were labeled 'seedling tissue' and 'mature tissue' respectively. Root material was collected from seedlings in the light and temperature experiments and from mature plants in the salt experiments. Leaf expansion zone tissue was taken from salt experiment plants only in 1 cm sections from the axil of the stem. Leaf, root and leaf expansion zone tissue was cut from the plant with a scalpel and immediately immersed in liquid nitrogen to prevent DNA degradation by nucleases.

2.3.4.1 Leaf samples

Leaf tissue was harvested from seedlings once the initial lead leaf had developed. All leaves were harvested or if the plant was to be potted, only the secondary leaf was taken for DNA extraction. For mature leaves, the lead leaf on the second tiller was chosen and/or any other leaf if more tissue was required. Leaf tissue was cut from the plant using a clean razor blade and immersed in liquid nitrogen for approximately 30 seconds, allowing the entire leaf to freeze. The frozen leaf was stored at -80°C until required. The leaf was ground using a pestle and mortar to a fine powder in liquid nitrogen and transferred to a 2 ml tube for DNA extraction.

2.3.4.2 Root samples

Root tissue was harvested from the Petri dish of light and temperature treated plants and from hydroponic tanks in the salt treated plants. The entire root system was cut from the plant, rinsed in water and immersed in liquid nitrogen for approximately 30 seconds, allowing the root tissue to freeze. The root tissue was ground using a mortar and pestle to a fine powder and transferred to a 2 ml tube for DNA extraction.

2.3.4.2 Leaf expansion zone samples

Leaf expansion zone tissue was harvested from plants at the same time that mature leaf tissue was taken. A 1cm sample was cut from the main stalk for DNA extraction. The tissue was immersed in liquid nitrogen for approximately 30 seconds, allowing the leaf expansion zone tissue to freeze. The tissue was ground using a mortar and pestle to a fine powder and transferred to a 2 ml tube for DNA extraction.

2.3.5 DNA preparation

Tissue samples were processed using the DNA isolation protocol developed by Doyle and Doyle (1987)[172] and modified by Jason Carling (2003)[173]. To the powered tissue samples, 1 ml of fresh buffer working solution (Appendix 1) was added that had been pre-warmed to approximately 60°C immediately after grinding. Samples were inverted 20 times and incubated at 60°C for approximately 3 hours. 1 ml of chloroform and isoamyl alcohol (24:1) was added to each tube and inverted a further 40 times. Samples were centrifuged for 30 minutes (6000 x g) and the supernatant containing the DNA transferred to a new tube. 1 ml of isopropanol was added to the supernatant and inverted 20 times then incubated at room temperature for 10 minutes before being centrifuged for 30 minutes (6000 x g). The supernatant was removed by inversion and blotted allowing the DNA pellet to be washed once with 1 ml of 70% ethanol, centrifuged (6000 x g) for 30 minutes and the ethanol discarded. The pellet was dried either at 37 degrees for several hours or at room temperature overnight. Once dried, the pellet was resuspended in 50ul of TE buffer, with 3ug of RNase if required to remove the RNA.

2.4 DArT Methodology

The standard DArT protocol was adapted as a basis for developing the complexity reduction, adapter ligation, amplification and hybridisation experiments. A primary restriction enzyme was used to digest wheat genomic DNA while a second restriction enzyme was added to further reduce the number of fragments in the genomic representation. Within this reaction, an adapter was ligated to the primary restriction enzyme cut site on the DNA molecule. The resulting digestion ligation reaction was used as a template for the amplification reaction and the DNA fragments between the adapters were amplified by PCR. The amplification mix was then purified, labeled with fluorescent dyes and hybridised to a microarray containing wheat DNA fragments. The fluorescence was captured using a fluorescent scanner and images analysed using DArTsoft version 7.4.3. The data was analysed using various criteria to generate a list of candidate polymorphisms between the samples analysed. Experiments were repeated and results compared.

2.4.1 DNA quality

DNA samples were checked for quality and quantity by running them on an agarose gel with a size/mass ladder or lambda DNA at various concentrations. The DNA band is then compared to the standard, and its concentration estimated. All DNA samples are adjusted to uniform amounts of approximately 100 ng/µl. 1µl of DNA was mixed with 4 µl of 1x loading dye (Fermentas, Canada) and run against a 1 kb DNA Ladder (Fermentas, Canada) as a reference. Samples were loaded on a 0.8% agarose gel and run in 1 x Tris-acetate-EDTA (TAE) buffer at 80 volts until the bromophenol blue migrates approximately 4 cm from the wells. The gel was stained after running with 0.5 µg/ml ethidium bromide in 1 x TAE for 20 min and photographed. DNA was then quantified using the 1 kb DNA ladder as an approximate reference. A defined high molecular weight band should be visible on the gel showing DNA of good quality. Poor quality DNA will show a smear and indicate degradation. RNA will be present as a smear at the bottom of the gel. It is not necessary to remove RNA for the DArT protocol as it has been shown in-house (unpublished data) that RNA does not interfere with the digestion and ligation or amplification steps. An example of good quality DNA with approximate quantity of 100 ng/µl is shown in Figure 2.3.

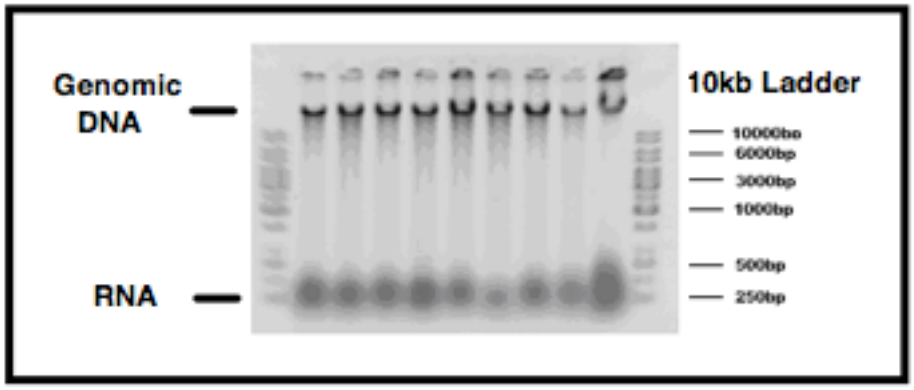


Figure 2.3: Good quality genomic DNA extracted from wheat

2.4.2 Adapter preparation

Adapters are used in the digestion and ligation step of the DArT protocol and are composed of two oligonucleotides with partially complimentary sequences. Adapters are designed to bind the cut sites of the primary restriction enzyme used. *PstI* (New England Biolabs, USA) cuts the sequence 5'-C^TGCAG-3'[174] and adapters were designed to bind to the 5'-TGCA-3' overhang created. PstI_adapter1 oligonucleotide was designed to bind to the overhang shown in figure 2.4 and the sequence is shown in table 2.2. As the genomic DNA fragment is double stranded, a second oligonucleotide, PstI_adapter2, was designed to bind complementary to PstI_adapter1. Rather than adding each oligonucleotide separately to the digestion ligation reaction, they are annealed to each other first and subsequently called PstI1+2 adapter. Adapters are adjusted to a 100 µM concentration and equal volumes mixed and incubate for 5 min at 80°C then cooled to room. The concentration is then adjusted to a 5 µM working concentration. Once the adapters have ligated to the restriction enzyme cut site, amplification of the fragment with an unknown sequence inside the adapters is possible, as primers are designed to bind to the known adapter sequence (figure 2.4).

Table 2.2: Adapter definitions.

Adapter name	Primary RE	Sequence 5' to 3'
PstI_adapter1	PstI	CAC GAT GGA TCC AGT GCA
PstI_adapter2		CTG GAT CCA TCG TGC A
C-fwd11	PfIMI	CTG AGT AGT GCC AGA ACG GTC NNG
AdaptC_rev		GAC CGT TCT GGC AC

2.4.3 Restriction Enzyme Selection

Restriction enzymes (endode-oxyribonucleases or restriction endonucleases) are a group of enzymes that catalyze the cleavage of double stranded DNA molecules at specific sites to produce discrete fragments. They are found in bacteria as a natural defense against foreign DNA and are routinely used in genetic manipulation experiments. In DArT, they are used to reduce the complexity of an organism's genome, that is, to reduce the genome into a subset of fragments. The primary restriction enzyme is used to cleave specific sites so that an adapter molecule can be ligated to the fragment, giving its ends a known sequence. Secondary restriction enzymes are used to eliminate fragments from this pool to further reduce the complexity of the genomic representation. Only fragments with an adapter ligated to each end will be amplified during PCR. The restriction enzymes used are shown in Table 2.3 and briefly described below.

The restriction enzymes were selected based on their recognition sequence as well as their sensitivity to methylation as shown in table 2.3. *PstI* (NEB, USA) is a 6 base cutter as it recognises the 6 bp DNA sequence CTGCA*G and is not sensitive to any methylation. Taq$^{\alpha}$I (NEB, USA) is a 4 base cutter that recognises the sequence T*CGA and is blocked by Dam methylation. The restriction enzyme McrBC (NEB, USA) was used in addition to PstI and Taq$^{\alpha}$I as it cleaves DNA containing 5-methylcytosine, 5-hydroxymethylcytosine or N4-methylcytosine on one or both DNA strands[175]. McrBC cleaves DNA between two (G/A)^{m}C sites separated by up to 3kb, with 55-103 bases optimal[176]. Comparisons between markers identified using PstI and Taq$^{\alpha}$I / and markers identified using PstI, Taq$^{\alpha}$I and McrBC should allow the identification of methylation polymorphisms as fragments will be destroyed that

contain a methylated cytosine thus being eliminated from the genomic representation and not hybridising to the array.

Table 2.3: Restriction Enzyme definitions.

Restriction Enzyme*	**Cut Sequence**	**Methylation status**
PstI	5′...CTGCA▼G...3′ 3′...G▲ACGTC...5′	Not sensitive
Taq$^{\alpha}$I	5′...T▼CGA...3′ 3′...AGC▲T...5′	Blocked by overlapping Dam methylation.
MseI	5′...T▼TAA...3′ 3′...AAT▲T...5′	Not sensitive
PflMI	5′...CCANNNN▼NTGG...3′ 3′...GGTN▲NNNNACC...5′	Blocked by overlapping Dcm methylation.
McrBC	…PumC (N40-3000) PumC…	Cleaves DNA containing methylcytosine (mC) on one or both strands (Pu = G or A)

*All enzymes and graphics supplied by New England Biolabs Inc (USA), www.neb.com.

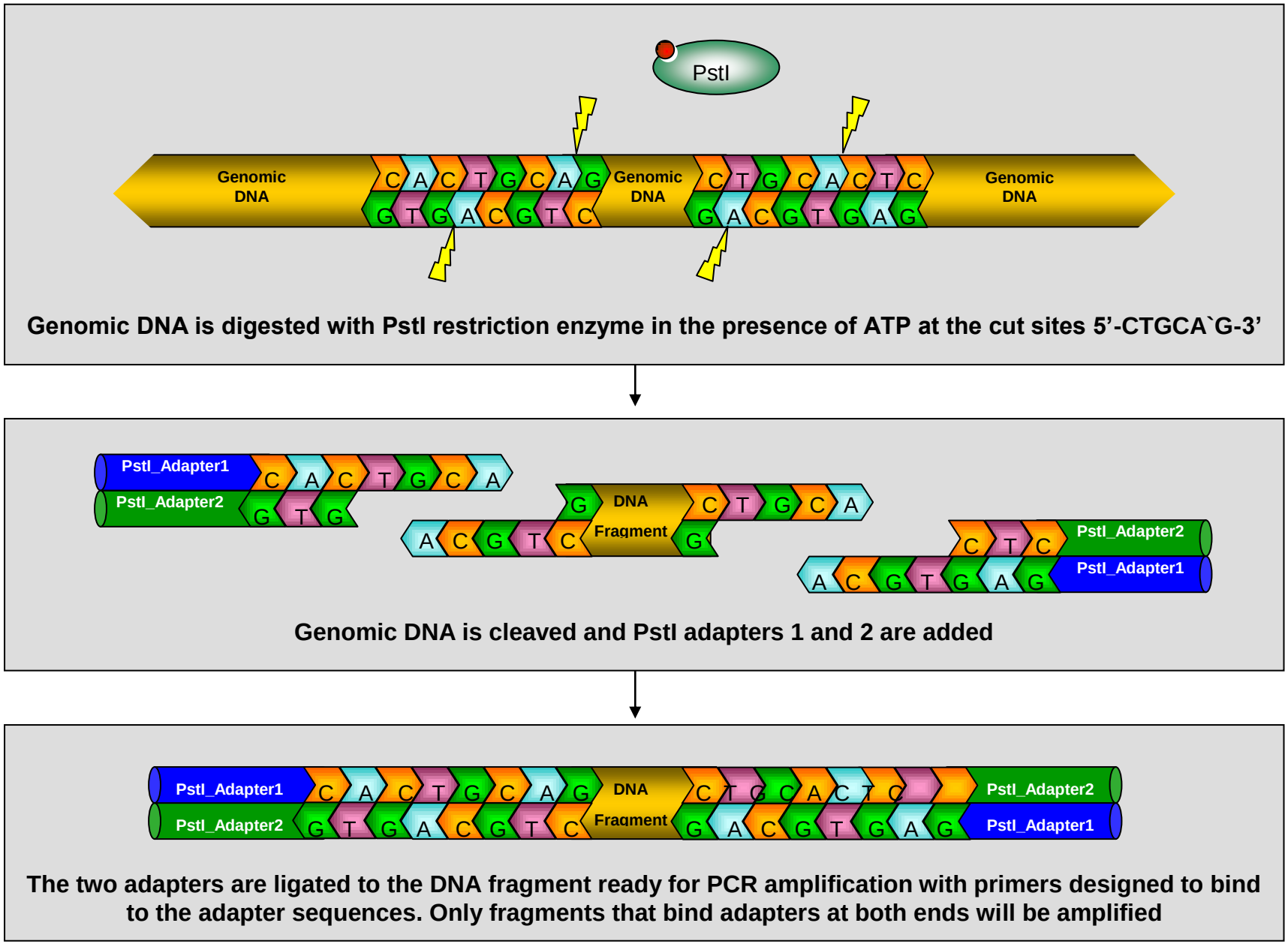


Figure 2.4: The Digestion / Ligation step in DArT procedure

Source: Diversity Arrays Technology Pty. Ltd.

2.4.4 Restriction Enzyme Digestion and Adapter Ligation

DNA is digested by a primary restriction enzyme, such as PstI, to generate a group of fragments that represent the entire genome. An adapter sequence is ligated to the cut site on these fragments (figure 2.4). A proportion of these fragments are destroyed using a secondary restriction enzyme that cuts frequently within the wheat genome, such as Taq$^{\alpha}$I. A subset of these fragments remain, those enclosed by the adapter sequences, comprising approximately 0.1% of the original genome. These fragments are referred to as the genomic representation. The PstI / Taq$^{\alpha}$I digestion ligation reaction consists of approximately 100 ng of the DNA sample, 1 µl 10X RE buffer (100 mM Tris-OAc, 500 mM KOAc, 100 mM $Mg(OAc)_2$, 50 mM DTT, pH 7.8), 0.1 µl 100 x BSA (NEB), 0.2 µl 50 mM ATP, 0.1 µl 5 µM PstI adapter, 0.1 µl PstI (20 U/µl, NEB), 0.1 µl Taq$^{\alpha}$I (20 U/µl NEB), 0.2 µl T4 DNA ligase (30 Weiss units/µl, NEB) and 7.2 µl molecular grade H_2O (Sigma-Aldrich). Samples are incubated at 37°C for 2 hours (for PstI) and at 60°C for 2 hours (Taq$^{\alpha}$I). The enzymes are heat inactivate at 80°C for 20 min and used as a template for amplification.

2.4.5 PCR Amplification primer preparation

PCR amplification primers are oligonucleotides designed to bind to the adapter sequence from the digestion and ligation reaction. The primers are used in the PCR to amplify the region between two adapters on a single DNA fragment. The PstI 1+0 primer was used for *PstI* reactions and has the sequence 5'-GATGGATCCAGTGCAG -3'. This primer binds to the adapter sequence and during PCR the DNA is denatured, the primers anneal, the polymerase adds bases and extends the fragment so that it is exponentially replicated. Primer sequences are shown in table 2.4.

2.4.6 PCR amplification of the genomic representation

PCR is used to amplification the genomic representation containing fragments bound by the adapter molecules. These fragments do not contain the restriction site for the secondary restriction enzyme and are short enough to be amplified efficiently by the RedTaq DNA polymerase (Sigma-Aldrich, Australia). The standard number of PCR cycles used is 30, unless stated otherwise. The number of PCR cycles is kept to a

minimum to reduce the bias towards fragments that are amplified more efficiently than others.

Table 2.4: PCR Primer definitions.

Primer name	Sequence 5' to 3'
M13f	GTT TTC CCA GTC ACG ACG TTG
M13r	TGA GCG GAT AAC AAT TTC ACA CAG
PstI+0	GAT GGA TCC AGT GCA G
AdaptC_PCR	GAG TAG TGC CAG AAC GGT C

The digestion and ligation reaction is used as a PCR template for amplification. 1 µl of the reaction is used with 5 µl of 10 x PCR buffer (100 mM Tris-Cl pH 8.3, 500 mM KCl, 15 mM MgCl2, 0.1 % gelatine (Sigma-Aldrich, Australia), 1 µl 10 mM dNTPs (Sigma-Aldrich, Australia), 2 µl 10 µM PstI+0 primer, 2 µl RedTaq (1 U/µl; Sigma-Aldrich, Australia) and 39 µl molecular grade H_2O (Sigma-Aldrich, Australia). The PCR amplification reaction was used in an MJ thermal cycler under the conditions shown in table 2.5. The annealing temperature of 58°C was calculated due to the number of bases in the primer oligonucleotide as well as the GC content.

5 µl of PCR product is analysed on a 1.2 % agarose TAE gel to confirm that a homogeneous smear of fragments is obtained and to visualise the size distribution. A homogenous smear indicates that the genomic distribution does not contain repetitive genomic sequences (large quantity of fragments of the same size) and/or mitochondrial or chloroplast DNA. These repetitive fragments generally show distinct bands. An example of a homogeneous smear for wheat PstI / Taq^{α}I amplified genomic representation is shown in figure 2.5.

Table 2.5: Standard PCR conditions for genomic representation amplification.

Step 1:	94°C for 1 minute
Step 2:	94°C for 20 seconds
Step 3:	58°C for 40 seconds
Step 4:	72°C for 1 minute
Step 5:	Go to step 2 (29 more times)
Step 6:	72°C for 7 minutes

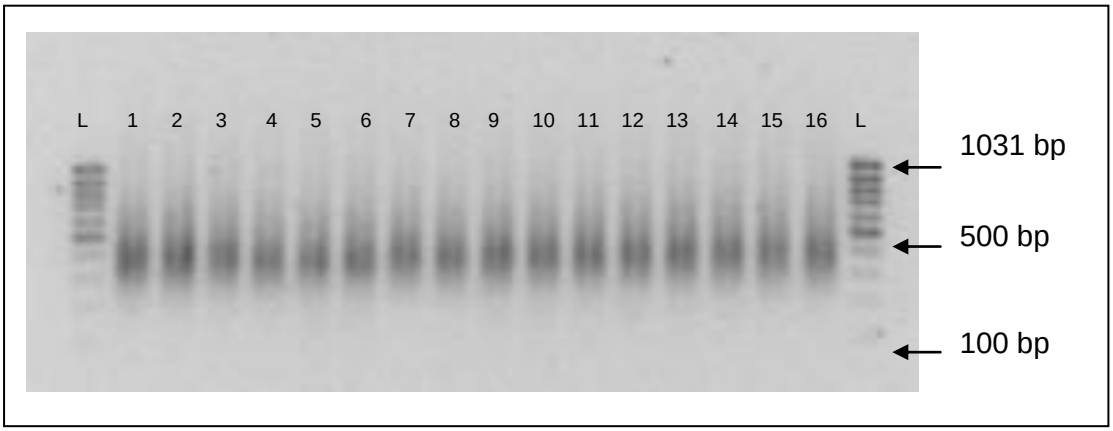


Figure 2.5: PCR amplified homogeneous smear of wheat

The 16 samples were digested with PstI and Taq$^{\alpha}$I, ligated with Pst_adapter 1+2 and amplified with PstI1+0 for 30 cycles of PCR and run on a 1.2% agarose gel showing a smear of fragments around the 500 bp range.

2.4.7 Genomic representation library creation

Once the genomic representation has been produced, a library is created where a subset of fragments will be represented. The pool of chosen genotypes representing a given species genetic diversity or parents of a cross are mixed, cloned and amplified for microarray printing.

2.4.7.1 TOPO cloning

The pooled genomic representations are cloned using the TOPO TA Cloning® system (Invitrogen, USA). The PCR fragments are cloned into the linear pCR2.1-TOPO® vector where the fragments are inserted within the LacZ gene and transformed into competent TOP10 *E. coli* cells, as shown in figure 2.6.

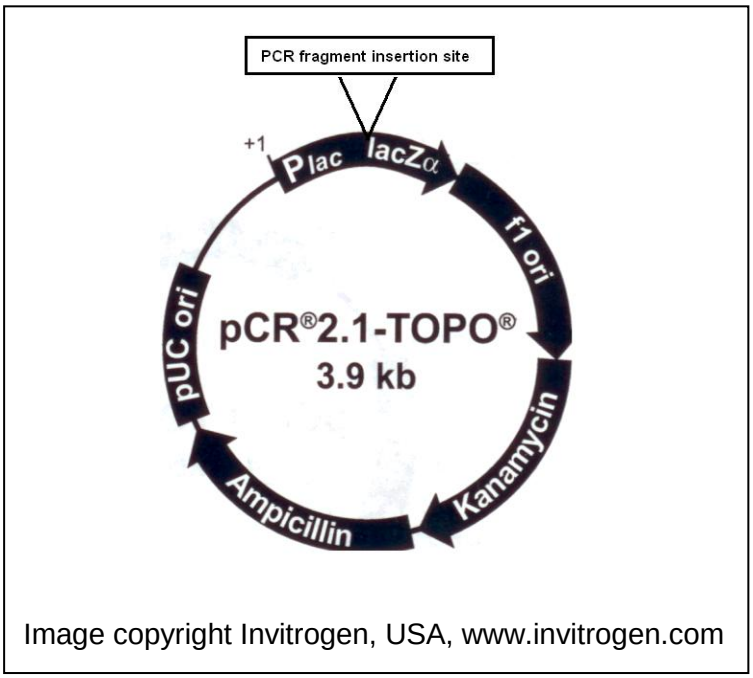


Image copyright Invitrogen, USA, www.invitrogen.com

Figure 2.6: pCR2.1-TOPO® vector showing the PCR fragment insertion site

Source: Modified from Invitrogen, USA[178].

The TOPO cloning reaction contains 4 µl of the combined genomic representations (PCR product) of each genotype, 1.0 µl of salt solution (1.2M NaCl, 0.06M $MgCl_2$) and 1.0 µl of pCR2.1-TOPO vector. The reaction is incubated at room temperature for 15 minutes followed by the addition of 2.0 µl of the TOPO ligation reaction. This is added to a vial of ONE Shot TOP10 Electrocomp E. coli cells (Invitrogen, USA) and mixed gently. The reaction is incubated on ice for 15 minutes, transferred to a cuvette and electroporated once at 1.5 kvolts. 500 µl of S.O.C. medium (Appendix A) is added on ice to the mixture and incubated with shaking at 37°C for 40 minutes. After the first round of growth, 20 µl of the bacterial suspension is plated onto LB plates with ampicillin (100 µg/ml) and X-gal (40 mg/ml) then incubated at 37°C overnight for a maximum of 16 hours.

2.4.7.2 Colony Selection

Once the genomic representation has been cloned into bacterial cells and allowed to grow, they are picked and the cloned DNA fragment amplified. Each colony is picked into 50 µL Freezing Medium (Appendix A) in 384-well format using a sterile toothpick. Blue/white selection is employed where individual white colonies are selected due to the disruption of the LacZ gene by the insertion of the DNA fragment (figure 2.6) [179]. Disruption of this gene destroys the enzymatic ability of the β-galactosidase subunit, which inhibits metabolism of X-gal. Cells that do not contain a cloned DNA fragment have a non-disrupted LacZ gene, thus producing an active subunit and metabolising X-gal to produce a blue substrate. White colonies only are transferred to each well, being careful to avoid cross contamination. This procedure is repeated to fill multiple 384-well plates. Plates are covered with a lid and the edges sealed, then incubated at 37°C for 20 to 24 hours.

2.4.8 Insert PCR amplification

Cloned genomic representation fragments are PCR amplified directly from the overnight grown bacteria plate into new plates containing 25 µl of insert amplification PCR mix (Appendix A). Plates are inoculated using sterile plastic 384-well replicators that transfer approximately 2µl of the bacterial suspension from the growth plate to the insert amplification plate. The inoculated insert amplification plates are sealed with PCR film and amplified in 384 well PCR thermal cycler (Eppendorf, USA) using conditions described in table 2.6.

2.4.9 Spotting plate preparation

Once a library has been created and the insert successfully amplified, the insert needs to be prepared so that it can be spotted onto the microarray substrate. The insert amplification plate contain the DNA fragment (insert) of interest plus the amplification reagents.

Table 2.6: Insert Amplification PCR conditions.

Step 1:	95°C for 4 minutes
Step 2:	57°C for 35 seconds
Step 3:	72°C for 1 minute
Step 4:	94°C for 35 seconds
Step 5:	52°C for 35 seconds
Step 6:	72°C for 1 minute
Step 7:	Go to step 4 (34 more times)
Step 8:	72°C for 7 minutes

The DNA fragment is purified by drying the insert amplification plate at 37°C overnight, then by washing the dried DNA fragment with 35 µl of 70% ethanol. The plate is then sealed, briefly centrifuged to collect the ethanol in the bottom of the well and incubate for 90 minutes at room temperature. Plates are centrifuged at maximum speed (3220 x g) for 40 minutes at 30°C. Immediately after centrifugation, the ethanol is removed by inverting the plate over a plastic collector and centrifuged at 200 rpm (8 x g) at 20°C for several seconds. Plates are then blotted dry and allowed to air dry at 37°C for 30 min to 1 hour. The DNA fragments are then dissolved in 20 µl of DArTSpotter2 (in-house, Appendix A), distributed to each well, then sealed with a PCR plastic seal (Qiagen, Australia), centrifuged for a few seconds and incubated at room temperature for at least 2 days. Plates should be shaken by hand and re-centrifuged several times to give sufficient time for the DNA to re-dissolve. From the insert amplification plate, 2.0µl of PCR product is analysed on a 1.2% agarose TAE gel. The plate is considered successfully amplified if more than 95% of inserts amplify with single bands, showing a single amplified DNA fragment from the genomic representation (figure 2.7). The bacterial plate is sealed and covered with a plastic lid and is stored at -80°C until required

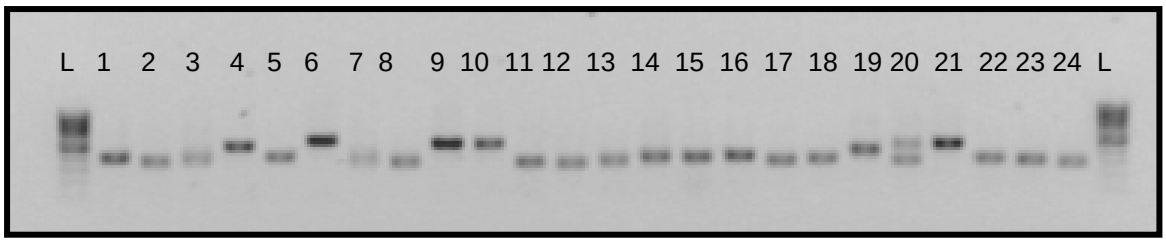


Figure 2.7: Insert amplification PCR of wheat.

Insert amplification by PCR showing single bands present in each well indicating that the cloned DNA fragment was successfully amplified from the TA cloning vector The exception is well 20, where the well was contaminated with two colonies during colony picking

2.4.10 Microarrays printing

Once the individual DNA fragments from the genomic representation have been dissolved in DartSpotter2, they are ready to be spotted onto 1 x 3 inch glass microarray slides (Erie Scientific, USA). The glass is supplied with a poly lysine surface to increase the surface adhesion of the DNA fragment being deposited onto it. The MicroGrid II microarrayer (BioRobotics, UK) was used to spot the samples using 64 tungsten split pins, collection the DNA fragments from the spotting plates, in a random distribution in duplicate over the slide. Once printed, the slides are processed so that the DNA fragments are heat fixed to the slide. The slides are immersed in almost boiling water for a few seconds, then centrifuged to dry them before being stored in a desiccator with silica gel connected to a vacuum pump for 30 minutes, followed by storage at room temperature for 24 hours minimum before use.

2.4.10.1 Rearraying of polymorphic markers

After DArTsoft polymorphism analysis has identified high quality polymorphic markers from the printed library, the markers can be condensed into a new library, thus removing non-informative, poor quality or heterogeneous markers. The new library is termed a ‘rearray library’ as clones are rearrayed from the original bacterial growth plates to the new library. The original bacterial growth plates are replicated into fresh freezing media (appendix B) and grown for 20 hours at 37°C. The original library plates are loaded into the MicroGrid microarray printer and using a single

sterile pin the bacterial cell suspension is transferred from the original well to inoculate the new well. After all wells are inoculated, the new rearray library plate is grown at 37 degrees, replicated for storage at -80°C and colony PCR amplified. The amplified plate containing the amplified polymorphic DNA fragment is then purified, re-suspended and printed in a new array.

2.4.11 Generation of Targets for Hybridisation

The term 'target' refers to the DNA fragments that will be fluorescently labeled and hybridised to the microarray. DNA quantity and quality is tested and adjusted to 100ng/ul. Targets are produced using a similar method to the library creation method with the exception that DNA fragments produced from the digestion ligation and subsequent PCR amplification are not cloned. Instead, the PCR amplification reaction is purified and labeled with fluorescent dyes and hybridised to a corresponding microarray.

2.4.11.1 Target Preparation

The DNA that is to be hybridised to the microarray is prepared using the same protocols as described in the restriction enzyme digestion, adapter ligation and PCR amplification steps within the genomic representation creation. Depending on the complexity of the genomic representation, it may be necessary to produce targets using a pool of several independent PCR reactions, rather than single PCR reactions.

2.4.11.2 Target Precipitation

After the PCR amplification of the genomic representation, targets are precipitated and washed to remove the PCR reagents that remain in the reaction. To the amplification reaction, 1 volume of Isopropanol is added and mixed several times. The reaction is incubated at room temperature for 15 min, followed by centrifugation at 4000 rpm (3220 x g) for 40 min at 30 °C. The supernatant is discarded by inversion of the tube with gentle shaking to remove as much of the supernatant as possible followed by blotting onto a lint free absorbent towel. The pellet is washed with 100 µl of 70% ethanol and incubated for 10 min at room temperature followed by further centrifugation at 4000 rpm (3220 x g) for 40 min at 30 °C. The supernatant is

discarded and the pellet is dried at 37°C for 30 min to 1 hour. The DNA pellet is then dissolved in 3.5 µl of H_2O or TE buffer (Appendix B).

2.4.12 Labeling of Targets

The DArT procedure uses 2'-deoxyuridine 5'-triphosphate (dUTP, GE health, USA) labeled targets on a microarray substrate. The dyes fluoresce under a laser where the intensity is recorded and scored for each target. A TOPO labeled control is used that hybridises to the TOPO vector sequence that is amplified as part of the cloned DNA fragment. The TOPO control is labeled in a different colour to the target and indicated the presence of DNA on the slide. Limited to the scanners laser and filter configuration, DArT uses Cy3 and/or Cy5 fluorescent dyes to label targets and Alexa488 as a TOPO control, details shown in table 2.7.

Targets are labeled by adding 3.5 µl of target (purified cloned DNA fragment in water or TE buffer) with 1.0 µl 10x NEB Buffer 2, 0.5 µl 500 µM random decamers (Fermentas, USA), 0.5 µl dNTPs (Sigma-Aldrich, Australia) and 2.0 µl MG H_2O (Sigma-Aldrich, Australia). Samples are denatured at 95°C for 3 min and put on ice after brief centrifugation to collect the liquid. 0.1 µl of 25 nmoles dUTP dye, 0.5 µl Klenow exo- (500 U/µl; New England Biolabs, USA) and 1.4 µl MG H_2O is added to the labeling mix then incubated in the dark for 3 hours at 37˚C.

2.4.13 Microarray hybridisation

DArThybridiser has been developed and prepared in house to aid in the hybridisation of DNA targets to DNA microarray slides. It is preheated to 65°C before use to reduce its viscosity. DArThybridiser consists of Alexa488 labeled pCR2.1 vector polylinker (approx. 80 ng per slide), ExpressHyb (Clontech, USA), 10mg/ml herring sperm DNA (Promega, USA) and 2 mM EDTA pH 8.0.

Table 2.7: Fluorescent dyes used in DArT.

Dye	Excitation wavelength	Emission wavelength	Laser
Cy3 dUTP *	550nm	570nm	Green 532nm
Cy5 dUTP *	649nm	670nm	Red 633nm
Alexa-488 ^	488nm	519nm	Blue 488nm

* Dye supplied by GE health, USA.

^ Dye supplied by Invitrogen, USA.

Microarray slides are hybridised in DArT hybridisation chambers that hold 8 slides per chamber (in-house design). The labeling reaction (target) is removed from incubation at 37°C and briefly centrifuged. 60 µl of DArThybridiser is transfer into each target sample and mixed by pipette. If two targets are to be used, both targets are mixed into 50 µl of DArThybridiser. Samples are denatured at 95ºC for 3 minutes and held at 55ºC until loaded onto the array. Individual samples are mixed and 60 µl of each sample is deposited onto the slide over the array and a cover slip (60 x 24 mm) added. The chambers are sealed and incubated in a water bath at 65ºC over night (12-16 hours) at 65ºC.

2.4.14 Microarray slide washing

After incubation, the slides are washed to remove the hybridisation buffer. DArT uses 4 wash solutions comprising of 1 x SSC plus 0.1 % SDS (wash 1), 1 x SSC (wash 2), 0.2 x SSC (wash 3) and 0.02 x SSC (wash 4). 200 µl of 0.5 M Dithiothreitol (DTT) (Sigma-Aldrich, USA) is added to each 1 litre volume of wash solution before use and mixed well. Hybridisation chambers are removed from the 65°C water bath and slides are removed individually, their cover slip removed and placed directly in wash 1. Once all slides have been removed, slides are agitated in wash 1 for 1 minute then incubated for 4 minutes. Slides are agitated in wash 2 for 1 minute followed by incubation for 4 minutes and transferred to wash 3 for 2 minutes including 1 minute of agitation. Slides are agitated in wash 4 for 30 seconds then centrifuged immediately at 500 x g for 7 min at 30°C to dry them. Slides are dried further in a light-protected

desiccator with silica gel connected to a vacuum pump for 30 min. Slides are then scanned with a laser at the corresponding wavelength for each of the dyes.

2.4.14 Image Acquisition

Microarray slides are scanned using a Tecan LS300 (Grödig, Austria) laser fluorescent microarray scanner equipped with 488 nm, 532 nm and 633 nm filter sets. Slides are placed in slide adapters and scanned individually for each wavelength. The fluorescent intensity data (signal) from the slides is stored in a TIFF image file for each wavelength scanned. The auto-gain function is enabled to allow for slight differences in fluorescent intensities for each slide. Images are then imported into DArTdb, an in-house database and analysed using DArTsoft. An example of a slide scanned in this way is shown in figure 2.8, where the Cy3 and Cy5 targets as well as the TOPO reference scans can be seen.

2.4.15 DArTdb Image Extraction

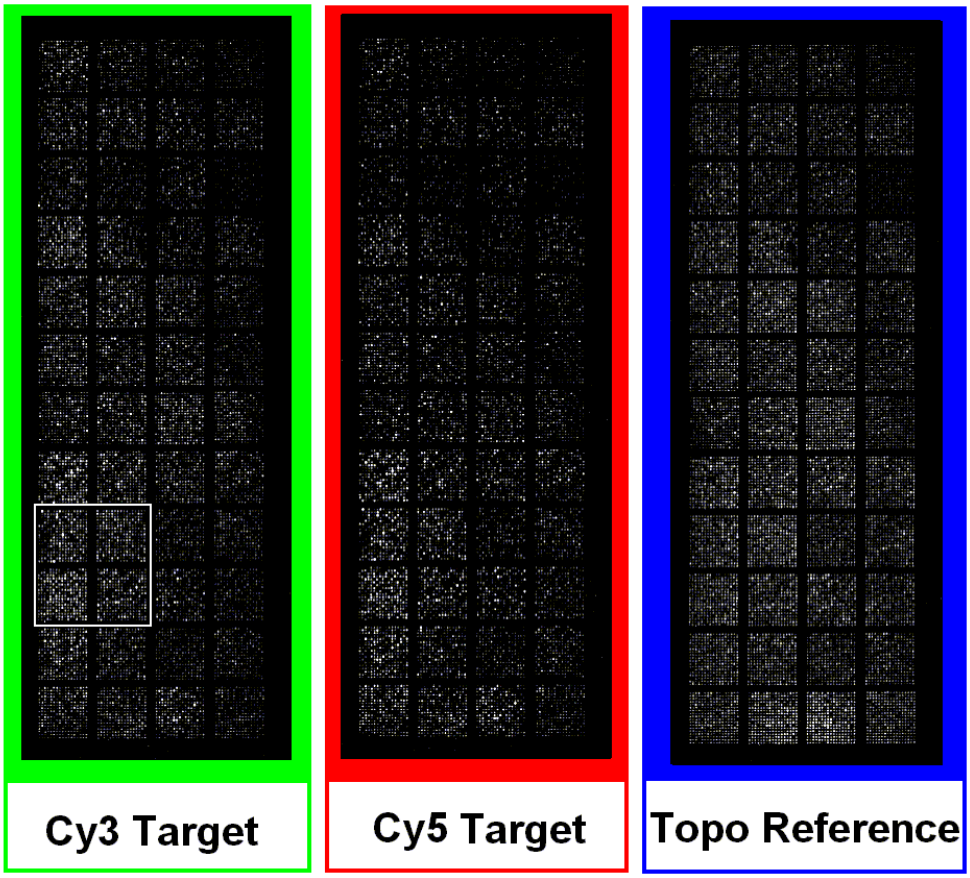
DArTdb is an in-house laboratory information management system (LIMS) that stores experimental information, array designs, protocols, DNA plate information, TIFF files and extracted image data. It also functions to extract fluorescent intensity data from scanned microarray TIFF files (figure 2.8b). DArTdb automates spot recognition (figure 2.8c) and image data extraction using a print map file generated from the microarrayer to find the approximate location and design of the microarray. The number of rows, columns and blocks is read and a grid is placed over the TIFF image. The software within DArTdb then makes slight adjustments for image rotation or pin variation and re-aligns the grid. Intensity values are extracted from each pixel within the spot as well as intensity values from pixels around the spot representing the local background value (figure 2.8d). The intensity values are analysed for RatioPix, RatioMed, RatioAvg and RatioCov then compared for homogeneity (Qratios) and signal-to-noise. Slides, blocks and/or spots are then flagged for rejection based on this information before further polymorphism analysis.

The RatioPix value is the pixel based ratio that is calculated from the logarithm of the ratio of the target against the reference at the pixel level. RatioPix is calculated by taking the log of the intensity of the target minus the background target divided by the intensity of the reference minus background of the reference.

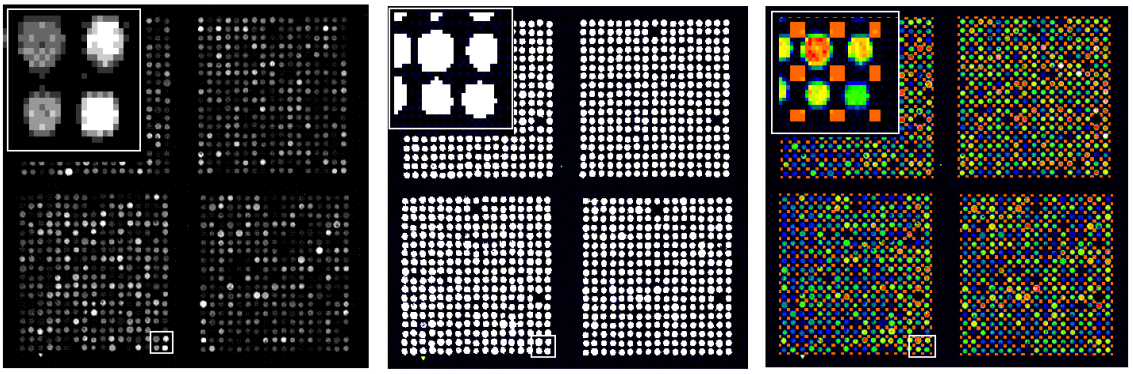
RatioMed is the logarithm of the ratio of the target against the reference using the median intensities of the spot. RatioMed is calculated from the log of the median intensity of the target minus the background of the target divided by the median intensity of the reference minus the background of the reference.

RatioAvg is the same as the RatioMed, except average intensities are used in place of median intensities. RatioAvg is calculated from the log of the average intensity of the target minus the background of the target divided by the average intensity of the reference minus the background of the reference.

RatioCov values are the logarithm of the ratio of target against reference using a covariance of pixels measurement. RatioCov is calculated using the log of $(yy - xx + \sqrt{(yy - xx)^2 + 4 * xy^2} / 2 * xy$, where yy is the covariance of target and target, xx is the covariance of reference and reference and xy is the covariance of the reference and target.



(a) Cy3 and Cy5 labeled target images with Topo labeled reference image.



(b) TIFF image **(c)** Spot recognition **(d)** Background calculation

Figure 2.8: Cy3, Cy5 and Topo reference channel target TIFF images (a), raw TIFF images (b), DArTsoft spot recognition grid overlay (c) and background calulation parameters (d).

The four characterizations of spots above, RatioPix, RatioMed, RatioAvg and RatioCov, are expected to be very similar (within 1%) when a spot is of good quality. A bigger difference between these ratios translates into a difficulty in measuring, possibly due to weak signal, fluorescent dust or scratch etc. The Qratio measures the homogeneity of the ratios giving a measurement of the quality of the ratios measured. The higher the Qratios value, the more reliable the ratios are that were extracted. It is calculated by taking the mean of the four ratios divided by the mean ratios plus the variance of the four ratios.

The signal-to-noise ratio is calculated as a further assessment for the quality of spots, measuring target signal strength relative to the background noise. Signal-to-noise values are calculated by taking the median value of signal on the reference channel divided by the median value of signal on the reference channel plus the background value of the spot in the reference channel. The reference channel is the wavelength (channel) scanned for the TOPO-labeled target that was hybridized to the array.

Background values are taken as high background signal can interfere with spot recognition, causing the addition or elimination of pixels into the spot. Local background values are calculated from all channels (images from all wavelengths) from pixels around each spot. The DevBack value defines the standard deviation of pixel intensities for each channel with the CFBack giving the coefficient of variation of background value.

Other factors are taken into account when the spot is being analysed for quality, including the expected size of the spot compared to actual size (in pixels), circularity of the spot and pixel saturation, that is, those pixels above the maximum count of 65,535.

2.4.16 DArTsoft polymorphism analysis

After data has been extracted from TIFF images, data for each spot is compared over spot replicates within the array, in multiple identical arrays and over replicate targets. Polymorphism analysis is performed through DArTsoft, an in-house microarray analysis software package. The clustering decision is used to identify markers that group together into distinct polymorphic groups. The clustering algorithm uses a bimodal score of 1 to 0 to distribute every point in three different groups. Points are assigned to a group of unclassified points if their maximum membership value is lower than the clustering decision membership threshold (default value is 0.9) or to any of the two other groups representing the crisp members of each cluster. DArTsoft then compares the cardinal of each group: if the number of unclassified points is lower than a certain percentage (default is 50%) of any other group and if the cardinal of each classified group is greater than the number of replicated slides then the final decision is positive (1), otherwise it is negative (0).

The Q value measures the fraction of the total variation across all individuals due to bimodality and is performed in one dimension. The Q value is used to rank markers based on their quality, with markers usually limited to scores greater than 75. As replicated individuals are supposed to give identical results, replicated points are expected to fall into the same cluster. After binarisation every single point is scored and DArTsoft controls the reproducibility of the experiment. The call rate value is an expression of reliability of the final scores, representing the number of scored slides against the maximum number of potential scores, with a score of 100 showing that all replicated were scored identically. Targets that are not replicated will have a reproducibility of 100. Polymorphism information content (PIC) value is used to measure the 'informativeness' of a genetic marker for linkage studies. It was originally defined by Botstein *et al* (1980)[180] with DArTsoft using a simplified version described by Anderson *et al* (1993)[181]. A marker with a PIC of 0.5 shows 50% of samples scored positive (1) that is clustered in one group and 50% scored absent (0) and clustered in the other group. Further DArTsoft and DArTDb definitions are shown in figure 2.9.

RatioPix

Pixel based ratio of the logarithm of the ratio of target against reference calculated at a pixel level.

$RatioPix = median\ of\ \alpha\%\ most\ homegenous\ Rp$

where:

$$Rp = \log\left(\frac{intensity_{target} - background_{target}}{intensity_{reference} - background_{reference}}\right)$$

Rp is calculated for each of the pixels of the spot.

α is the fraction of pixels retained

RatioMed

Median Intensity based ratio of the logarithm of the ratio of the target against the reference using the median intensities of the spot

$$RatioMed = \log\left(\frac{median\ intensity_{target} - background_{target}}{median\ intensity_{reference} - background_{reference}}\right)$$

Clustering

Uses a value of 0 or 1 to distribute every point into three different groups:

$decision = 1\ if\ card(ug) < \alpha * card(g_0) \wedge card(ug) < \alpha * card(g_1) \wedge card(ug) \geq n$

$decision = 0\ otherwise$

where ug is the group of unclassified points,

g_0 is the group of crisp elements belonging to the lower cluster,

g_1 is the group of crisp elements belonging to the upper cluster,

α is the clustering decision membership threshold,

card is the number of elements of a group

$\wedge$ is the symbol representing a logical "and"

PIC

Polymorphism information content value used to measure the 'informativeness' of a genetic marker for linkage studies

$PIC = 1 - \sum_{i=1 \Rightarrow n} P_i^2$

where P_i is the population frequency of the i-th allele

n the total number of allelic states

Dicsordance

The discordance value is a complementary value of the reproducibility, expressing the overall variation of scores within the replicates.

$Discordance = \sum_{1 \Rightarrow sx} [1 = \sum_{i=1 \Rightarrow c} (n_i / n)^2] / sx$

where c is the maximum number of clusters

n_i is the number of scores of the i-th cluster,

n is the total number of scored targets (excluding the "X" scores)

sx is the total number of samples

RatioAvg

Average intensity based ratio of the logarithm of the ratio of target against reference using the average intensities of the spot

$$RatioAvg = \log\left(\frac{average\ intensity_{target} - background_{target}}{average\ intensity_{reference} - background_{reference}}\right)$$

RatioCov

Covariance of pixels based on the ratio of the logarithm of the ratio of target against reference using a covariance of pixels measurement

$RatioCov = \log(yy - xx + \sqrt{(yy - xx)^2 + 4 * xy^2} / 2 * xy)$

where:

yy = covariance(target,target)

xx = covariance(reference,reference)

xy = covariance(reference,target)

Q

Q value measures the fraction of the total variation across all individuals due to bimodality, performed on one dimension

$Q = max(i * (\bar{x} - \overline{xlo_i})^2 + (N - i) * (\bar{x} - \overline{xhi_i})^2) / N * var(x) * 100\ ,\ 0 \leq i < N$

where x is the average value of the N points,

xlo_i is the average value of the i elements of lowest value,

xhi_i is the average value of the N-i points of highest value

var(x) represents the variance of the N points.

Call rate

An expression of reliability of the final scores, representing the number of scored slides against the maximum number of potential scores

$$Callrate = 100 * (\sum_{i \Rightarrow N} score_i / N\ ,\ with\ \begin{cases} score_i = 0\ if\ s_i = X \\ score_i = 1\ otherwise \end{cases}$$

where N is the total number of points (including the added average points)

s_i is the score of the i^{th} individual for a determine clone

P

A measurement of variation across individual performed on one or more dimensions

$P = (n1 * (g - g1)^2 + n2 * (g - g2)^2 / \sum_{i=1 \Rightarrow n1+n2} (x_i - g)^2) * 100$

where n1 and n2 being the number of points respectively belonging to the crisp cluster1 and cluster2 (decided by the highest membership value),

g is the centre of gravity of the n1+n2 d-dimensional x points,

g' and g2 are the centres of gravity respectively of the n1 and n2 d-dimensional x points.

Reproducibility

As replicated individuals are supposed to give identical results, replicated points are expected to fall into the same cluster. After binarisation every single point is scored and Dartsoft controls the reproducibility of the experiment

$$Repro = \sum_{i=0 \Rightarrow Nrep} dif_i / Nrep\ ,\ with\ \begin{cases} dif_i = 0\ if\ s_1 = s_2 = \ldots = s_n \\ dif_i = 1\ if\ at\ least\ one\ of\ the\ n\ scores\ is\ different \end{cases}$$

where Nrep is the number of replicates (different individuals)

n the number of replicated slides for each replicate

s_j the score of the j^{th} slide of a replicate.

Figure 2.9: DArTdb and DArTsoft analysis criteria definitions

2.5 Bulk Segregant Analysis (BSA)

2.5.1 Sample selection

A breeding experiment was performed between Wollaroi and *ssp. carthlicum* cultivars with DNA extracted from parents and 94 progeny grown in an increased salt environment. Plants were grown under normal conditions, with the increase in salt delivered via irrigation. Salt concentrations were increased to 200mM for 2 days, before leaf samples were harvested for DNA extraction. Leaf DNA samples for the BSA experiment were arranged into two phenotypic groups based on average SPAD readings. SPAD readings measure the total leaf chlorophyll content, an indication of leaf photosynthesis activity and plant growth. Plants with high chlorophyll contend should correlate with plants that are tolerant to growth in an elevated salt environment. Plants that have lower chlorophyll contend and SPAD reading should have a lower tolerance or even intolerance to a salt enriched environment.

2.6 RIL Linkage map creation

2.5.1 Linkage group assignment

Individual maps were constructed for all recombinant inbred line (RIL) populations using Easy Map, a program developed in-house for high-throughput mapping of double haploid (DH) and RIL populations. Easy Map automates the distribution of markers into linkage groups, the ordering of markers within linkage groups (based on the RECORD algorithm), the detection of potential genotyping errors, the re-optimisation of marker orders after replacing potential errors with unknown genotype calls and the estimation of map distances. Linkage groups were then assigned to chromosome / linkage groups based the existing chromosome assignments of markers printed on the array. At the current level of marker coverage, most chromosomes are represented by more than a single linkage group. The order and orientation of linkage groups within chromosomes is unknown. In addition, there were a number of loci that were excluded because they were not sufficiently linked to any other linkage group. The ordering and orienting of linkage groups was correlated with an in-house created linkage map where the marker order and chromosome positions are known.

2.5.1 Map image generation

Linkage group maps were generated using data output from Easy Map imported into Map Manager QTX 0.30 (www.mapmanager.org)[182] and then graphically represented using Map Chart 2.2 (www.biometris.nl).[183] Map Manager QTX was used to sort Easy Map data into linkage groups using self RI linkage evaluation, a linkage criterion of P=0.0001 and the map function Kosambi. Linkage group data was exported as linkage distance in centimorgans (cM) and imported into Map Chart, where images were produced.

Chapter 3

Developmental Diversity

“Anyone who has never made a mistake has never tried anything new”

Albert Einstein

3.0 Developmental Diversity

3.1 Introduction

Developmental diversity is the study of the principles and processes that underlie growth, development and evolution between groups of biological organisms. In the context used in this chapter, developing tissue in wheat plants, that is the leaf, root and leaf expansion zone (growing region or inter collating meristem) tissue was analysed using DArT across various cultivars. Although cells within an organism contain the same genome, availability of DNA for transcription and translation differs, including differences in histone folding, protein binding and methylation patterns[184]. Methylation differences can effect gene expression within in the cell or tissue type and can block restriction enzymes from cutting DNA. DArT will be used to analyse differences in tissue types across wheat cultivars using restriction enzymes to digest genomic DNA, creating a genomic representation. The genomic representation is then fluorescently labeled and hybridised to a microarray and the fluorescence of each DNA fragment or feature, measured.

3.2 Janz and Kukri tissue analysis

3.2.1 Aims

An initial experiment was designed to evaluate the effectiveness of using DArT to detect tissue specific polymorphisms between wheat cultivars. Developmental diversity was explored using tissues from various cultivars including DNA from Janz and Kukri, two common Australian cultivated hexaploid bread wheats. The aim of these experiments is to use DArT to detect and evaluate polymorphisms between leaf and root tissue from the 2 wheat samples, and also 30 progeny samples from a cross between them. Polymorphisms between tissue types will be analysed and data presented, with cultivar specific polymorphisms explored in chapter 5.

3.2.2 Specific Methods

Janz and Kukri DNA samples, provided by the South Australian Research and Development Institute (SARDI), from the 2 parents and 30 progeny was extracted

from leaf and corresponding root tissue. Libraries were generated in-house for leaf and root separately using PstI and Taq$^{\alpha}$I restriction enzymes from 29 wheat cultivars including a cross between Halberd and Cranbrook hexaploid wheat cultivars. Libraries were printed so that duplicate copies of each of the 4,608 clones were randomly arranged on the slide. Targets were generated and hybridised in duplicate, so that one target was labeled with Cy3 dye and the duplicate labeled with Cy5 dye, then hybridised to a single slide. Both replicates were generated from a separate digestion ligation reaction and separate PCR amplification reactions. The entire experiment was then duplicated, producing data for each clone from the 2 spots per slide, over 4 slides for each of the 32 DNA samples.

3.2.3 Results

3.2.3.1 Leaf and root targets on a leaf library

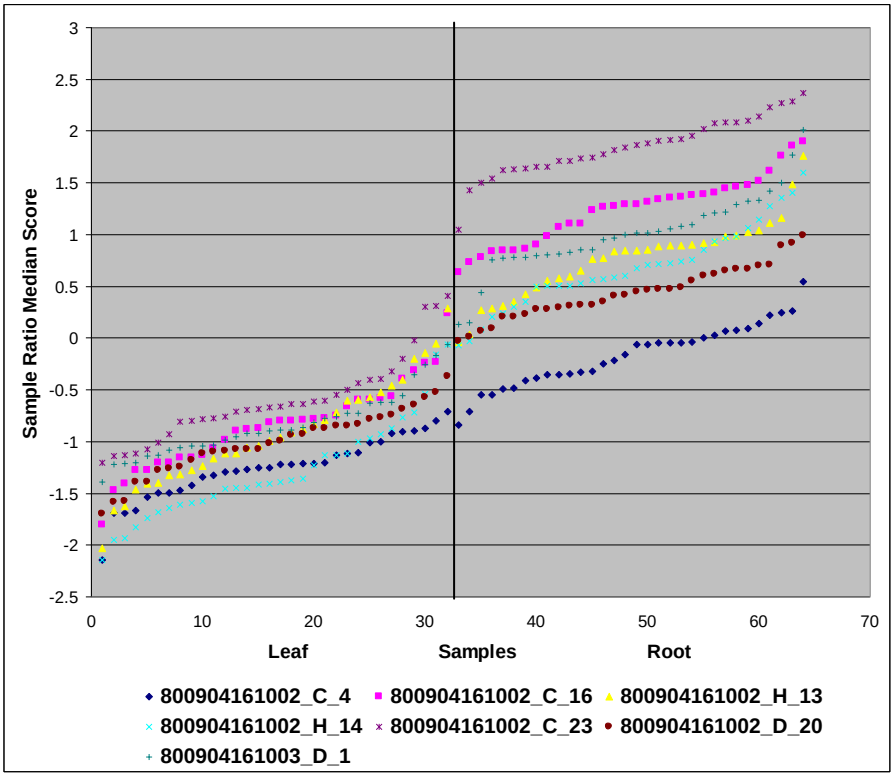
After sample hybridisation, targets were scanned for each dye and the fluorescent signal intensity extracted from each pixel on the image by DArTdb. DArTsoft polymorphism analysis was then performed to identify polymorphic markers between leaf and root tissue samples for each parental cultivar and for the 30 progeny. Data was consolidated from two independent experiments using score merger, an in-house Perl script that compares information generated from both experiments. It was found that from the 4,608 markers on the array, data passed quality control parameters for 3,452 markers (74.9%) in the initial experiment and 4,575 markers (99.2%) in the replicated experiment, with 3,436 markers (74.6%) passed in both experiments. Failure to pass the quality control parameters in image extraction function by DArTdb can be due to a weak signal for that spot, failure of that spot to print, morphological issues within the spot, dust or debris on the slide interfering with spot intensity, etc.

The data generated was combined in score merger and a list of 4,591 markers scored from both experiments. Of these, 851 markers were scored with a marker consistency of 75% or greater over the 32 leaf and 32 root samples for both experiments. The 1,156 markers scored in only one experiment were also included. The 851 markers with a high consensus score between experiments and the 1,156, markers identified in only one experiment were limited to a Q value (quality of bimodal variation) of above 75. The number of markers scoring a Q value greater

than 75 in at least one of the experiments totaled 142, 137 markers scored in both experiments and only 5 markers that scored in only 1 experiment. Of these 142 markers, 68 markers were scored with a reproducibility of 1 scoring inconsistency (98.43) or less in the initial experiment and 29 markers in the duplicated experiment, with 85 markers and 12 markers identified with 1 scoring inconsistency or less in one or both experiments respectively. Of the 85 high quality markers identified, 7 markers were scored polymorphic between tissue samples, with all 7 markers scored absent in leaf samples and present in root samples for both Janz and Kukri parental samples.

Further analysis of the 30 progeny samples shows that 25 samples (83.3%) were scored polymorphic in the same way as parental samples, with the fragment present in root and absent in leaf tissues. Of the remaining 5 samples, markers were scored with various inconsistencies, that is, scored with an ‘X’ rather than the bimodal present ‘1’ or absent ‘0’ score. An ‘X’ represents a score where the raw data did not fall into a distinct bimodal (present or absent) clusters or was scored in one cluster at one data point and in the other for a replicate data point. Of these 5 samples, 2 samples were scored in the same fashion as the parental samples for the 7 markers, but with several ‘X’ scores, 1 sample was scored with a majority of ‘X’ scores and 1 marker was scored with all ‘X’ scores. The remaining sample was scored in reverse, with 6 out of 7 markers scored present and 1 marker scored ‘X’ in leaf samples and the reverse in root samples, with 1 marker scored absent and 6 markers scored ‘X’. The 7 markers are shown in figure 3.1 for all 32 leaf and 32 root samples, with sample ratio median data graphed against root and leaf samples. Marker 800904161002_C_16 is shown in addition, where the bimodal distribution of leaf and root markers can be seen for the 32 leaf and 32 root samples.

(a)



(b)

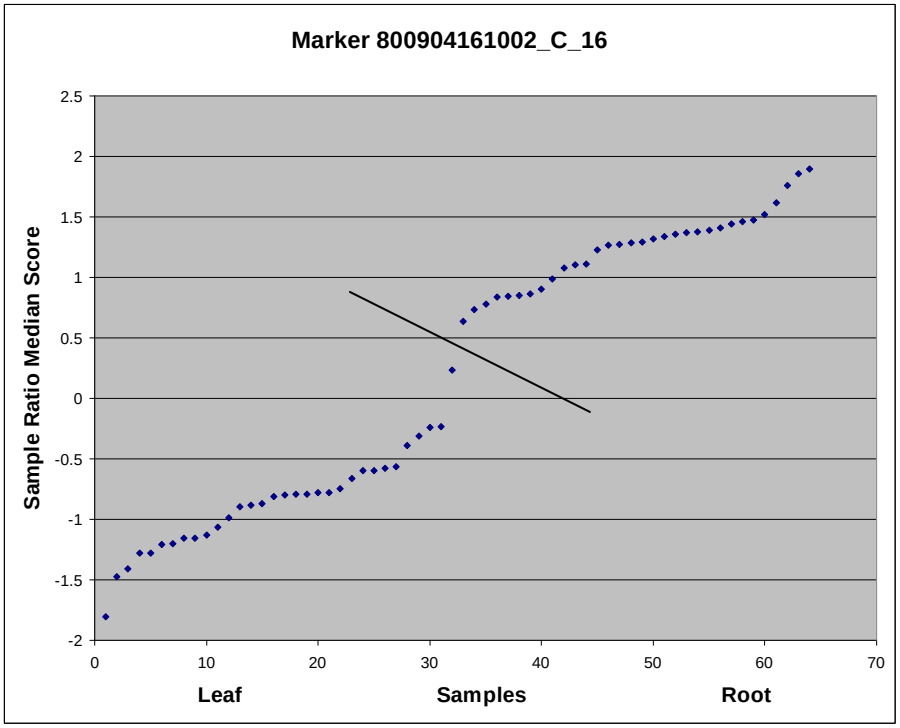


Figure 3.1: A clear distinction between leaf and root samples can be seen in both examples. **(a) Leaf and root tissue specific markers found on leaf array showing 7 sample ratio median scores against tissue type, (b) Marker 800904161002_C_16 showing the bimodal distribution of sample ratio median scores into 2 distinct clusters based on leaf and root tissue samples.**

Of the remaining 78 high quality markers identified from both experiments, a further 6 markers were identified that were polymorphic between leaf and root tissue in Kukri only, with 1 marker scored present in leaf tissue and 5 markers scored present in root tissue. An additional 2 markers were scored present in Janz leaf tissue but absent in Janz root tissue. Analysis identified 35 markers that scored polymorphic between Janz and Kukri cultivars (described in Chapter 5: Genetic Diversity Analysis), 17 markers that are suspected to be cultivar polymorphisms, containing one scoring inconsistency (an 'X') in analysis, 3 markers that were non-polymorphic between parental samples and 14 markers that contained scoring inconsistencies. These non-polymorphic parental markers scored polymorphic across the 30 progeny samples.

3.2.3.2 Leaf and root targets on a root library

After the successful identification of tissue specific markers using the leaf array, the experiment was repeated using the same targets and hybridised to an array containing DNA fragments generated from root tissue. The experiment was replicated and quality control analysis identified 4,472 (97.04%) markers in the initial experiment and 4,608 (100%) markers in the replicated experiment. Of these, 4,472 markers were identified in both experiments with an additional 136 markers passing quality control parameters in the replicated experiment only. Data was extracted and DArTsoft polymorphism analysis used to score markers from the 32 leaf and 32 root samples within the Janz and Kukri cross. Marker scores for each experiment were compared using score merger and limited to a marker consensus of 75 or greater. Analysis identified 1,421 markers, with 89 markers scored with a Q value of 75 or greater, all of which were identified in both experiments. Of the 89 high quality markers, 47 markers scored a call rate of 80 or greater (additional quality check) in one or both experiments. Analysis of these 47 markers in the parental samples identified 5 markers that were scored polymorphic between leaf and root samples in Kukri, 1 marker scored polymorphic between leaf and root samples in Janz and no markers that were polymorphic between tissue samples in both cultivars. A further 7 markers were scored polymorphic between Janz and Kukri cultivar samples, either scoring present or absent for the marker in one but not both cultivar samples. The majority of remaining markers scored at least one 'X', that is, not in a delimitative present '1' or absent '0' cluster, or scored opposite amongst replicates for one or both of the leaf and root sample pairs.

3.2.4 Conclusions

From the 4 experiments performed using Janz and Kukri DNA samples and the 30 progeny, it can be seen that it is possible to differentiate between tissue samples using the DArT method. Analyses identified markers that were only present in leaf or root tissue from and between Janz and Kukri cultivars. Analysis shows that the leaf array was more successful in identifying polymorphic markets compared to the root library. This experiment further shows that although the genomic DNA is essentially the same in both tissue types, genomic analysis using DArT does reveal candidate polymorphic DNA markers that can be used for identifying tissue types from and between Janz and Kukri wheat cultivars.

3.3 Light and temperature stress in bread wheat

3.3.1 Aims

Once it was confirmed that DArT was able to detect tissue specific polymorphisms in wheat cultivars, experiments were designed to target developmental tissue diversity by subjecting plants to a range of environmental stimuli. Janz and Kukri plants were grown at 2 light and 3 temperature conditions and DNA extracted from leaf and root tissue for analysis. The aim is to focus on developmental diversity, that is, to detect polymorphisms identified between leaf and root tissue and between seedling and mature samples. Chapter 4 will focus on molecular physiology looking at the different temperature and light condition polymorphisms and Chapter 5 will analyse cultivar polymorphisms.

3.3.2 Specific Methods

Janz and Kukri plants were grown at 10°C, 20°C and 30°C in growth cabinets at the Research School of Biological Science (RSBS) at the Australian National University (ANU) and at the Centre for the Application of Molecular Biology to International Agriculture (CAMBIA) in Canberra. Within each growth cabinet, plants were grown in full light conditions of approximately 250 lumens and shaded light conditions of around 75 lumens, with a day cycle of 16 hours. Plants were allowed to grow for 60 days then allowed to flower at 20°C. Seeds were collected and then germinated on filter paper with a selection sacrificed for leaf and root tissue samples. These

samples were termed ‘seedling’ DNA. The remaining seedlings were planted and grown at the same environmental condition as the previous generation for 30-40 days before leaf tissue was harvested. These samples were termed ‘mature’ leaf DNA. Samples were DNA extracted. labeled, hybridised, scanned and analysed using DArTsoft. Data was restricted to a Q value of 75 or greater and a call rate of 80 or greater and markers within this range are termed ‘high quality’ markers.

3.3.3 Results

3.3.3.1 Tissue specific polymorphism at 10°C

Leaf and root samples were analysed from Janz and Kukri samples grown at 10°C. Analysis identified 258 high quality polymorphic markers, all of which were cultivar specific, varying between Janz and Kukri but scored the same for all tissue types. As plants were growing in sub-optimal conditions, leaf quality was poor and viability low.

3.3.3.2 Tissue specific polymorphism at 20°C

Leaf and root tissue samples were analysed from Janz and Kukri cultivars grown at 20°C. Analysis identified 986 high quality polymorphic markers, the majority of which were Janz and Kukri cultivar specific. Of these, 39 markers were identified as tissue specific in Janz. Comparison of the 4 leaf and 3 root samples identified 19 markers scored present in leaf and 20 markers in root. Further analysis of marker 801504280005_E_15 (wPt-5967) shows differentiation between sample ratio median scores for leaf and root tissue in Janz samples, as shown in figure 3.2. The divergence between leaf (scored present) and root (scored absent) samples can be clearly seen for this marker. Analysis of Kukri samples identified 22 tissue specific markers, 8 markers in leaf and 14 markers in root.

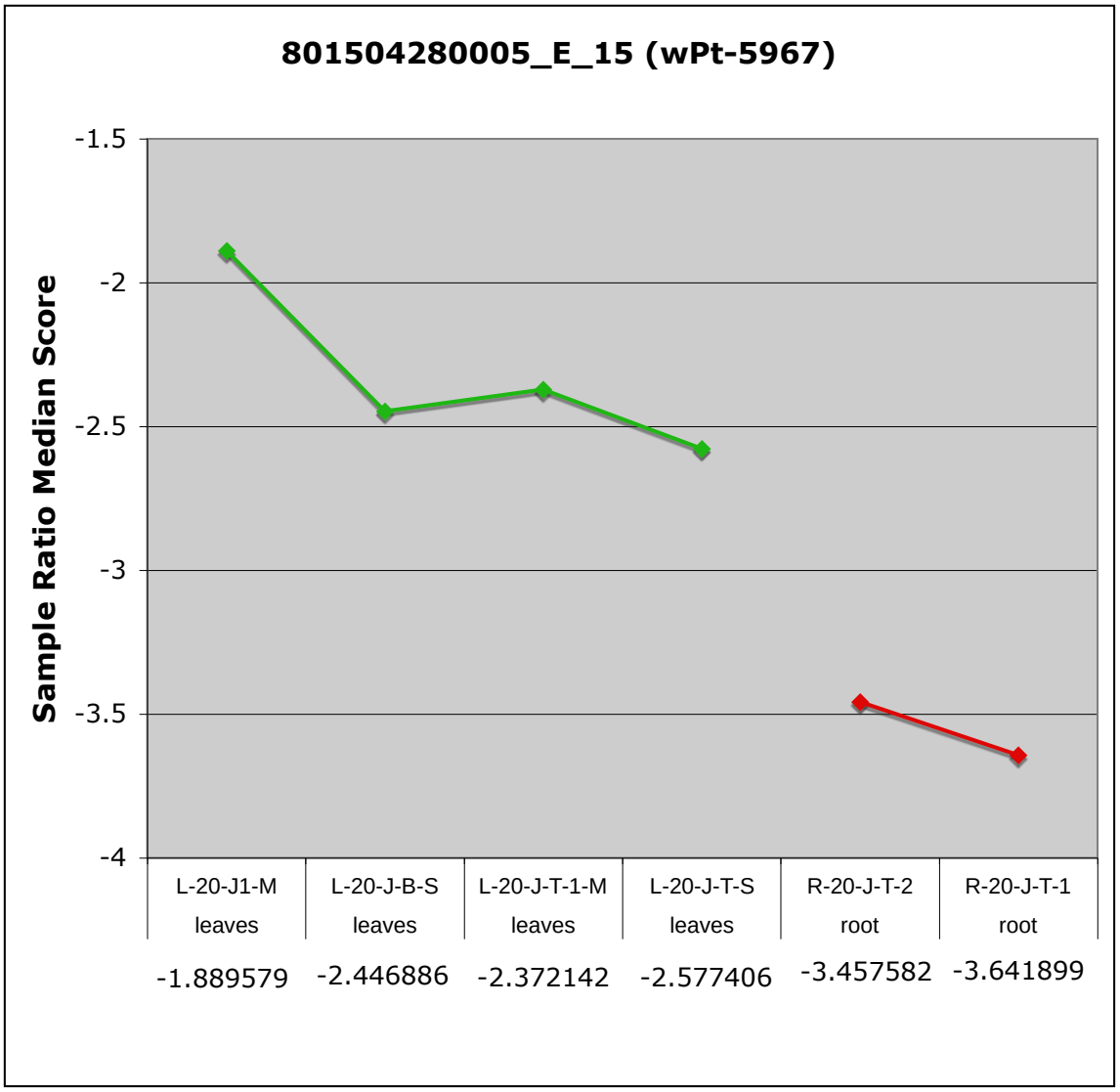


Figure 3.2: Tissue specific Janz marker identified from samples grown at 20°C. Marker 801504280005_E_15 can be seen to be polymorphic when sample ratio median scores are compared across tissue samples.

3.3.3.3 Tissue specific polymorphism at 30°C

Leaf and root samples were analysed from Janz and Kukri samples grown at 30°C. Analysis identified 1,120 high quality polymorphic markers, the majority of which were cultivar specific. Of these, 97 markers were identified as tissue specific in Janz. Comparison of the 5 leaf and 1 root Janz sample identified 5 markers that scored present in leaf and 92 markers in root. The higher number of root markers is due to the smaller sample size in root tissue analysis. Analysis of Kukri samples identified 102 tissue specific markers, 37 markers in leaf and 56 markers in root.

3.3.3.3 Comparison of tissue specific polymorphism at 20°C and 30°C

Comparison of the 39 Janz markers identified at 20°C and the 97 markers identified at 30°C found 0 markers common to both analyses. Comparisons of the 22 Kukri markers identified at 20°C and the 102 markers identified at 30°C also identified no overlapping markers. This indicates that these markers are potentially only polymorphic at certain temperatures and are not polymorphic at all temperatures. An example is marker 801504280002_E_2 that at 30°C was scored polymorphic in Kukri but not Janz, however at 20°C was scored polymorphic in Janz but not Kukri between leaf and root tissue samples.

3.3.3.2 Comparison of seedling and mature leaf samples

DArTsoft analysis of seedling and mature leaf samples from Janz and Kukri identified 1,286 high quality polymorphic markers. Of these markers, 974 were found to be polymorphic in the majority of samples between seedling and mature leaf tissue in Kukri. Markers were limited to those that scored the same bimodal score for all 5 seedling and 3 mature leaf samples, identifying 197 markers. Of these, 90 markers scored present in Kukri seedling leaf samples and 107 markers in Kukri mature leaf samples. As will be further discussed in Chapter 4: Molecular Physiology, the seedling samples were grown at different temperatures. All mature leaf samples were extracted from plants grown at 10°C while seedling samples were extracted from plants grown at 10°C, 20°C and 30°C. A further 255 markers can be identified between seedling and mature leaf samples in Kukri grown at 10°C. Of these, 155 markers were scored present in 10°C seedlings and 84 markers present in 10°C mature leaf. In addition, 4 markers were scored absent for both seedling and mature leaf samples grown at 10°C but present in seedlings grown at 20°C and 30°C and 12 markers scored present in seedling and mature samples grown at 10°C but absent in seedlings grown at 20°C and 30°C. Comparisons in Janz identified very few polymorphisms between seedling and mature samples, with the majority of the variation between samples grown at differing temperatures, as discussed in Chapter 4: Molecular Physiology.

3.3.4 Conclusions

From experimentation, it can be seen that temperature and light conditions do effect tissue development as demonstrated through DArT analysis of genomic DNA. Polymorphisms were detected across replicates for wheat from at 20°C and 30°C and between seedling and mature tissue types. The main limitation to these experiments was sample size, being to small to allow proper marker testing, hence why in-depth results were not presented here. However, as a proof of concept experiment, this leads to further investigations. Tissue polymorphisms in these samples exist due to differences in cellular conditions within each tissue, with DNA availability at certain loci differing depending on the regions of the genome that are actively being transcribed. These differences, such as DNA methylation, protein binding, histone complexes etc. can restrict/alter the action of restriction enzymes for example, during genomic representation generation.

3.4 Salt stress in durum wheat

3.4.1 Aims

Two cultivars of tetraploid durum wheat, *T. turgidum L. ssp. polonicum* and *ssp. durum* were grown in differing salt concentrations. Leaf and leaf expansion zone tissue was sampled with the aim to detect polymorphisms between tissue types and between salt treatments. Leaf expansion zone was used in replacement or in addition to root tissue as potential contamination with soils when using mature (not seedling) can occur.

3.4.2 Specific Methods

Ssp. *polonicum* and ssp. *durum* tetraploid wheat was grown in a glass house at the Crown Scientific and Industrial Research Organisation (CSIRO), Black Mountain, Canberra. Plants were grown in quadruplicate under control, incrementally increased salt and all-at-once shock salt conditions. The control conditions contained 1mM NaCl with the final concentration of the increasing and shock phenotypes at 200mM. Sample were hybridised in duplicate per slide, with the addition of McrBC methylation sensitive restriction enzyme to Cy5-labelled samples. The DArT standard PstI and Taq$^{\alpha}$I wheat 8 plate (V2.2) array was used to analyse polonicum and durum samples.

In addition, two arrays were made using PflMI and MseI followed by PstI and MseI restriction enzymes. These two arrays were used in addition to the standard PstI and Taq$^{\alpha}$I arrays to further diversity studies and to evaluate DAM methylation which is blocked by Taq$^{\alpha}$I and PflMI restriction enzymes.

3.4.3 Results

3.4.3.1 Analysis of seedling samples: Cy3 and Cy5 targets

DArTsoft polymorphism analysis of seedling DNA from leaf and root tissue taken from durum and polonicum cultivars identified 450 high quality polymorphic markers. Cy3 and Cy5 labeled target replicates were analysed together to identify 36 tissue specific polymorphisms, 6 markers from polonicum with 2 markers scored present in leaf samples and 4 markers in root samples. Durum analysis identified 30 tissue specific markers, with 13 scored present in leaf tissue and 17 present in root tissue. Of the 36 tissue specific markers identified, 4 markers were common to both cultivars for a specific tissue type.

3.4.3.2 Analysis of seedling samples: Cy3 targets

DArTsoft polymorphism analysis of the Cy3 targets only, that is, targets produced using PstI and Taq$^{\alpha}$I restriction enzymes, identified 1036 high quality markers. Of these, 365 were identified as polymorphic between tissue samples, 290 markers in polonicum with 136 markers present in leaf samples and 154 markers present in root samples. Analysis of durum samples identified 173 markers, 55 markers present in leaf samples and 117 markers present in root samples. Between cultivars, 98 markers were identified as being polymorphic in tissue samples for both polonicum and durum, with 29 markers present in leaf tissue, 37 markers for root tissue, 30 markers present in polonicum leaf and durum root and 2 markers present in polonicum root and durum leaf.

3.4.3.3 Analysis of seedling samples: Cy5 targets

DArTsoft polymorphism analysis of the Cy5 labeled targets only, that is, targets produced using PstI, Taq$^{\alpha}$I and McrBC restriction enzymes, identified 1125 high quality markers. Of these, 404 markers were identified as polymorphic between

tissue samples, 334 markers in polonicum with 218 markers present in leaf samples and 116 markers present in root samples. Analysis in durum samples identified 141 markers, 45 markers present in leaf samples and 96 markers present in root samples. Between cultivars, 71 markers were identified as being polymorphic for both polonicum and durum, with 21 markers present in leaf tissue, 32 markers for root tissue, 16 markers present in polonicum leaf and durum root and 2 markers present in polonicum root and durum leaf.

3.4.3.4 Comparison of markers identified in Cy3 against Cy5 targets

Analysis of Cy3 targets identified 365 tissue specific markers while analysis of Cy5 targets identified 404 markers. The difference in target preparation is the addition of McrBC restriction enzyme to the digestion of Cy5 targets during the DArT protocol. Theoretically McrBC should destroy any methylated fragments within the genomic representation as McrBC cleaves DNA containing methylcytosine (5-methylcytosine, 5-hydroxymethylcytosine or N4-methylcytosine) on one or both strands[185]. McrBC will not act upon unmethylated DNA, only cutting sites that consist of two half-sites of the form (G/A)mC[186]. Comparisons between the two sets of markers identified 103 markers common to both Cy3 and Cy5 analyses, with 261 markers found only in the Cy3 analysis and 301 markers only in the Cy5 analysis. The 261 markers identified in Cy3 analysis alone represent markers that may be methylated, as they were not present in Cy5 analysis where they would have been destroyed by the *McrBC* enzyme and removed from amplification. The 301 markers identified in Cy5 analysis may have been generated by a change in the genomic representation caused by a change in amplification frequencies of fragments after a subset were removed by the McrBC restriction enzyme.

3.4.3.5 Comparison of markers identified in Cy3 against Cy5 analysis and Cy3 and Cy5 analysis

Analysis of the 103 markers identified in the Cy3 and the Cy5 analysis individually that were common to both sets of analyses (section 3.4.3.4) compared to the analysis of 36 markers identified from Cy3 and Cy5 together (section 3.4.3.1) identified 19 markers from both analyses. There were 13 markers identified in the Cy3 and Cy5 analysis and 84 markers found in the individual Cy3 and Cy5 analysis. The 13 markers from both analyses are shown in Table 3.1, with Q values for the

Cy3 analysis, Cy5 analysis and the Cy3 and Cy5 analysis. From these 13 markers, all are shown to discriminate between leaf and root tissue samples in polonicum, with 6 markers present in leaf tissue and 7 markers present in root tissue. In durum samples, 3 of the 13 markers are polymorphic between tissue samples, with 1 marker present in leaf and 2 markers present in root samples. Specifically, marker 800904300004_K_19 is shown to discriminate between leaf and root tissue, being scored present in both durum and polonium samples. Conversely, marker 801504280001_F_13 is shown to be scored present in root tissue from durum and polonium samples. These markers can potentially be used to genetically determine what tissue the DNA was extracted from. Limitations are that the marker has only been tested between durum and polonicum samples thus the presence or absence score may not hold for other cultivars.

3.4.3.6 Comparison of markers identified in Cy3 but not Cy5 or Cy3 and Cy5 analyses

Markers scored that were identified in Cy3 (PstI and Taq$^{\alpha}$I) but not Cy5 (PstI, Taq$^{\alpha}$I and McrBC) or Cy3 and Cy5 analysis were further analysed. Theoretically, markers that were detected in Cy3 analysis should also be detected in Cy5 analysis with the exception of markers that were destroyed by the secondary McrBC restriction enzyme digestion. Markers that were 'destroyed' refers to markers that contained adapter sequences at each end of the DNA fragment that were to be amplified by PCR, but were cut at an internal McrBC methylated restriction enzyme site and thus removed from amplification. Results presented in section 3.4.3.1: Cy3 analysis identified 365 tissue specific markers, whereas Cy5 analysis identified 404 markers, with 103 markers common to both groups. The 103 markers identified in both analyses contain markers that were not destroyed my McrBC digestion, the methylation sensitive restriction enzyme, and thus theoretically must not be methylated. This leaves 562 markers, 261 markers identified in Cy3 but not Cy5 and 301 markers in Cy5 but not Cy3.

Table 3.1: Durum and Polonicum tissue specific candidate polymorphisms identified from Cy3, Cy5 and Cy3 and Cy 5 analysis

Marker Name	Q Cy3	Q Cy5	Q Cy3 & Cy5	Durum Leaf	Durum Root	Polonicum Leaf	Polonicum Root
800904090002_K_21	79.60523	78.55907	81.40054	0	1	1	0
800904161004_L_21	80.58101	81.96921	78.86291	0	0	0	1
800904300001_H_8	81.99237	82.30264	75.92136	0	0	0	1
800904300001_L_12	81.07729	85.01852	78.12206	0	0	0	1
800904300003_K_13	76.81053	82.67812	75.05973	0	0	0	1
800904300004_J_19	75.68087	81.5649	78.2348	1	1	1	0
800904300004_K_19	75.0373	75.29667	75.98404	1	0	1	0
800904300005_G_14	80.75331	81.48739	79.7954	1	1	1	0
800904300006_B_17	86.78667	88.45581	80.4019	1	1	1	0
800904300006_B_2	77.51624	81.11934	79.44077	1	1	1	0
800904300006_K_22	82.93696	75.64129	75.75141	0	0	0	1
801504280001_E_5	77.94738	84.90601	80.4389	1	1	0	1
801504280001_F_13	78.85131	79.44883	78.42138	0	1	0	1

To check the quality of these markers, the 261 markers identified as tissues specific polymorphism in Cy3 were compared to the same 261 markers scored for Cy5. It was found that 190 markers scored a Q value of below 75 and 52 markers had a reproducibility of below 1 scoring discrepancy, thus being removed from the Cy5 high analysis due to quality control. The remaining 19 high quality markers were scored non-polymorphic in Cy5 in one or both cultivars between leaf and root tissue samples.

Of the 301 markers identified in Cy5 but not Cy3, the same comparison was made. The 301 Cy5 markers were compared to the corresponding Cy3 marker scores where 234 markers were rejected based on low Q scores and a further 48 markers rejected on reproducibility. The remaining 19 high quality markers were scored non-polymorphic in Cy3 in one or both cultivars between leaf and root tissue samples.

The 301 markers identified in Cy5 but not Cy3 analyses were markers found from differing experimental conditions, such as differences in labeling efficiency or differences in the genomic representation (PCR amplification efficiencies). The 261 markers identified in Cy3 but not Cy5 analysis are, theoretically, markers that were amplified from the Cy3 representation but destroyed in the Cy5 representation by McrBC digestion. The fragments destroyed, that is, the DNA cleaved by McrBC between the two adapter sequences so they were unable to be amplified by PCR, must then contain a methylation site on the fragment/marker. However, as shown in the Cy5 analysis, there are addition markers that are identified between Cy3 and Cy5 analysis, so it is likely that the majority of markers identified in Cy3 and not Cy5 are due to differing experimental conditions and not due to the presence of a methylation site within the marker. By completing the reverse experiment, the level of error can be seen between samples that would normally have been attributed to the McrBC enzyme.

3.4.3.7 Analysis of Seedling and Control leaf samples: Cy3 and Cy5

Seedling samples grown from Petri dishes in the laboratory were compared to control plants grown in the CSIRO glass house. Both samples should be scored by DArTsoft polymorphism analysis in the same fashion as they both represent the untreated population of samples. Analysis of seedling samples from leaf tissue was compared to control leaf tissue samples. Durum and polonicum samples were generated in quadruplicate and labeled in pairs with Cy3 and Cy5 dyes. Analysis identified 455 high quality markers, 30 of which were polymorphic between tissue samples within a cultivar. Of these, 2 markers were identified in durum scoring present for the control samples and not the seedling samples. In polonicum, 6 markers were identified, 3 markers scored present in each of the seedling and control samples.

3.4.3.8 Analysis of Seedling and Control leaf samples: Cy3

Analysis of seedling and control leaf samples using only the Cy3 images, that is, targets generated using only PstI and Taq$^{\alpha}$I restriction enzymes, identified 823 high quality markers with a call rate of 75 or greater. The increase in the number of markers identified is due to the reduced sample size. Of these 823 markers, 133 markers were polymorphic between seedlings and control samples, with 63 markers identified in durum, 70 markers in polonicum and 14 polymorphic in both cultivars. Of the 133 polymorphic markers identified, 21 markers were scored present for durum seedlings, 42 markers in durum control samples, 38 markers in polonicum seedlings and 32 markers in polonicum control samples.

3.4.3.9 Analysis of Seedling and Control leaf samples: Cy5

Analysis of seedling and control leaf samples using only the Cy5 images, that is, targets generated using only PstI, Taq$^{\alpha}$I and McrBC restriction enzymes, identified 918 high quality markers with a call rate of 75 or greater. The increase in the number of markers identified is due also to the reduced sample size. Of these 918 markers, 151 markers were polymorphic between seedlings and control samples, with 53 markers identified in durum, 98 markers in polonicum and 14 polymorphic in both cultivars. Of the 151 polymorphic markers identified, 17 markers were scored present for durum seedlings, 36 markers in durum control samples, 48 markers in polonicum seedlings and 49 markers in polonicum control samples.

3.4.3.10 Analysis of Seedling and salt treated samples

DArTsoft polymorphism analysis of seedling, control and incremental salt samples from leaf and root tissue from polonicum and durum on a wheat 8 plate array revealed 419 high quality markers, that is, with a Q value of 75 or greater and a reproducibility maximum of one scoring discrepancy. Of these markers, clustering scores (1's or 0's) were highly polymorphic between cultivars but almost all were homogenous between tissue samples, with the exception of 5 markers. Of these 5 markers, 3 markers were discriminatory between polonicum with 1 marker scored present in leaf tissue samples and 2 markers present root tissue. Analysis of durum samples identified 2 markers, both scored present in root samples. These markers are shown in table 3.2.

Table 3.2: Durum and Polonicum tissue specific candidate polymorphisms

Marker Name	Q	Reproducibility	Durum Leaf	Durum Root	Polonicum Leaf	Polonicum Root
800904090002_C_18	84.276619	100	0	0	1	0
800904300002_N_22	79.189568	100	0	1	0	0
801504280001_E_5	77.524239	100	1	1	0	1
800904300006_J_20	78.824608	91.666664	0	1	0	0
801504280004_L_10	77.961136	91.666664	1	1	0	1

Of these 5 markers, marker 801504280001_E_5 was identified in the analysis of Cy3, Cy5 and Cy3 + Cy5 in section 3.4.3.1. Further analysis of low quality markers, that is, markers below a Q value of 75 and a reproducibility value of 90 identified 712 additional tissue specific polymorphic markers. Of these, 547 markers differentiated between polonicum leaf and root samples and 254 markers between durum leaf and root samples, with 89 markers polymorphic in both cultivars. As these markers have such low Q and reproducibility values, their values cannot be fully trusted.

3.4.4 Conclusions

From the results it can be seen that DArT can discriminate between tissue types in durum and polonicum wheat samples. There are numerous markers presented that can reproducibly determine genetically whether a sample was extracted from leaf or root tissue or whether it was extracted from seedling or mature control samples. These markers have only been tested and analysed using these two cultivars with limited reproducibility, so further testing is required before they can be accurately used as a molecular marker for tissue selection in wheat.

Apart from scoring errors and technical errors, non-heterogenic leaf and root tissue samples can have an impact in creating lower threshold values thus excluding the majority of markers from analysis. Polymorphisms that are detected between tissue types may not always be present in all cells within the leaf or the root material that is taken for sampling.

3.5 Cultivated durum wheat diversity

3.5.1 Aims

To further the study of developmental diversity, the variety of germplasm was increased to include four libraries containing various wheat collections from wild and cultivated durum materials. Each library was hybridised with replicate targets from leaf and root DNA extracted from 8 cultivated durum wheats (table 2.1). Analysis of the 8 durum samples was performed with the aim to identify markers selected for consistent polymorphic scoring between leaf and root samples.

3.5.2 Specific Methods

Libraries were created in-house from a diverse selection of durum wheats. Material was supplied in-house and from Ali Mehrabi, University of Tehran, Iran. All libraries were created using PstI and Taq$^{\alpha}$I restriction enzymes and printed in-house. Library 1 (7,296 clones) was created from a selection of 86 durum cultivars using leaf tissue and wheat species containing the AB+AG genomes. Library 2 (6,912 clones) contained wheat AB+AG cultivars, cultivars with ABD genomes, synthetic lines (AUS17020, 17023, EGA Hume, Meering SP-3, Minto SP-1), Ryson and Westwood Rye varieties, 9 Kofta lines and 10 triticale cultivars. Library 3 (6,912 clones) was printed from wheat ABD genome cultivars, D genome cultivars, A genome cultivars, synthetic lines and Aegilops biuncialis samples. Library 4 (6,528 clones) was made using wheat lines containing AB+AB genomes, ABD genomes, Trisomic lines, Cranbrook and Halberg cultivars and over 20 wheat cultivars using leaf including Janz and Kukri.

Targets were made from leaf and root tissue from 8 cultivated samples extracted from seedlings grown in the laboratory under the same environmental conditions. DNA was extracted and a genomic representation produced using PstI and Taq$^{\alpha}$I restriction enzymes. Targets for leaf (duplicated) and root were hybridised to 24 arrays from each of the 4 libraries, hybridised and scanned.

3.5.3 Results

DArTsoft analysis was performed using duplicated leaf samples and root samples from 8 cultivated durum wheat samples. Targets were prepared and hybridised to the 4 durum discovery arrays. Markers were selected that scored polymorphic between leaf and root tissue without exception for each cultivar. Markers were also selected that scored polymorphic between the *ssp. durum* samples (Wolllaroi, Tamaroi, durum line 139 and durum line 149) and markers that scored polymorphic for all 8 samples. Table 3.3 shows the number of tissue specific markers identified in each of these groups across the 4 durum libraries. Markers designated leaf specific refer to markers (DNA fragments) that were scored present (1) in leaf samples but absent in root samples (0). The biomodal 1 and 0 scoring system is used by DArTsoft to score each marker based on the hybridisation signal after normalisation. Similarly, markers designated root specific were scored present in root but absent in leaf tissue samples. Further analysis uses the sample ratio median scores to look at the seperation between marker scores for each marker, as this is more accurate than the bimodal DArTsoft assigned values.

Table 3.3: Tissue specific markers identified from analysis of leaf and root tissue in 8 cultivated durum wheats

	Library 1		Library 2		Library 3		Library 4		Total		
Cultivar Tissue	**Leaf**	**Root**	**Leaf**	**Root**	**Leaf**	**Root**	**Leaf**	**Root**	**Leaf**	**Root**	**Both**
T. turgidum L. ssp. polonicum	44	29	43	135	33	106	39	65	**159**	**335**	**494**
T. turgidum L. ssp. turanicum	14	15	28	87	28	94	42	35	**112**	**231**	**343**
T. turgidum L. ssp. turgidum	42	19	35	122	33	43	43	56	**153**	**240**	**393**
T. turgidum L. ssp. carthlicum	42	65	45	155	36	114	15	40	**138**	**374**	**512**
T. turgidum L. ssp. durum 149	28	19	48	165	25	113	48	42	**149**	**339**	**488**
T. turgidum L. ssp. durum Wollaroi	17	14	26	117	25	24	47	51	**115**	**206**	**321**
T. turgidum L. ssp. durum Tamaroi	18	8	15	93	16	21	71	62	**120**	**184**	**304**
T. turgidum L. ssp. durum 139	16	17	35	47	13	20	48	42	**112**	**126**	**238**
Total	**221**	**186**	**275**	**921**	**209**	**535**	**353**	**393**	**1,058**	**2,035**	**3,093**
All 8 cultivars	4	0	3	32	6	2	3	3	**16**	**37**	**53**
All 4 *ssp. durum* samples	7	2	3	50	3	9	14	10	**27**	**71**	**98**
Total high quality markers	700		1,092		535		1,144		3,461		

Analysis shows that 3,461 high quality tissue specific markers were identified when comparing DArT scores from 8 durum wheat leaf and root tissue samples from all 4 libraries. From the 3,461 markers, 3,093 markers were identified as tissue specific in one of the 8 durum samples over the 4 libraries (totaling 27,648 clones). Analysis of polonicum samples identified 335 root specific and 159 leaf specific markers across the 4 durum libraries. In library 1, 29 root specific markers were identified, with 2 markers present in all 4 durum samples (durum line 149, durum line 139, Wollaroi and Tamaroi) but no markers present in all 8 cultivars. Of the 44 leaf specific markers, 4 markers were present in all 8 cultivars and 7 markers were present in the 4 durum cultivars (including the 4 markers present in all 8 cultivars). Figure 3.3 shows the 4 markers identified in all 8 cultivars with their sample ratio median scores. As can be seen, marker 801705322001_H-19 has all 8 leaf and root scores clearly separated with a clear divergence between the upper leaf cluster (scored present) and the lower root cluster (scored absent). This is similar for all 4 markers even though the graph appears as if marker 801705321008_B_1 has a reduced divergence between clusters. However, when comparing each leaf and root pair for each cultivar, it can be seen that this is not the case. Even though the leaf score for durum line 149 is close to the root score for the carthlicium line 414 sample, it shows considerable divergence from its durum line 149 root pair.

Marker 801705322001_H_19 is one of the 7 durum specific markers that score polymorphic for leaf and root tissue in library 1. The sample ratio median scores for marker 801705322001_H_19 are shown in figure 3.3 and separately in figure 3.4, where there is distinct divergence between leaf and root scores in all 4 durum cultivars. If the data was shown bimodally as in table 3.2, the green leaf scores would be given a value of ‘1’ being scored present and the red root scores ‘0’ scored absent. Further analysis can be performed for all markers giving similar results.

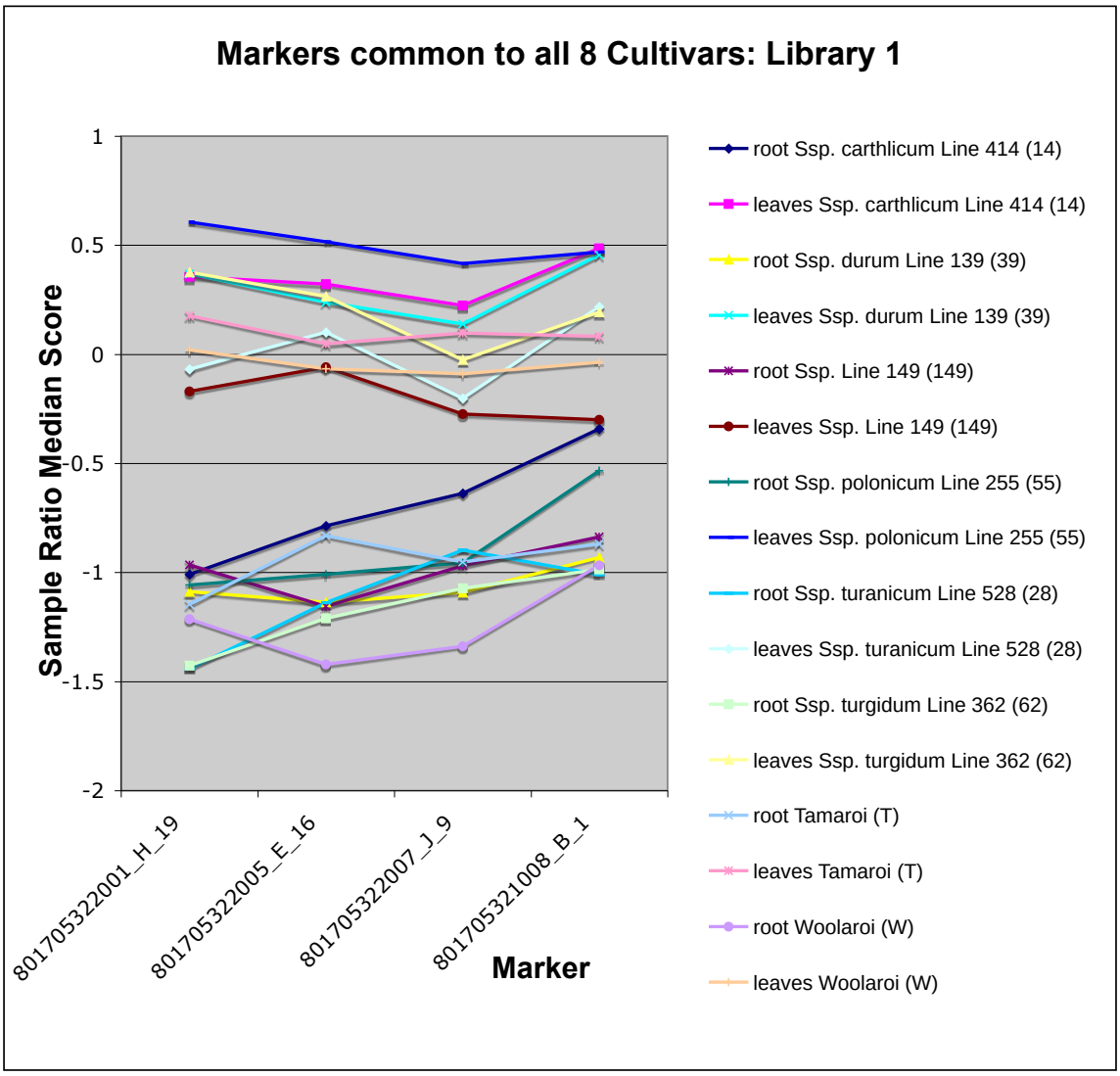


Figure 3.3: Library 1 tissue specific markers common to all 8 cultivars

The 4 markers shown are scored for all 8 cultivar leaf and root samples, however are scored bimodially as present or absent. This is shown in the graph as two distinct clusters. Some marker scores are more segregated than others.

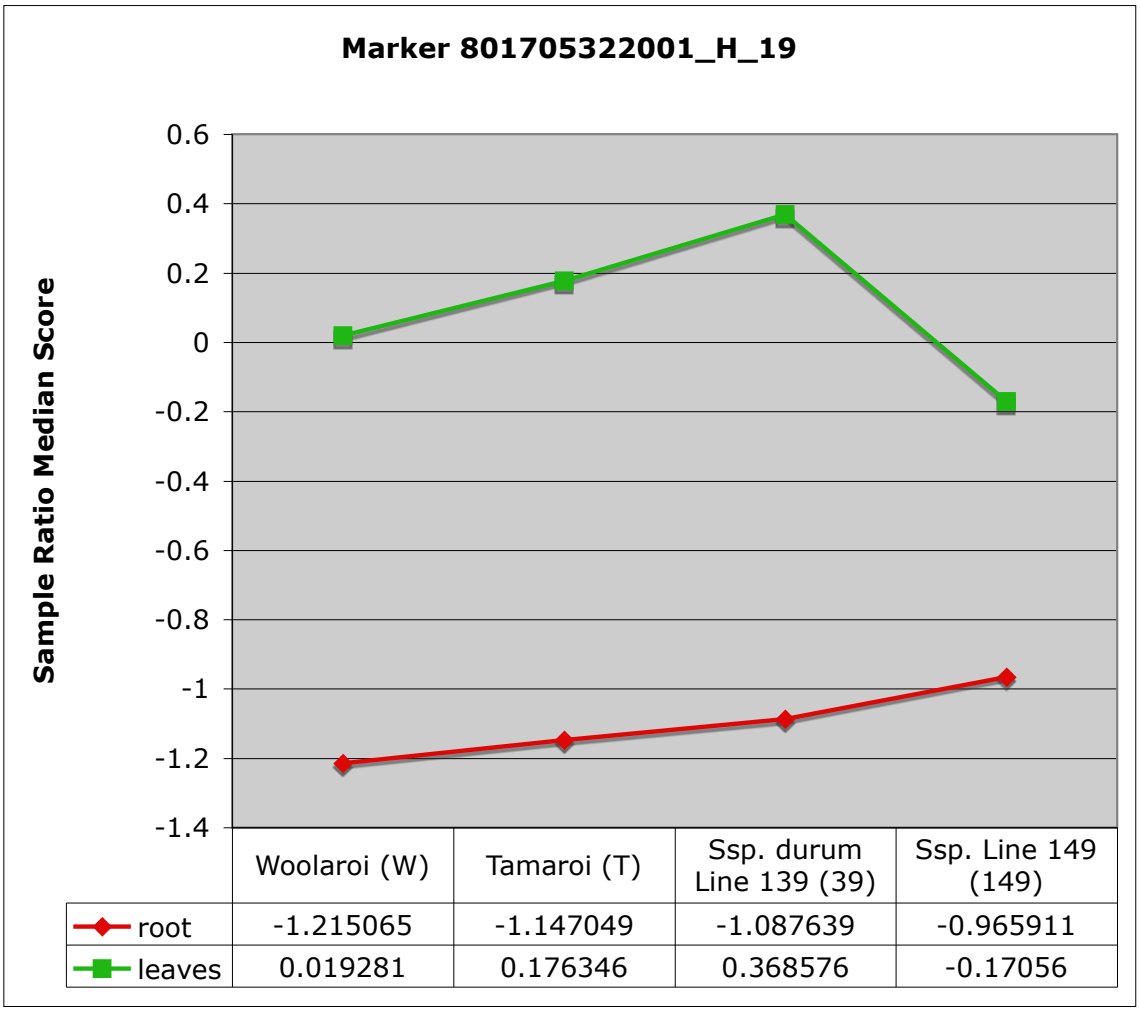


Figure 3.4: Library 1 tissue specific marker common to 4 durum cultivars**. When leaf and root tissue samples are compared against sample ratio median scores there is a clear separation between the two polymorphic groups for marker 801705322001_H_19.**

3.5.4 Conclusion

From the data analysed, there is considerable variation in leaf and root tissue samples from the 8 cultivars analysed using the DArT method. Markers were identified across all 8 cultivars that can distinguish between leaf and root samples. Due to experimental design, leaf tissue samples were duplicated and scores averaged to give a sample score. This process allows for scores with a high variance to be eliminated from analysis, as they are less reliable. The root samples were performed singularly, with no replicate for comparison. From this, it is expected that more root tissue specific markers will be identified compared to leaf specific markers.

3.6 Conclusion and Discussion

Developmental diversity studies employ many techniques to identify and measure differences in plant growth stages and tissue types. Described here is the DArT method that was successfully used to identify potential molecular markers that can be used in plant breeding programs. These markers can be identified as present or absent in seedlings before they are used in breeding experiments. Markers that have been identified should be used in screening experiments or genotyping experiments where the presence or absence of the marker can be linked to a trait of interest, such as a disease resistance locus or a physiological trait, such as dough strength. These areas of interest are called a quantative trait locus (QTLs). QTLs are stretches of DNA, in this case the marker fragment, that are closely linked to the genes that underlie the trait in question. This concept will be explored in subsequent chapters.

Molecular Physiology

“I have not failed. I've just found 10,000

ways that won't work.”

Thomas A. Edison

Chapter 4

4.0 Molecular Physiology

4.1 Introduction

Molecular physiology refers to the study of how processes are regulated at the molecular level within an organism. On a broad scale, molecular physiology investigates structures, biophysics, ion channels, transporters and pump functions, protein trafficking, cell membrane functions, cellular interactions, signal transduction, intracellular messengers and the integration of signaling. Specifically, as used in this chapter, cell growth regulation by the use of differential genomic DNA mechanisms will be explored, looking at molecular diversity in wheat plants grown under differing environmental conditions. Cultivated hexaploid bread wheats were subjected to varying light and temperature conditions with DNA extracted and analysed using DArT. In addition, wild and cultivated tetraploid durum wheats were subjected to differing salt treatments, with similar DArT analysis performed. A breeding experiment between wild and cultivated durum wheat was performed, with phenotypic data characterised and compared to molecular data generated using DArT. DArT was also used to detect methylated DNA polymorphic markers in a diverse durum wheat experiment and between 2 cultivated bread wheats.

4.2 Light and temperature stress in bread wheat

4.2.1 Aims

Molecular physiological diversity was explored using Janz and Kukri, two common Australian bread wheat cultivars. Leaf and root tissue was used to study polymorphisms between plants grown at varying environmental conditions. Control plants were grown at an optimal 20°C and at approximately 250 lumens, with varying environmental conditions used to ‘stress’ the plants. Below optimal temperatures of 10°C and above optimum temperatures at 30°C were used on both cultivars. In addition, plants were subjected to optimal light conditions of approximately 250 lumens and below optimum at 75 lumens. The aim of these experiments is to determine if there are any potential molecular markers that can be identified to predict plant viability under these conditions.

4.2.2 Specific Methods

Hexaploid Janz and Kukri wheat seeds were germinated and grown at the Research School of Biological Sciences (RSBS) at the Australian National University (ANU) and at the Centre for Molecular Biology to International Agriculture (CAMBIA). All plants were grown in temperature and light controlled growth cabinets. Cabinets were set to 10°C, 20°C and 30°C at both optimal (250 lumens) and low (75 lumens) lighting conditions.

4.2.3 Results

4.2.3.1 Comparison of temperature conditions

DArTsoft polymorphism analysis of 25 Janz and 21 Kukri samples grown at 10°C, 20°C and 30°C identified 539 high quality polymorphic markers. Of these, the majority of polymorphisms were cultivar specific, differentiating between Janz and Kukri samples. Cultivar comparisons are described further in Chapter 5. There were no markers that reproducibly scored polymorphic between plants grown at the three temperature treated groups.

4.2.3.2 Comparison of light conditions

Further analysis compared samples grown at optimal light conditions of approximately 250 lumens and low light conditions of approximately 75 lumens. Analysis was performed using 20 optimal light Janz samples and 6 shaded light Janz samples. The reduced sample size for low light samples was due to the reduction in viable plants and tissue samples. DArT analysis found no polymorphisms between the two groups. Analysis of 13 optimal light and 2 low light Kukri samples identified 31 polymorphic markers. As the number of Kukri samples was smaller than for Janz, a lower proportion of markers are eliminated from analysis for failing to fall into the bimodal clusters across replicates, thus a greater number of polymorphic markers are identified. Analysis was further limited to samples grown at 30°C from leaf samples and optimal and low light conditions compared. In Janz, 3 markers were scored present in optimal light samples and absent in low light samples, and in Kukri, 78 markers were identified as polymorphic, with 21 scoring present in optimal light and 57 markers present in low light samples. Similarly, the Kukri numbers are higher

due to lower sample sizes, indicative that there is less material to compare. Results are summarised in table 4.1.

Table 4.1: Polymorphic markers identified when comparing optimal and shaded light growth conditions in Janz and Kukri

Cultivar	Number of high quality polymorphic markers	Number of light specific polymorphic markers	Number of polymorphisms scored present in	
			Shaded light	Optimum light
Janz	562	0	0	0
Kukri	562	31	31	0
Janz 30°C leaf	562	3	0	3
Kukri 30°C leaf	562	78	57	21

4.2.4 Conclusions

From the analysis, it can be seen that the detection of polymorphic markers between wheat that was grown at varying temperatures and light conditions is not efficient and reliable. Reasons could be that the stress incurred wasn't sufficient or wasn't applied long enough to detect any differences between genomic representations. Sample sizes also varied, contributing to the variance in numbers. A more reliable test is needed.

4.3 Salt stress in durum wheat

4.3.1 Aims

As climate change increasingly affects crop production, tolerance to salt is of major phenotypic importance with soil salinity the major abiotic stress in plant agriculture worldwide[187]. Salt tolerance is investigated using two durum wheats with estimated phenotypic tolerance to salt conditions, both being screened for rates of Na+ uptake

and K+/Na+ discrimination.[188] Polonicum is a wild durum cultivar from Poland that is relatively salt in tolerant. Analysis of replicated plants grown in 200 mM NaCl produced a percentage dead leaf of 4.2%, mean chlorophyll estimate of leaves 1, 2 and 3 from the main stem of 31.5 SPAD units, total Na+ percentage of dead leaf of 93 µmol and Na+ concentration in dead leaf of 3.81 mmol g DW^{-1}.[189] In comparison, durum line 139, a cultivated durum grown commercially in Australia is also relatively salt tolerant, with a percentage dead leaf of 5.1%, mean chlorophyll estimate of leaves 1, 2 and 3 from the main stem of 30.8 SPAD units, total Na+ per percentage of dead leaf of 42 µmol and Na+ concentration in dead leaf of 3.94 mmol g DW^{-1}.[190] This data shows that polonicum has a ~2 fold higher total Na+ per percentage of dead leaf score durum 139 and diversity analysis between them could identify possible polymorphic markers relevant to salt tolerance studies and plant breeding programs.

4.3.2 Specific Methods

Polonicum and durum tetraploid wheat was grown in a glass house at the Crown Scientific and Industrial Research Organisation (CSIRO) on Black Mountain. Plants were grown in quadruplicate under control, incrementally increased salt and all-at-once shock salt conditions (as described in table 4.2). The control conditions contained 1 mM NaCl with the incremental samples increased to 200 mM NaCl over 21 days. Salt shock samples were grown as controls with the addition of 200 mM NaCl several days before harvesting. Experiments were performed that used replicate targets fluorescently labeled in 2 and 3 colours. Genomic representations were created using PstI and $Taq^{\alpha}I$ restriction enzymes in comparison to PstI, $Taq^{\alpha}I$ and McrBC genomic representations. The DArT wheat 8 plate V2.2 array was used to analyse polonicum and durum samples. In addition, a PflMI + MseI +/- McrBC and PstI + MseI +/- McrBC array were created for further analysis. These two 4-plate arrays were used to increase the complexity of the analysis by substituting the standard PstI and $Taq^{\alpha}I$ restriction enzymes for PflMI and MseI. A further reason was to determine if the percentage of polymorphic markers found differed between complexity reduction restriction enzymes and to see if $Taq^{\alpha}I$, which is Dam methylation sensitive, changed the proportion of polymorphic markers found.

Table 4.2: DNA samples used in Experiment One

Sample	Genotype	Treatment	Tissue
A1 durum	61-39	Control	Leaves
B1 durum	62-39	Control	Leaf expansion zone
A3 durum	146-39	Incremental	Leaves
B3 durum	147-39	Incremental	Leaf expansion zone
G1 polonicum	73-55	Control	Leaves
H1 polonicum	74-55	Control	Leaf expansion zone
G3 polonicum	155-55	Incremental	Leaves
H3 polonicum	156-55	Incremental	Leaf expansion zone
G7 durum	P-39	Seedling	Leaves
G8 durum	P-39	Seedling	Root
C7 polonicum	P-55	Seedling	Leaves
C8 polonicum	P-55	Seedling	Root

Analysis was performed using 12 samples in duplicate from plants grown from seedlings, control conditions and from incrementally increasing salt conditions. DNA was used from leaf, root and the leaf expansion zone tissue and labeled using Cy3. Targets were prepared identically but with the addition of McrBC to cut methylated DNA. These targets were labeled with Cy5 and hybridised with the corresponding Cy3 sample on the same microarray slide. Analysis was performed on control and incrementally increased salt samples, excluding the seedling samples as they were shown to be scored almost the same as controls. The experiment was expanded to include 4 replicates of each target, made using PstI and Taq$^{\alpha}$I labeled with Cy3, and PstI, Taq$^{\alpha}$I and McrBC labeled with Cy5. The 8 samples in quadruplicate were hybridised to 32 microarray slides printed with the same microarray as experiment one.

4.3.3 Results

4.3.3.1 Comparison of Seedling and Control samples

Samples extracted from seedling leaf tissue, seedling root tissue, control leaf tissue and control leaf expansion zone tissues were compared using DArTsoft as all should give a similar result. Duplicate samples from both targets (Cy3 and Cy5) were analysed together as replicates. A '0 threshold' setting was used so that all clones were analysed, as not to reject any clone. Clones were then sorted in descending order for Q values and then for Reproducibility. Limiting clones to a Q value higher than 75 and reproducibility above 90%, 419 clones were identified. For the *durum* samples, 3 clones out of the 419 clones were scored oppositely in one of the 4 tissue samples over all 4 replicates. The polonicum samples produced 4 clones that varied within the 4 tissue samples. This included one clone that was scored absent for both leaf samples but present for root and leaf expansion zone (clone 801504280001_E_5) and one clone that was scored absent in control tissue samples but present in seedling tissue samples (clone 801504280003_C_6). These markers are shown in figure 4.3 Interestingly, this clone was scored identically for incrementally increased salt conditions for leaf and leaf expansion zone tissue samples. The other 2 clones were scored oppositely in one of the 4 tissue samples over all 4 replicates. A scoring inconsistency is show in the analysis as a 'X', as apposed to a presence (1) or absence (0), when there is a discrepancy in scoring across replicates. This can be due to technical errors, such as scratches or debris on or near the spot that is being analyse. Durum samples showed 14 scoring discrepancies out of the 1,676 scores (419 clones x 4 tissue samples), while the polonicum samples showed 58 discrepancies. Thus we can conclude that the durum and polonicum samples are scored upwards of 96% identically for each sub species for leaf, root and leaf expansion zone tissue samples in both Cy3 and Cy5 replicates. This is important, as by using two different dyes, polymorphism detection is highly accurate with only a small variation in samples reported. This will play an important role when the methylation sensitive enzyme McrBC is added to targets and comparisons made.

4.3.3.2 Comparison of Seedling, Control and Incremental samples

Experiment one also included corresponding samples that were grown with increasing salt concentrations. This data matched the seedling and control data in that sequence based candidate polymorphisms were scored in the same way for durum and polonicum. This specifically identified clones that showed tissue polymorphisms, such as clone 801504280001_E_5 which showed tissue specificity in polonicum for the absence of the fragment in control leaves but present for the leaf expansion zone or roots, showed absence in both incremental leaf and leaf expansion zone samples. The same clone was scored present in durum over all seedling, control and incremental samples. Clone 801504280003_C_6 was scored absent in polonicum control samples and present for seedling samples, however for the incremental samples, the leaf tissue was scored absent and leaf expansion zone present. Interestingly, the same clone was scored present in control and seedling samples for durum, but was absent in both tissues for the incremental samples. This suggests that for durum, the fragment was eliminated from the DNA pools indicating that it may have been methylated and destroyed from amplification by McrBC. For the polonicum samples, it suggests that for the fragment is absent in control tissue and leaf incrementally salt treated, but present in seedling tissue and incrementally salt treated leaf expansion zone.

4.3.3.3 Comparison of Control and Incremental samples

Similar to experiment one, experiment two looked at diversity among the two sub species but also between treated and non-treated samples. Control samples were grown with normal salt concentrations, however incremental samples were grown with increasing concentrations of salt. Analysis was performed using the ‘0 threshold’ setting in DArTsoft between all samples and limited the list to only clones that scored equal or above 75 for Q and equal of above 90% for reproducibility. From this, 151 candidate polymorphic clones were identified, all of which scored identically over the 8 replicates (4 replicates in Cy3 and Cy5) for each species, with only one inconsistent score. These clones represent potential sequence based polymorphisms as they were scored with high Q values and were highly reproducible, both in the Cy3 and Cy5 channel and over spot and slide replicates.

Table 4.3: Experiment One Results

Clone Identification		P	Q	Reproducibility	Ssp durum				Ssp polonicum			
					Control		Seedling		Control		Seedling	
					leaves	GR	leaves	root	leaves	GR	leaves	root
1	800904300002_I_17	96.374	94.76	100	0	0	0	0	1	1	1	1
2	801504280004_K_14	96.364	94.75	100	0	0	0	0	1	1	1	1
3	801504280005_E_1	96.640	94.746	100	0	0	0	0	1	1	1	1
4	800904300001_I_1	96.233	94.629	100	1	1	1	1	0	0	0	0
5	801504280004_P_24	96.111	94.50	100	1	1	1	1	0	0	0	0
6	801504280004_L_23	96.020	94.420	100	1	1	1	1	0	0	0	0
7	801504280003_C_6	89.691	87.932	91.6	1	1	1	1	0	0	1	1
8	801504280001_E_5	78.702	77.524	100	1	1	1	1	0	1	0	1

Candidate polymorphic clones identified with DNA fragments present ‘1’ or absent ‘0’ for ssp. durum and ssp. polonicum, showing seedling (treatment) variation in polonicum sample 7 for leaf and root tissue and tissue variation in polonicum sample 8 for both leaf samples.

To look outside this range, it can be seen that clones vary quite considerably. Reproducibility of 100 indicates that all samples were scored the same for all replicated targets (either on different slides or labeled with different dyes on the same slide). As we have 8 target replicates in experiment two, one discrepancy in scoring would reduce the reproducibility by 12.5%, below that if the 90% cutoff. Looking at scores above one discrepancy, the number of clones increases from 151 to 285. Analysis of these clones reveals much more diversity in scoring patterns, in that durum shows one candidate polymorphism that is scored present for the control sample but absent for the incrementally increased salt sample (801504280003_M_5). Polonicum showed two treatment based candidate polymorphisms, one where the control is scored absent and the sat treated present (801504280005_D_8) and the other showing the opposite pattern (800904090002_D_6). Moving to a 2 out of 8 discrepancy, that is a reproducibility of 75%, 365 candidate polymorphisms are found, an extra 80 that are all cultivar specific and not tissue or treatment based.

4.3.3.4 Comparison of PstI and Taq$^{\alpha}$I (Cy3) targets

As experiment one and two both used the same control and incremental samples, their individual analysis results can be compared. Looking at PstI and Taq$^{\alpha}$I data only, that is Cy3 targets, a list of candidate clones can be compiled for each experiment. As we have used the '0 threshold' setting, the list needs to be sorted so that all clones are not included. A Q value of 75 and a reproducibility of over 85% was used, to include clones with 0 or 1 scoring discrepancy. From this, experiment one generated 471 clones, with several different categories of polymorphisms. A treatment specific polymorphism means that the sample is scored either present of absent for a treatment for both samples, and the opposite score for the other treatment. A tissue and treatment specific polymorphism shows a score for a treatment in both tissues and then one treatment on the other treatment with the same score in one tissue and the opposite score in the other tissue. A tissue specific polymorphism shows a score for both treatments in the same tissue and opposite in the other tissue (leaf versus leaf expansion zone or root). For durum samples, there were 5 treatment specific and 5 treatment and tissue specific polymorphisms identified. For polonicum, 4 treatment and 12 treatment and tissue and 3 tissue specific polymorphisms were observed. While experiment two generated 387 clones, there was considerable less diversity.

In polonicum, only 2 treatment and 1 treatment and tissue specific polymorphic clones were identified. In durum, only 2 treatment and tissue specific polymorphisms were identified. All other clones were sequence based between the two cultivars.

Comparisons of these two candidate polymorphism lists showed 512 clones, with 345 (67.40%) identified in both experiments as sequence polymorphisms. 125 (24.40%) were identified in experiment one only, with durum showing 6 treatment specific and 6 treatment and tissue specific polymorphisms. Polonicum samples showed 4 treatment, 11 treatment and tissue specific and 3 tissue specific polymorphisms. Experiment two identified 42 (8.20%) clones with durum showing 1 treatment and tissue specific polymorphism while polonicum showed 2 treatment polymorphisms.

4.3.3.5 Comparison of PstI, Taq$^{\alpha}$I and McrBC (Cy5) targets

In the same way as clones were analysed for Cy3 (section 5.3.3), Cy5 targets in experiments one and two were analysed separately to generate two lists of clones limited by scores equal or above 75 for Q and 80% for reproducibility. Both experiments shared 374 (57.98%) out of the total 645 clones all being sequence based between durum and polonicum. A further 240 (37.20%) clones were identified in experiment one, and 31 (4.80%) clones in experiment two. From durum samples, experiment one identified 10 treatment specific, 24 treatment and tissue specific and 7 tissue specific polymorphisms. Experiment two only identified 1 tissue specific polymorphism, differing between leaf and root/leaf expansion zone between both treatments. Looking at polonicum, experiment one identified 4 treatment specific, 42 treatment and tissue specific and 9 tissue specific polymorphisms. Experiment two identified 3 treatment specific and 4 treatment and tissue specific polymorphisms.

4.3.4 Conclusions

From the data presented, it can be seen that polymorphism between durum and poloncium samples canbe used as potrential DNA markers to differenciate between cultivar samples. The first 6 markets in table 4.3 show that the presence or absence of a single or multiple markers can be used to genotype these samples as durum or

polonicum line. These potential markers will have varying expression patterns when compared to other cultivares.

4.4 RIL Salt stress in durum wheat

4.4.1 Aims

A breeding experiment was performed between two diverse durum cultivars, a wild carthlicum cultivar and the Australian durum cultivar Wollaroi. The 2 parents and 94 progeny were selfed and DNA extracted from the 6th filial generation (F6). Experiments were designed to look at the segregation of polymorphic markers with the ultimate aim to identify salt tolerant associated markers.

4.4.2 Specific Methods

4.4.2.1 Tissue samples

Durum wheat was grown in conjunction with CSIRO Plant Industries at the laboratory at Black Mountain in Canberra[191]. The wheat samples were sourced and grown by Dr Rana Munns and Dr Richard James in plant growth rooms at constant temperatures and lighting conditions, including day/night cycles. All samples were grown in either duplicates or quadruplicate and tissue samples taken from the leaf and the leaf expansion zone (tiller) of the plant. Table 4.4 shows the durum wheat samples grown.

4.4.2.2 Seedling tissue samples

In addition to the DNA extracted from the 8 genotypes in table 4.4, seeds were germinated from the same species as an additional control. Leaf and root DNA was extracted from seedlings grown for an average of 14 days as shown in table 4.5.

Table 4.4: Durum Wheat samples used for Salt tolerance experiments

Designation	Species	Tissue
14	*T. turgidum L. ssp.* carthlicum	Leaves / Growing region
149	*T. turgidum L. ssp.* durum 149	Leaves / Growing region
55	*T. turgidum L. ssp.* polonicum	Leaves / Growing region
W	*T. turgidum L. ssp.* durum Wollaroi	Leaves / Growing region
62	*T. turgidum L. ssp.* turgidum	Leaves / Growing region
28	*T. turgidum L. ssp.* turanicum	Leaves / Growing region
39	*T. turgidum L. ssp.* durum 39	Leaves / Growing region
T	*T. turgidum L. ssp. durum* Tamaroi	Leaves / Growing region

Table 4.5: Durum Wheat seedling samples used for Array development

Species	Genotype #	Treatment	Tissue
T. turgidum L. ssp. carthlicum	14	Control	Leaves / Roots
T. turgidum L. ssp. durum 39	39	Control	Leaves / Roots
T. turgidum L. ssp. durum 149	149	Control	Leaves / Roots
T. turgidum L. ssp. polonicum	55	Control	Leaves / Roots
T. turgidum L. ssp. turanicum	28	Control	Leaves / Roots
T. turgidum L. ssp. turgidum	62	Control	Leaves / Roots
T. turgidum L. ssp. durum Tamaroi	T	Control	Leaves / Roots
T. turgidum L. ssp. durum Wollaroi	W	Control	Leaves / Roots

4.4.2.3 Recombinant Inbred Line (RIL) tissue samples

A cross between Wollaroi (W) and T. turgidum L. ssp. polonicum (55) was performed and 96 progeny self-fertilised until the 6th filial generation (F6). All plants were grown in climate controlled glass house at CSIRO Black Mountain. This work was performed by Rana Munns *et al* at CSIRO Plant Industry.[192] Leaf material was harvested from the first leaf of the second tiller where 5cm of the leaf tip was removed from each genotype in quadruplicate.

4.4.2.4 Array design

Four arrays were developed to explore diversity between durum wheats. Three arrays were made using DNA samples from T. turgidum L. ssp. durum 39 and T. turgidum L. ssp. polonicum. DNA was extracted from both sub species and included two treatment conditions, Control and Shock. Table 4.6 below describes the tissue samples used for the following three arrays.

Table 4.6: Durum Wheat samples used for array development

Species	Genotype	Treatment	Tissue
T. turgidum L. ssp. durum 39	246-39	Shock	leaves
T. turgidum L. ssp. durum 39	247-39	Shock	leaf expansion zone
T. turgidum L. ssp. polonicum	255-55	Shock	leaves
T. turgidum L. ssp. polonicum	256-55	Shock	leaf expansion zone
T. turgidum L. ssp. durum 39	61-39	Control	leaves
T. turgidum L. ssp. durum 39	62-39	Control	leaf expansion zone
T. turgidum L. ssp. polonicum	73-55	Control	leaves
T. turgidum L. ssp. polonicum	74-55	Control	leaf expansion zone

4.4.2.5 PstI and Taq$^{\alpha}$I library created from Durum and Polonicum

Restriction enzymes PstI and Taq$^{\alpha}$I were used to reduce the complexity of the genomic representation and included the use of adapter PstI_adapter1 ligated to PstI_adapter2. Four 384 well libraries were created totaling 1,536 clones from DNA in table 4.6. Arrays were printed in duplicate on a standard glass microarray slide.

4.4.2.6 PstI and MseI library created from Durum and Polonicum

Restriction enzymes PstI and *MseI* were used to reduce the complexity of the genomic representation and included the use of adapter PstI_adapter1 ligated to PstI_adapter2. Two 384 well libraries were created totaling 768 clones from DNA in table 4.6. Arrays were printed in quadruplicate on a standard glass microarray slide.

4.4.2.7 PflMI and MseI library created from Durum and Polonicum

Restriction enzymes PflMI and MseI were used to reduce the complexity of the genomic representation and included the use of AdaptC_rev ligated to C-fwd11. Two 384 well libraries were created totaling 768 clones from DNA in table 4.6. Arrays were printed in quadruplicate on a standard glass microarray slide so that 3,072 spots were present on each array.

4.4.2.8 PstI and Taq$^{\alpha}$I library created from 8 genotypes

The fourth array was created using DNA extracted from seedlings grown from 8 different sub species, as listed in table 4.4. PstI and Taq$^{\alpha}$I restriction enzymes were used with PstI_adapter1 ligated to PstI_adapter2, to produce a 4 plate, 1,536 clone library. The array was printed in duplicate on a standard glass microarray slide.

4.4.2.9 Target Production

Targets were prepared using DNA from tissue samples using the same methods described for library creation. Targets were made in duplicate using the corresponding restriction enzymes to the library that they are to be hybridised to, with one set of targets further digested with McrBC. The McrBC restriction enzyme

cleaves DNA (either or both strands) between two methylated A/TmG sites, thus allowing for the comparison between McrBC treated and non-treated samples.

4.4.3 Results

4.4.3.1 Analysis of seedling and control samples

DArTsoft polymorphism analysis of seedling and control leaf samples from polonicum and durum 39 leaf material was compared using a PstI and Taq$^{\alpha}$I array. Analysis of PstI and Taq$^{\alpha}$I restriction enzyme generated targets (Cy3 labeled) identified 591 high quality markers, limited to a Q value of 75 or greater, reproducibility maximum of 1 scoring discrepancy and a call rate of 80 or greater. Of these markers, the majority are polymorphic between cultivars, however, 28 markers were identified as polymorphic between durum seedling leaf and control leaf samples. Of these, 17 markers were scored present in control leaf and 11 markers present in seedling leaf samples. In polonicum, 32 markers were identified that differentiated between leaf tissue in seedling and control samples, with 16 markers scored present in control leaf and 16 markers in seedling leaf samples. It would be expected that seedling and control material be highly similar, as they both represent the non-treated phenotype.

Analysis of PstI, Taq$^{\alpha}$I and McrBC restriction enzyme generated targets (Cy5 labeled) identified 592 high quality polymorphic markers. Of these, 17 markers were identified as polymorphic between durum seedling leaf and control leaf samples. Of these, 7 markers were scored present in seedling leaf and 10 markers in control leaf samples. In polonicum, 27 markers were identified as polymorphic, with 14 markers scored present in seedling leaf and 13 markers in control leaf samples.

Comparison of the 28 markers identified in durum Cy3 targets with the 17 markers identified in durum Cy5 analysis shows 2 markers found in both analyses, 26 markers only identified in Cy3 and 15 markers only identified in Cy5 analysis. Comparison of the 32 markers identified in polonicum Cy3 targets with the 28 markers identified in polonicum Cy5 analysis shows 1 markers found in both analyses, 31 markers only identified in Cy3 and 37 markers only identified in Cy5 analysis.

4.4.3.2 Analysis of control and incremental samples

4.4.3.2.1 PstI and Taq$^{\alpha}$I library

DArTsoft analysis was used to detect polymorphic markers between control samples and incremental samples, those grown with increasing quantities of salt. The experiment was duplicated, and the two experiments are compared, with the number of polymorphisms identified between samples shown in Table 4.7. Experiment 1 identified 591 high quality markers in Cy3 and 592 markers in Cy5, with 146 (25%) markers identified in both wavelengths. Experiment 2 identified 389 high quality markers in Cy3 and 476 markers in Cy5, with 339 (87%) markers identified in both wavelengths. Theoretically, all Cy5 markers should be identified in the Cy3 analysis, and those identified in Cy3 but not Cy5 can be attributed to the McrBC restriction enzyme destroying the fragment during genomic representation construction.

Table 4.7: Number of markers identified between control and incremental samples for leaf and leaf expansion zone for each cultivar

	Cy3						Cy5					
Tissue	**Leaf**		**LEZ**		**Total**		**Leaf**		**LEZ**		**Total**	
Control (C) or Incremental (I)	C	I	C	I	C	I	C	I	C	I	C	I
Experiment 1 – durum	9	10	9	9	18	19	22	16	19	9	41	25
Experiment 2 – durum	0	1	1	0	1	1	0	0	0	0	0	0
Experiment 1 – polonicum	13	25	14	23	27	53	22	12	8	27	30	39
Experiment 2 - polonicum	1	2	0	3	1	5	3	4	9	9	12	13

* LEZ = Leaf expansion zone.

From table 4.7, experiment 1 identified considerable more polymorphic markers between the control and incremental samples in both leaf and leaf expansion zone samples. This could be contributed to the poor (25%) correlation between Cy3 and Cy5 markers identified as polymorphic between cultivars, probably due to non-optimal experimental conditions.

4.4.3.2.2 PflMI and MseI library

Analysis of polonicum and durum samples from control and incremental environmental stress conditions was performed using the PflMI and MseI array. DArTsoft polymorphism analysis was used to identify polymorphic markers between the 2 durum cultivated wheat samples. Targets were generated using PflMI and MseI restriction enzymes, labeled with Cy3 and hybridised in quadruplicate with 4 copies of each of the 768 clones printed on the array. McrBC methylation sensitive restriction enzyme was added to a replicate of each target, labeled with Cy5 and hybridised to the corresponding Cy3 target on the same array. DArTsoft polymorphism analysis identified 31 high quality markers in Cy3 and 41 in Cy5, limited to a Q value of 75 or greater, a reproducibility maximum of one scoring discrepancy and a call rate of 80 or greater. Of these, 12 markers were polymorphic between treatments groups. Comparisons between control and incremental durum leaf samples identified 1 marker present in control samples only and 11 markers scored present in incremental samples only. Comparison of durum leaf expansion zone tissue between control and incremental samples identified 10 treatment specific polymorphisms, all being scored present in incrementally increasing salt samples and absent in the control samples. Analysis of polonicum samples identified 5 treatment specific markers between control and incremental lead samples, 1 marker scored present in control samples and 4 markers present in incremental samples. Analysis of leaf expansion zone samples identified 2 markers, both scored present in incremental samples only.

Comparison between leaf samples from both cultivars identified 5 markers that were scored polymorphic, all present in the salt treated incremental samples and absent in control samples in durum and polonicum samples. Between the leaf expansion zone samples for durum and polonicum, 1 marker was identified that was polymorphic for both cultivars, scored present in incremental samples and absent in control samples. Analysis of the 41 high quality markers identified in Cy5 samples that is targets generated using PflMI, MseI and McrBC, found 3 that were polymorphic between control and incremental treatments. Of these, 1 marker was scored present in control durum leaf samples, 1 marker present in durum leaf expansion zone samples and 1 marker present in incremental leaf samples. These 3 markers were also identified in Cy3 analysis, however were scored non polymorphic between treatments. They also scored Q values of below the 75 threshold, thus not included in the high quality marker analysis. As the McrBC restriction enzyme destroys methylated fragments, it

can be theorized that the fragments not identified in Cy5 analysis but scored polymorphic in Cy3 analysis were methylated. Data from the three markers does not necessarily show this due to the poor scoring in Cy3 analysis.

4.4.3.3 Analysis of control, incremental and shock samples

4.4.3.3.1 PfiMI and MseI library

The PflMI and MseI experiment was replicated from section 4.3.3.2.1 with the addition of shock salt treatments samples with all targets duplicated. Durum and polonicum samples were compared for leaf tissue across control, incremental and shock treatments and compared against incremental and shock conditions for leaf expansion zone tissue. The addition of McrBC was used in targets labeled with Cy5. Analysis of durum leaf in Cy3 identified 96 polymorphisms between the 5 samples, with 12 markers polymorphic between treatments. Of these, 2 markers were identified that scored absent in control samples, but present in both salt treated incremental and shock samples. A further 2 markers wear identified as polymorphic in incremental samples, being scored absent but present for control and shock samples. Comparisons identified 4 markers that were scored shock treatment specific, with 3 scored present in shock samples but absent in control and incremental samples. Analysis of durum leaf expansion zone tissue identified 8 polymorphisms, with 7 polymorphic markers scored present in shock samples and absent in incremental samples, and 1 marker scored in reverse. Between tissue types, 3 markers were scored absent for incremental and present for shock in both tissue types.

Analysis of polonicum samples identified 38 polymorphisms. Of these, 2 markers were identified as control treatment specific, 1 marker scoring present and I marker absent only in control leaf polonicum samples. Analysis found 3 markers that were Incremental specific, 2 markers scoring absent and 1 marker present only in incremental samples. Analysis of shock treatment samples identified 14 polymorphisms, 6 markers scoring present and 8 markers absent only in polonicum leaf shock samples. Analysis of leaf expansion zone identified 16 markers scored present in polonicum incremental samples and 8 markers present in shock samples. Comparison between tissue types for polonicum showed that 4 markers were scored present in leaf and leaf expansion zone tissue for shock samples only, with 3 of these

markers also scoring absent in control samples. Another 4 markers were identified that scored present in shock samples and absent in incremental samples for both tissue types, and 3 of these markers were scored absent in control leaf tissue. The 6 markers identified that scored polymorphic between shock and control/incremental samples for both tissue types are shown in table 4.8, where the sample ratio median scores are shown. It can be seen that 3 markers are scored present (green) and 3 markers absent (red) for shock salt treated samples. This data is shown graphically in figure 4.1.

Analysis of Cy5 targets with the addition of the McrBC restriction enzyme identified 31 high quality markers, 3 of which were polymorphic in durum leaf samples. Of these, 2 markers were polymorphic in incremental samples, one scored present and one absent, and 1 marker scored present in shock samples and absent in both control and incremental samples. Analysis of leaf expansion zone tissue identified 6 markers, 5 of which were scored present in incremental and absent in shock samples, and 1 marker present in shock samples only. Analysis of polonicum samples identified 8 polymorphic markers, with 2 markers shock treatment specific in leaf tissue, with 1 marker scored present and 1 marker scored absent. This correlated with 1 marker in leaf tissue that was scored in the same way. In the leaf expansion zone analysis, 6 markers were scored polymorphic, 5 of which were scored present in shock samples only, and 1 marker scored present in incremental samples.

Table 4.8: Polonicum leaf and leaf expansion zone shock treatment specific markers identified in Cy3 analysis showing sample ratio median scores

Clone Name	Control Leaf	Incremental Leaf	Incremental LEZ	Shock Leaf	Shock LEZ
801006230011_I_18	-2.409953	-2.451162	-2.433643	-1.308388	-1.071233
801006230011_A_2	-2.391762	-2.468609	-1.837966	-1.290243	-0.787558
801006230012_I_12	-1.612212	-1.470964	-1.24437	-0.610708	-0.154834
801006230011_D_8	1.566978	1.628372	1.836191	0.071145	0.731194
801006230011_L_4	1.80551	2.005878	1.986671	1.402277	1.30839
801006230012_J_11	2.302072	2.477009	2.426589	1.115787	1.63469

* LEZ = Leaf Expansion Zone

4.4.3.3.2 PstI and MseI library

7Analysis of durum and polonicum samples from control, incremental and shock salt treatments was performed on the PstI and MseI array. DArTsoft polymorphism analysis was used to identify polymorphic markers between the 2 cultivated wheat samples. Targets were hybridised in duplicate, with 4 copies of each clone printed on the array. Targets were generated using PstI and MseI and labeled with Cy3 and replicated with the addition of McrBC and labeled with Cy5. Analysis identified 27 high quality markers in Cy3 with 1 marker scored present and 1 marker absent in durum control leaf samples. An additional marker was scored absent in durum shock leaf samples and present in the control and incremental leaf samples. Analysis of leaf expansion zone, from control, incremental and shock samples, identified 2 polymorphic markers, 1 scored present and 1 absent in shock samples. Analysis of polonicum identified 5 treatment specific polymorphic markers, with 1 marker scored present in control leaf samples and 1 marker scored absent in shock samples. Analysis of leaf expansion zone tissue identified 2 markers that were scored present in control samples only, 1 marker scored absent in shock samples and 2 markers, 1 marker scored present and 1 marker absent in incremental salt samples.

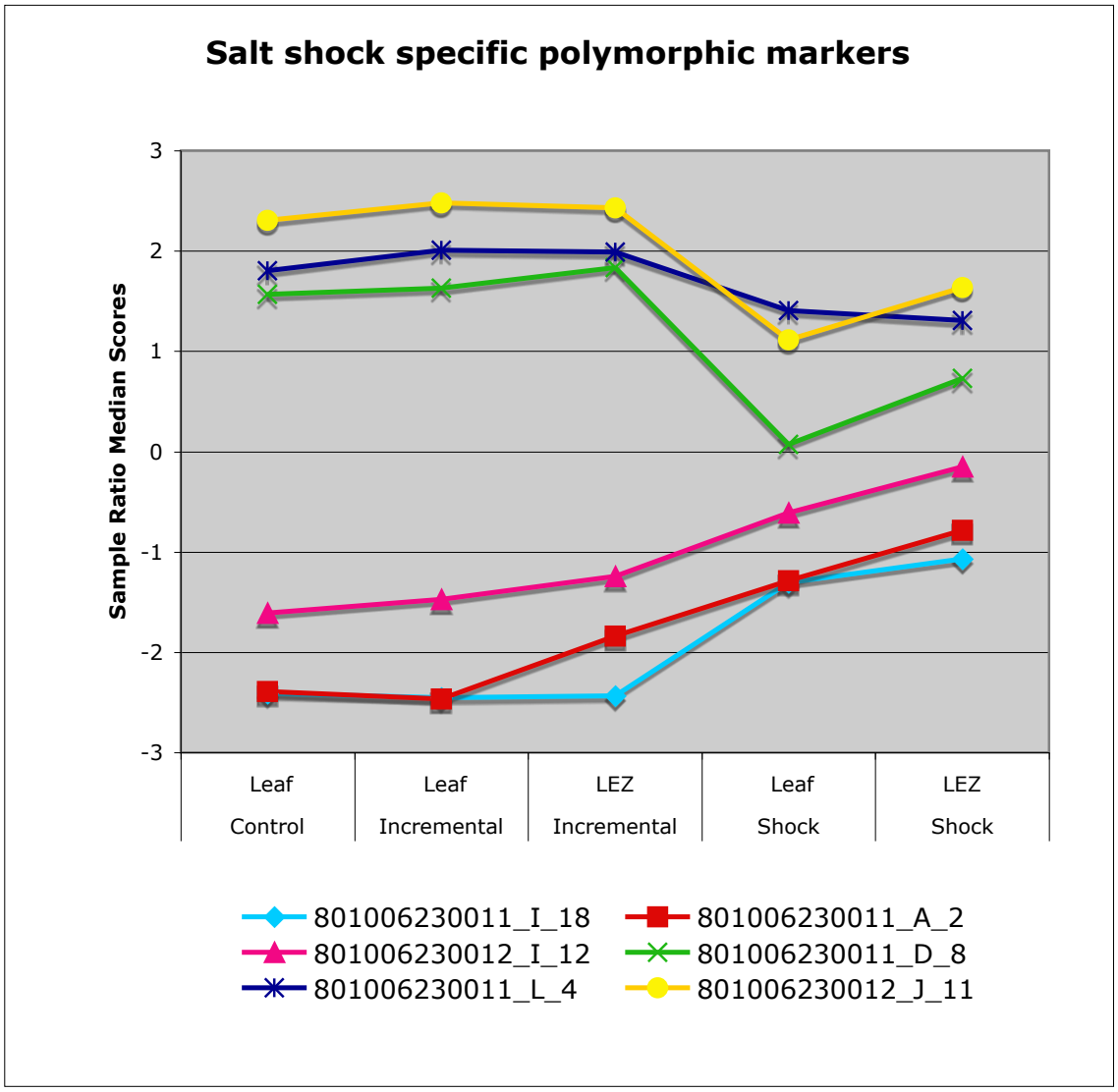


Figure 4.1: Distribution of shock specific polymorphic markers.

Distribution of ratio median scores for 6 markers from control, incremental and shock samples. There is a clear bimodal distribution of scores between control and incremental and the shock samples, showing a bimodal distribution. These markers may be able to identify a shock phenotype when scored against control samples.

Analysis of Cy5 targets with the addition of McrBC restriction enzyme identified 47 high quality polymorphic markers in the initial experiment. Of these, 10 markers were identified as polymorphic between salt treatments in durum, including 4 markers that were scored present in only control leaf tissue. A further 1 marker was scored absent in only control leaf samples and 2 markers scored absent in shock samples but present in control and incremental samples. In leaf expansion zone samples, 3 markers were scored present only in control samples, and 1 marker scored present and 1 marker absent in shock samples. Polonicum analysis identified 8 polymorphic markers, with 1 marker scored present in leaf control samples and 1 marker scored absent in leaf shock samples. In leaf expansion zone tissue, 2 markers were scored present for control samples, 3 markers scored absent in incremental samples, 1 marker scored present in incremental samples and 1 marker scored absent in shock samples.

Comparison of marker 801006230005_N_11 between leaf and leaf expansion zone tissue show that it was scored present in control and incremental samples but absent in both shock samples in polonicum. In comparisons to durum samples, the marker scored the opposite in durum leaf expansion zone tissue, scoring absent in control and incremental and present in shocks samples. In durum leaf, this marker scored absent in control and present in incremental and shock samples. Marker 801006230005_F_21 was scored present in polonicum control samples from leaf and leaf expansion zone tissue, but absent in both incremental and shock samples. Comparisons to durum samples show the marker scored present in all tissue samples except scoring absent in durum leaf shock samples. This is summarised in table 4.9 showing the binary scores for each Polonicum marker and condition (Durum data also shown).

The experiment was duplicated, with the addition of seedling leaf and root tissue. Durum analysis from seedling, control, incremental and shock samples from leaf identified 74 high quality polymorphisms in Cy3. Of these, 15 markers were polymorphic between treatments. For durum, 7 markers were scored absent in the non-treated seedling and control samples, and present in the treated incremental and shock samples.

Table 4.9: Polonicum polymorphic clones that score the same in leaf and leaf expansion zone tissue.

	Durum						Polonicum					
	Leaf			LEZ			Leaf			LEZ		
Marker	**C**	**I**	**S**	**C**	**I**	**S**	**C**	**I**	**S**	**C**	**I**	**S**
801006230005_N_11	0	1	1	0	0	1	1	1	0	1	1	0
801006230005_F_21	1	1	0	1	1	1	1	0	0	1	0	0

LEZ = Leaf expansion zone.

C = Control samples.

I = Incremental salt treated samples.

S= Shock salt treated samples.

A further 2 markers were scored present in seedling, control and incremental but absent in shock samples. Analysis of root and leaf expansion zone tissue identified 4 markers that were scored polymorphic between root and the control, incremental and shock samples, 2 markers scored present and 2 markers absent in root. An additional 5 markers were scored absent for all leaf expansion zone and root tissue samples except incremental samples which were scored present. Analysis of polonicum samples identified 20 markers that were scored polymorphic between seedling root tissue, control, incremental and shock leaf expansion zone tissue. Analysis of polonicum leaf samples identified 9 markers that differentiated between seedling samples and control, incremental and shock salt treated samples and only 1 marker that scored present in control samples only. Root and Leaf expansion zone samples were compared, with 12 markers identified were scored polymorphic, 6 markers scored present and 6 markers scored absent in root tissue. An additional marker was scored present in control and incremental samples but absent in seedling root and shock samples.

Analysis of the duplicated experiment for Cy5 targets identified 10 markers that were scored polymorphic in durum, with 4 markers scored present and 1 marker absent in control leaf samples. A further 2 markers were scored absent only in leaf shock

samples. Analysis of leaf expansion zone tissue identified 2 markers that scored present in control samples, 1 marker that scored present and 1 marker absent in durum shock samples. Polonicum analysis identified 47 high quality markers, 12 of which were polymorphic between samples. Leaf sample analysis identified 1 marker scored present and 1 marker absent in control samples and 2 markers scored absent for shock samples. In leaf expansion samples, 3 markers were scored present for control samples, 2 markers scored present and 4 markers scored absent in incremental samples and 1 marker scored absent in shock samples.

The duplicated experiment analysis identified 74 high quality markers, 15 of which were polymorphic between treatments in durum Cy3 analysis. Of these, 7 markers were scored absent in non salt treated seedling and control leaf samples, and scored present in salt treated incremental and shock samples. A further 2 markers were scored present in seedling, control and incremental but absent in shock treated samples. Analysis of root and leaf expansion zone tissue identified 4 markers, 2 markers scored present and 2 markers scored absent in seedling root tissue, Analysis of polonicum Cy3 samples identified 16 treatment specific polymorphisms, 7 markers scoring present and 1 marker absent for seedling root tissue samples. Analysis identified 9 markers polymorphic between root and control, incremental and shock leaf expansion zone samples and a further 1 marker scored present in control and incremental samples but absent in seedling and shock samples.

The duplicated experiment analysis identified 93 high quality markers, 28 of which were polymorphic between treatments in durum Cy5 analysis. Analysis of leaf tissue samples revealed no polymorphic markers, with all variation present in leaf expansion zone tissue. Of these, 4 markers were scored absent in shock samples and present in all other samples, including seedling root tissue. A further 3 markers were scored absent and 1 marker present in seedling root tissue compared to control, incremental and shock leaf expansion zone tissue. 2 markers were scored present in root and incremental tissue and absent in control and shock tissue and 6 markers were scored present in control and incremental but absent in root and shock samples. Analysis of polonicum Cy5 targets identified 35 polymorphic markers between treatments, with 3 markers scored present in seedling leaf samples, A further 14 markers scored present in seedling and shock samples, 1 markers absent in incremental samples and 1 marker absent in shock samples. Leaf expansion zone

analysis identified 7 markers, 3 present and 4 absent, in seedling root samples, 2 markers absent in shock samples, 2 markers present in control and incremental samples and 1 marker present in seedling root and incremental salt treated samples.

4.4.4 Conclusions

From the data presented in table 4.8 and figure 4.1, several polymorphisms were identificed that could distinguish between shock and control and incremental treatments. This could imply that the DNA sequence was either methylated or unmethylated in shock samples compared to control and incremental samples, thus being differentially expressed in on the array. These sequences were not cut with the McrBC restriction enzyme, thus no adapter sequence and subsequent PCR amplification was performce, eliminating them from the genomic representation.

4.5 Methylation Polymorphism Detection

4.5.1 Aims

To further methylation detection analysis, 94 diverse durum samples were analysed using PstI and Taq$^{\alpha}$I restriction enzymes. Targets were made and labeled in 3 colours, with each sample replicated in Cy3 and in Cy5, containing the additional McrBC restriction enzyme to detect methylated DNA samples. The aim is to differentiate between methyl states within these samples.

4.5.2 Specific Methods

The 94 durum samples labeled with Cy3 (-McrBC) and Cy5 (+McrBC) were hybridised to a 17 plate microarray containing 6,528 clones. The array consisted of clones made up of hexaploid leaf samples, rearrayed polymorphic clones, Trisomic lines, A, B, and D Genome lines and AB+AG genome lines.

4.5.3 Results

Analysis of 94 wheat samples on the PstI and Taq$^{\alpha}$I array identified 836 high quality polymorphic markers in PstI and Taq$^{\alpha}$I Cy3 analysis, limited to a Q value of 75 or

greater and a call rate of 80 or greater. Reproducibility for all markers was scored 100, due to no target replicates in the experiment. Analysis of Cy5 targets, with the addition of the McrBC methylation sensitive restriction enzyme, identified 784 high quality markers. Comparison between the 836 Cy3 markers and the 784 Cy5 markers show 635 markers (81.41%) that were identified in both analyses. As these markers were scored in both analyses, the McrBC enzyme did not destroy the marker (DNA fragment) and thus the marker theoretically should not be methylated.

Analysis identified 48 markers (6.15%) that were present in Cy5 but absent in Cy3 analysis that can be theorized to be methylated, thus digested by McrBC in Cy5 target preparation and removed from PCR amplification, due to the lack of adapter-fragment-adapter formation. To check their quality, the 48 Cy5 identified markers were compared to all 6,528 markers on the array scored for Cy3 targets. Comparison shows that 34 markers were eliminated from Cy3 analysis based on their Q values scoring below the 75 threshold. The remaining 14 markers were eliminated based on their below call rate threshold scores of less than 80. These 48 markers can be further analysed to look at consensus scores between the bimodal scores for the 94 samples in both Cy3 and Cy5 analysis. Comparison shows 18 markers score a consensus of 100, meaning the 94 bimodal scores for each progeny were scored the same, either present or absent, between Cy3 and Cy5 results. These markers can not be methylated as scores are consistent between McrBC treatments. A further 22 markers scored a consensus of 90 or greater for bimodal scores in Cy3 and Cy5, reducing the number of samples in which the fragment could have been eliminated due to McrBC digestion. The remaining 8 markers were scored with a consensus of 90 or below, with 4 markers above 80, 1 marker above 70, 2 markers above 60 and one marker scored 48. These markers have a proportion of scoring discrepancies and thus a higher proportion of possible samples scoring polymorphic between McrBC treatments. Looking at the call rate between Cy3 and Cy5 samples shows that 43 of the 48 markers scored above the 80 thresholds, with 5 markers scoring between 80 and 46. These 5 markers correlate with the 5 lowest consensus markers, and thus are probably polymorphic due to poor scoring on the array and not due to McrBC treatment.

Comparisons identified 97 markers (12.44%) that were present in Cy3 analysis but absent in Cy5 analysis. To check the quality of these markers, they were compared to the full 6,528 set of polymorphic markers identified in Cy5, including both high and low quality markers. All 97 markers were scored in both Cy5 and Cy3 analysis between the 94 samples, with 22 markers rejected from Cy5 high quality analysis due to a Q value of less than 75. The remaining 75 markers were rejected from high quality Cy3 analysis due to their call rate scoring less that 80. This shows that these markers were scored during DArTsoft analysis, but were rejected from the high quality analysis comparisons. Analysis of consensus scores between Cy5 and Cy3 analyses identified 40 out of the 97 markers scoring 100 between McrBC treatments. A further 48 markers are scored 90 or greater with the remaining 9 markers scored lower than 90, with 1 markers scored in the 80's, 1 marker in the 70's, 4 markers in the 60's 1 markers in the 50's and 2 markers in the 30's. Analysis of call rates between Cy5 and Cy3 analysis show 89 markers with a call rate of above the 80 threshold, with 8 markers scored below a call rate of 80. Correlation of the 9 markers identified with a low consensus and the 8 makers found with a low call rate found 100% homology, with all poor call rate markers scoring a low correlation, thus suggesting a high probability that these markers were not scored in Cy3 analysis because of poor scoring, not because they were polymorphic between McrBC treatments. Markers that scored a low consensus but were scored a high call rate, meaning the majority of markers were scored for the 94 sample, may have suggested that these markers were McrBC specific polymorphisms.

4.5.4 Conclusions

Comparisons of Cy3 (-McrBC) and Cy5 (+McrBC) experiments from 94 durum samples hybridised to a 17 plate microarray containing 6,528 clones generated a huge amount of data. This data was analysed using DArTSoft software for quality and polymoprhisms that passed further analysed. These potential polymorphic markers were then compared to non-quality rejected markers in the opposite colour (Cy3 or Cy5) to eliminate any polymorphic data due to poor performance (such as poor spot morphology, dust that fluoresces on the slide, scratches, poor hybridization etc) looking at why it was rejected from one colour analysis and not the other. The vast majority of markers unfortunatyley were scored but failed quality analysis thus having a high probability of being polymorphic due to experimental error and not due to the effects of the McrBC enzyme on methylated DNA.

4.6 Janz and Westonia Methylation

4.6.1 Aims

Methylation diversity was explored in bread wheats using Janz and Westonia, two common Australian cultivated wheats. Samples were grown and DNA extracted from leaf tissue. Replicate targets were generated using PstI and Taq$^{\alpha}$I with and without the addition of McrBC restriction enzyme. The aim is to look for differences between restriction enzyme treatments from these samples, identifying potential molecular markers.

4.6.2 Specific Methods

Janz and Westonia wheat cultivars were germinated from seed in-house and grown in soil until plants matured. Leaf material was collected from both cultivars approximately 5 weeks after emergence from the soil, taking the entire leaf from the first or second tiller. DNA was extracted from leaf samples and targets generated using PstI and Taq$^{\alpha}$I restriction enzymes and hybridised to a wheat 10 plate array made from various wheat cultivars. McrBC restriction enzyme was added to replicate targets and polymorphisms compared.

4.6.3 Results

Janz and Westonia leaf samples were analysed using PstI and Taq$^{\alpha}$I restriction enzymes. Samples were duplicated with the addition of McrBC to destroy any methylated DNA fragments from the genomic representation. Targets were generated in quadruplicate over four experiments, with all targets in the first and second experiments labeled with Cy3 and targets in the third experiment labeled in duplicate in Cy3 and in duplicate in Cy5. Experiment 4 was doubled, so that 8 replicates of each cy3 labeled targets were hybridised. Experiment 1 identified 296 high quality polymorphic markers, with 56 markers varying between McrBC treatments. Of these, 55 markers were scored present in Janz samples, with 32 markers in the PstI and Taq$^{\alpha}$I samples and 23 markers identified with the addition of McrBC. In Westonia, only 1 marker was identified as polymorphic between McrBC treatments, being scored present in the PstI and Taq$^{\alpha}$I samples. The experiment was repeated a total of 4 times, with results shown in table 26.

Table 4.10: Markers identified in Janz and Westonia polymorphism analysis

Experiment	Number of high qualit5y polymorphisms (%)	Number of markers scored present in						
		Janz	Janz + McrBC	Janz Total	Westonia	Westonia + McrBC	Westonia Total	Both Janz and Westonia
1: 4 Cy3 replicates	296 (7.71%)	32	23	55	1	0	1	1
2: 4 Cy3 replicates	954 (24.84%)	200	278	478	238	235	473	359
3: 2 Cy3 + 2 Cy5 replicates	1,283 (33.41%)	424	500	924	401	484	885	811
4: 8 Cy3 replicates	600 (15.63%)	64	61	125	73	53	126	120

* The number of scoring differences between Janz and Westonia samples treated with and without McrBC restriction enzyme.

From table 4.10, it can be seen that there is a high degree of variation between experiments in the number of high quality markers that were identified. Similarly, there is a high degree of variation between the numbers of polymorphisms detected between McrBC treatments. To determine the reliability of experimentation, experiment 2 and 3 results, both giving the highest number of markers, were compared. Comparison of the 478 markers identified in Janz experiment 2 and 924 markers in Janz experiment 3 found 354 markers (74.06%) in both experiments, with 124 markers found only in experiment 2 and 570 markers only found in experiment 3. Comparisons of the 473 markers identified in Westonia experiment 2 and 885 markers in Westonia experiment 3 found 315 markers (66.60%) in both experiments, with 158 markers found only in experiment 2 and 570 markers only found in experiment 3. Comparison of the 55 markers identified in Janz experiment 1 with the 354 markers in experiments 2 and 3 show 14 markers (25.45%) found in all three experiments, with the majority of the remaining 41 markers removed from analysis due to poor Q value, reproducibility or call rate scores. The 1 marker identified in Westonia experiment 1 was found in both experiments 2 and 3. Experiment 4 containing the 8 replicated samples from the 4 treatments identified 125 Janz polymorphic markers between McrBC treatments, with comparison of these markers to the 354 markers found in experiment 2 and 3 show 3 markers found in both. Similarly, comparison of the 126 markers identified in Westonia experiment 4 with the 315 markers identified in both experiments 2 and 3 showed 1 markers common to all three experiments.

From these comparisons, the 354 Janz markers and 315 Westonia markers identified in both experiments 2 and 3 appear to be scored consistently to some extent in experiments 1 and 4. These 2 experiments identified considerable less polymorphic markers overall, probably due to technical issues during experimentation.

4.6.3.1 McrBC polymorphic markers

The 354 Janz and 315 Westonia McrBC specific polymorphisms can be divided into two groups, those present or absent with the additional restriction enzyme treatment. Of the 354 Janz markers, 209 markers are scored present only in McrBC treated samples and 145 markers scored absent in treated samples. In Westonia, of the 315 markers, 173 markers are scored present in McrBC treated samples and 142

markers scored absent. Markers scored absent in McrBC treated samples were also scored present in PstI and Taq$^{\alpha}$I samples, hence being polymorphic. These markers, may have been destroyed in the digestion/ligation reaction and subsequently not amplified in the PCR reaction. As McrBC only cuts methylated DNA, it can be hypothesized that these fragments were methylated and detected as methylation polymorphisms between treatments for each cultivar. Markers scored present in McrBC treated samples were also scored absent in PstI and Taq$^{\alpha}$I samples. These markers may have been created by the addition of the McrBC restriction enzyme to the digestion/ligation reaction. Changes to the abundance of fragments in both reactions changes the rate of PCR amplification, thus appearing to create a fragment, when in reality, the fragments are just amplified at a higher rate and are more abundant in the genomic representation. In any case, the proportion of markers identified in the two groups, the 354 Janz and 315 Westonia McrBC specific polymorphisms was considerably higher than was expected.

Further analysis looking at experiment 2 and 3 only, compared PstI and Taq$^{\alpha}$I samples to each other and then separately, PstI, Taq$^{\alpha}$I and McrBC samples together. From this set of analysis, markers were selected that were of high quality for Janz and Westonia that were scored either present or absent consistently over the 4 target replicates. Results are shown in table 4.11 where it can be seen that homology of markers between experiments is high with almost 100% of markers being scored between experiment 2 and 3 for each of the 2 cultivars and 2 McrBC treatments

Table 4.11: Markers found from Experiment 2 and 3 comparing targets generated with PstI and Taq$^{\alpha}$I and PstI, Taq$^{\alpha}$I and McrBC restriction enzymes.

Condition		Number of high quality markers scored present	Number of high quality markers scored absent	Number of high quality markers scored in both	Number of high quality markers scored in PstI / Taq$^{\alpha}$I (-)	Number of high quality markers scored in PstI, Taq$^{\alpha}$I + McrBC (+)
Scored Present	exp2 (-) Janz	389	389 (100%)	200	185	238
	Exp3 (-) Janz	389				
	exp2 (+) Janz	445	445 (100%)			
	Exp3 (+) Janz	445				
	exp2 no w	407	407 (99.75%)	182	217	152
	exp3 no w	408				
	exp 2 yes w	336	336 (100%)			
	exp3 yes w	336				
Scored Absent	exp2 no j	559	536 (95.88%)	191	338	124
	exp3 no j	560				
	exp 2 yes j	319	319 (100%)			
	exp3 yes j	319				
	exp no w	364	364 (100%)	199	163	241
	exp3 no w	364				
	exp 2 yes w	448	448 (100%)			
	exp3 yes w	448				

4.6.4 Conclusions

Analysis identified 287 candidate methylation polymorphisms, 145 from Janz and 142 from Westonia cultivars, out of the 3,840 clones printed on the array. These markers were scored consistently over 4 target replicates and over 2 experiments. A small proportion of these markers were also scored consistently in 2 additional experiments. These markets indicate a methylation site within these fragments that an be used for gene epression analysis or for gene silencing experiments.

4.7 Conclusion

It can be seen that the detection of polymorphic markers between wheat grown at varying temperatures and light conditions is not efficient and reliable. Reasons could be that the stress incurred wasn't sufficient or wasn't applied long enough to detect any differences between genomic representations. Sample sizes also varied, contributing to the variance in numbers. A more reliable test is needed.

Detection of polymorphisms in salt stressed wheat plants was more reliable, with candidate polymorphisms identified bewtween durum and polonicum cultivars and between leaf and growing region/root tissues. Polymorphisms between contol and seedling tissue was also observed.

Differential expression of candidate polymorphic markers was also observed between contol and incremental leaf samples and salt shock leaf samples, with markers scored reproducibility scored present and absent in replicate samples.

The vast majority of markers identified in the comparison of Cy3 (-McrBC) and Cy5 (+McrBC) experiments most probably failed quality analysis having a high probability of being polymorphic due to experimental error and not due to the effects of the McrBC enzyme on methylated DNA.

Thus the lack of hard evidence to support the idea that DArT can be used to identify polymorphic DNA sequences between environmentally variable samples, tissue samples or methylation states is limited. This new experimental application of the

technology is in its infancy and still requires further development. The traditional methods for DNA sequence variation are well used and robust, as described in chapter 5.

Chapter 5

Genetic Diversity Analysis

"Imagination is more important than knowledge"

Albert Einstein

5.0 Genetic Diversity

5.1 Introduction

Genetic diversity refers to any variation in nucleotide, gene, chromosome or whole genome of an organism. In the context used here, genetic diversity refers to differences in genomic DNA that are identified using DArT. These polymorphic DNA fragments can be detected between tissue samples as described in Chapter 3, in samples treated with varying environmental stresses, described in Chapter 4 and between wheat cultivars as described here in Chapter 5. To examine DNA sequence polymorphisms between wheat cultivars, various tetraploid durum wheats and hexaploid bread wheats were used. Janz and Kukri as well as 30 progeny from a breeding experiment were compared to identify Janz-like or Kukri-like DNA molecular markers. Sequence diversity between Janz and Westonia was examined to detect genomic sequence polymorphisms between leaf samples. A durum wheat diversity experiment between 8 cultivated and 52 wild samples analysed over 4 arrays totaling 27,648 features was also performed. Salt treated samples were also compared for cultivar polymorphisms, with samples grouped according to phenotypic data and analysed using bulk segregant analysis (BSA). BSA used 94 recombinant inbred lines (RIL) to analyse samples for polymorphic markers, with results organised into linkage groups creating a molecular map.

5.2 Janz and Kukri cultivar specific polymorphisms

5.2.1 Aims

Janz and Kukri cultivated bread wheats were crossed and DNA extracted from leaf and root tissue from 2 parents and 30 progeny. Samples were compared and data analysed for cultivar specific polymorphisms.

5.2.2 Specific Methods

A breeding experiment was performed by the South Australian Research and Development Institute (SARDI) between Janz and Kukri cultivated wheat plants. DNA from the 2 parents and 30 progeny was extracted from leaf and corresponding root

tissue. Libraries were generated in-house for leaf and root separately using PstI and Taq$^{\alpha}$I restriction enzymes from various DNA sources including a cross between Halberd and Cranbrook hexaploid wheat cultivars. Targets were produced and hybridised in 2 or 3 colour, so that there were a minimum of 2 targets for each leaf and root DNA sample. The entire experiment was then duplicated, so that 4 replicates of each target were available for analysis.

5.2.3 Results

5.2.3.1 Janz and Kukri cultivar comparison on leaf array

Janz and Kukri parental samples as well as the 30 progeny leaf and root samples were hybridized to the wheat leaf array. DArTsoft analysis identified 115 candidate polymorphic markers in the initial experiment, limited to Q values greater than 75 and a reproducibility maximum of one scoring inconsistency. Of these, 67 markers were identified as being cultivar specific, in that they differentiated between Janz and Kukri samples. Of these 67 markers, 38 markers were present in Janz cultivar samples and 29 present in Kukri samples. The replicated experiment found 42 high quality polymorphic markers, with 11 markers differentiating between cultivars. Of these, 5 markers were present in Janz and 6 markers in Kukri. Comparisons between the 67 markers identified in the initial experiment and the 11 markers in the repeated experiment show that 6 markers were identified in both experiments. These 6 cultivar specific makers are shown in table 28, with their Q and reproducibility values. It can be seen that 2 Janz-like cultivar markers and 4 Kukri-like markers were identified. These markers are termed Janz-like and Kukri-like as they may not only be specific to the cultivar analysed, and may belong to a group of cultivars that either contain or lack the marker.

Table 5.1: Janz and Kukri cultivar specific candidate polymorphisms

			Combined bimodal distribution			
Marker Name	**Q Exp 1 / Exp 2**	**Reproducibility Exp1 / Exp2**	**Janz leaf**	**Janz root**	**Kukri leaf**	**Kukri root**
800904090003_F_10	92.26488 / 76.79535	96.875 / 93.75	0	0	1	1
800904163004_K_22	81.35962 / 81.87173	98.4375 / 93.75	0	0	1	1
800904163005_P_16	83.47417 / 79.23715	98.4375 / 95.3125	0	0	1	1
800904163006_M_23	92.91141 / 79.25245	100 / 96.875	0	0	1	1
800904163005_G_8	92.77435 / 86.52944	100 / 98.4375	1	1	0	0
800904163006_D_24	93.36114 / 86.54593	100 / 95.3125	1	1	0	0

Comparisons for these markers across the 30 progeny from the breeding experiment show that for a given marker, the progeny will either have the marker present or absent, thus able to be clustered into Janz-like or Kukri-like groups for a given marker. Figure 5.1 shows the leaf and root tissue samples from the Janz and Kukri parents and 30 progeny from the breeding experiment. Figure 5.1 compares the sample ratio median values for each sample across the 6 identified markers. It can be seen that scores segregate well for the last 3 markers, 800904163006_M_23, 800904163005_G_8 and 800904163006_D_2, that is, the binary assignment of 1 or 0, or present or absent, is clearly defined. These 3 markers cluster into groups with a distinct divergence between them. The first 3 markers, 800904090003_F_10, 800904163004_K_22 and 800904163005_P_16, are not as clearly grouped into defined clusters and the divergence between them is reduced. The data in figure 5.1 shows leaf and root sample ratio median vales, and even though the binary value of 1 or 0 was identical for each tissue (table 5.1), the actual ratios were more widespread and are more informative. When the leaf ratio data is graphed (figure 5.2), the divergence between the two clusters is more defined than with both tissue samples represented, thus discrimination in leaf tissue is more defined than in root tissue for those markers.

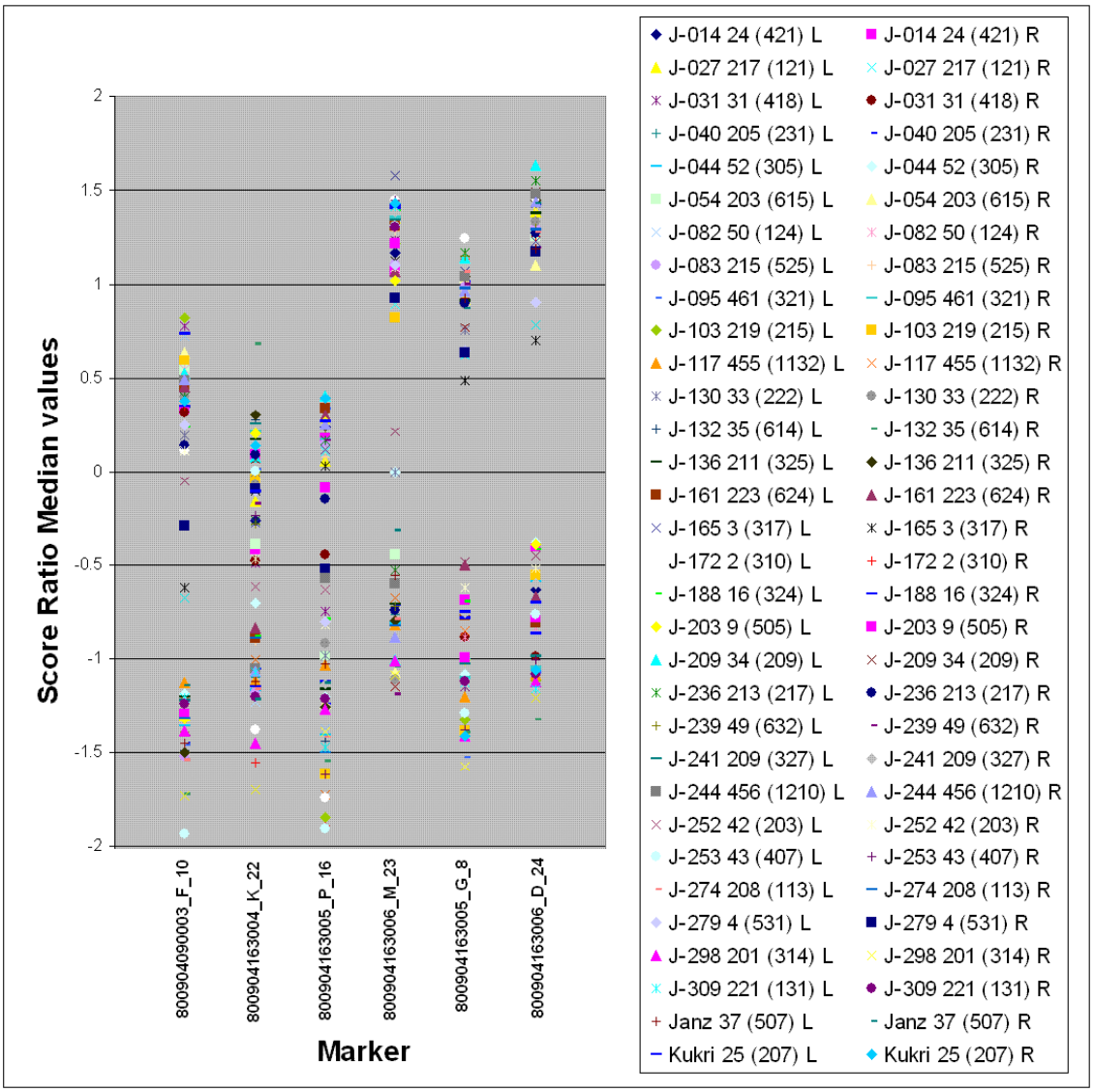


Figure 5.1: Sample ratio median values for 6 markers identified as cultivar specific.

DArTsoft polymorphism analysis showing sample ratio median values for 2 parents and 30 progeny from leaf (L) and root (R) samples for 6 markers identified as cultivar specific between Janz (J) and Kukri (K) in replicated leaf array experiments.

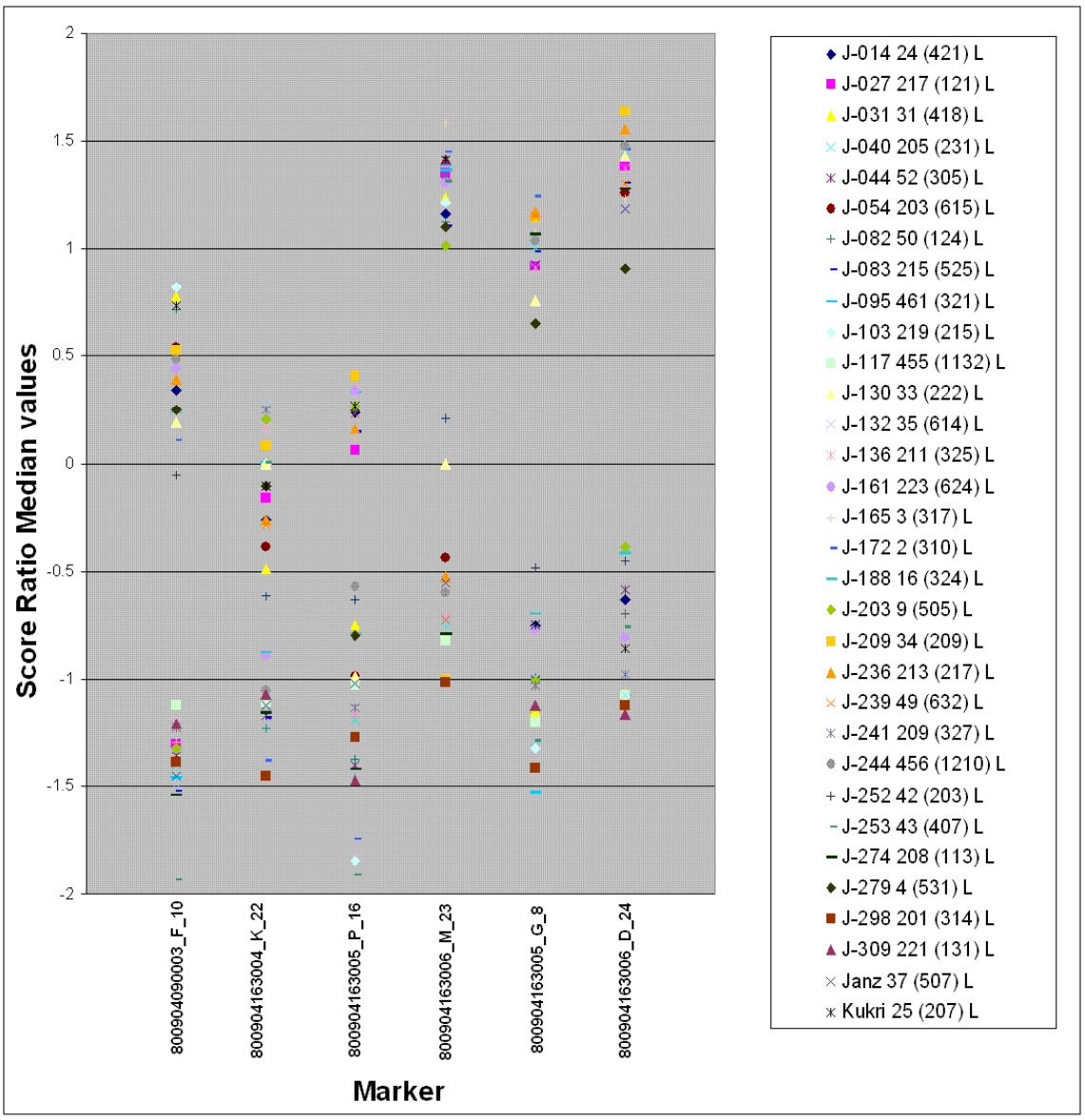
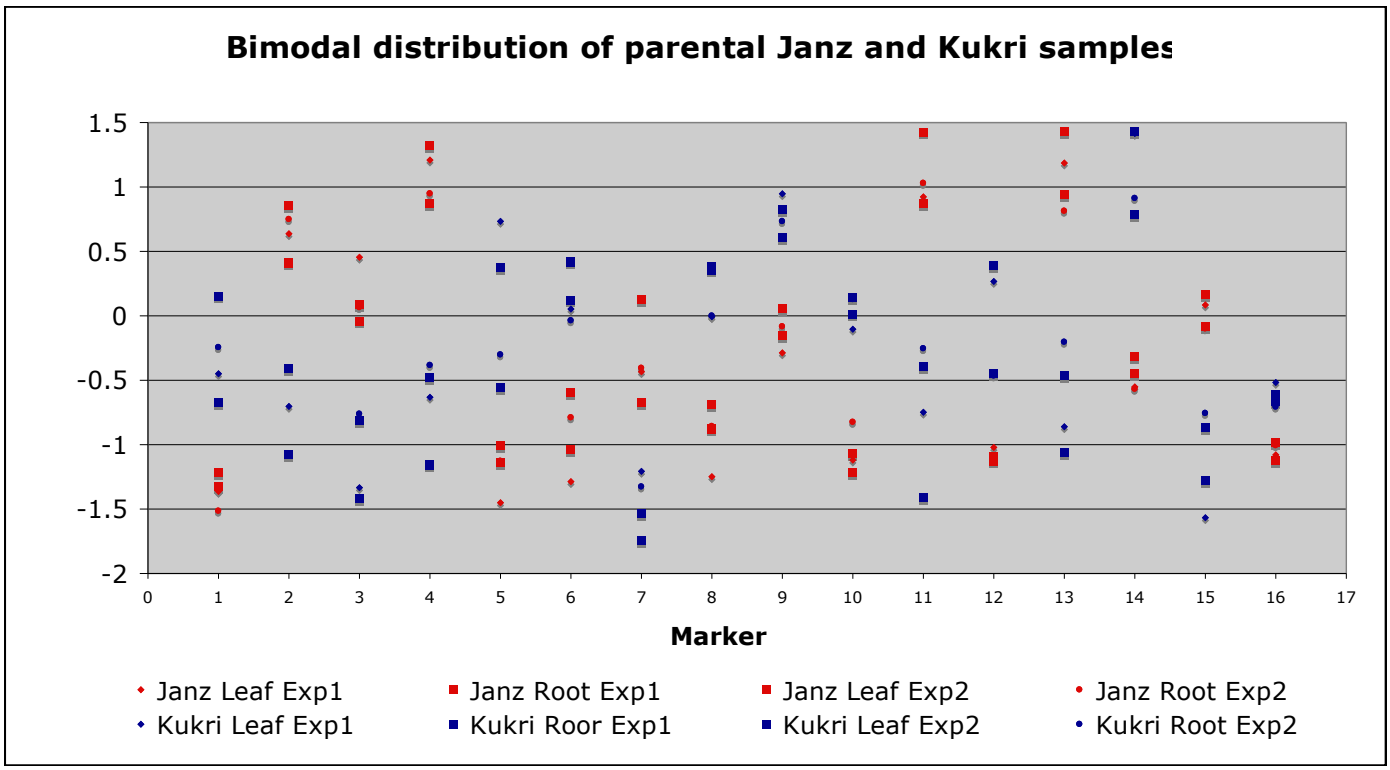


Figure 5.2: Leaf sample ratio median values for 6 markers identified as cultivar specific.

DArTsoft polymorphism analysis showing the sample ratio median values for 2 parents and 30 progeny from leaf (L) samples for 6 markers identified as cultivar specific between Janz (J) and Kukri (K) in both leaf array experiments.

5.2.3.2 Janz and Kukri parental leaf and root cultivar analysis

Comparison of the Janz and Kukri parental samples in the breeding experiment alone, without the 30 progeny, reduces the sample size within the DArTsoft polymorphism analysis, creating a higher number of good quality, Janz and Kukri cultivar specific markers. Parental analysis from leaf and root samples identified 249 high quality cultivar specific candidate polymorphisms in the initial experiment, 146 markers present in Janz samples and 103 in markers in Kukri samples. The duplicated experiment identified 407 high quality cultivar specific markers, 156 markers present in Janz samples and 251 markers in Kukri samples. Comparisons between the 249 markers from the initial experiment and 407 markers from the duplicated experiment identified 63 markers common to both experiments. Of these, 29 markers were scored present in Janz and 34 markers present Kukri, as shown in figure 5.3, where the distribution of sample ratio median scores are graphed for both experiments for a selected 16 out of the 63 markers. The 6 markers identified from the parental and progeny analysis as show in table 5.1 are all represented and scored identically in the bimodal parental analysis and are shown in figure 5.3 in red text. From figure 5.3, the distribution of markers clearly form two distinct bimodal clusters, where Janz and Kukri samples are separated as they are scored either present or absent during analysis. From the replicated results, it can be seen that the leaf and root samples for both experiments clusted closer together in both cultivars as apposed to replicate samples. The leaf and root samples from the initial experiment score closer than the initial leaf and replicated leaf samples. However, both experiments do correlate, as they can be clusted together into clearly defined Janz and Kukri cultivar clusters.



Marker	Clone Name	Marker	Clone Name
1	800904084001_A_23	**9**	800904163004_H_6
2	800904090001_C_21	**10**	800904163004_K_22
3	800904090001_H_19	**11**	800904163005_G_8
4	800904090003_C_9	**12**	800904163005_P_16
5	800904090003_F_10	**13**	800904163006_D_24
6	800904090004_A_3	**14**	800904163006_M_23
7	800904161003_I_11	**15**	800904163007_P_5
8	800904163004_E_12	**16**	800904163007_P_9

Figure 5.3: Parental leaf and root sample ratio median values for 16 markers identified as cultivar specific.

DArTsoft polymorphism analysis showing sample ratio median values for 16 polymorphic markers identified as cultivar specific between Janz and Kukri in both leaf array experiments between parental samples. Markers in red text were also identified in the parental and 30 progeny samples analysis (table 5.1).

5.2.3.3 Janz and Kukri cultivar comparison on root array

Comparison of parental and progeny Janz and Kukri leaf and roof DNA samples using the DArT wheat root array identified 74 high quality markers. Of these, 15 markers were identified that scored polymorphic between cultivars in the initial experiment, 8 markers present in Janz and 7 markers present in Kukri. The experiment was duplicated, with 225 high quality markers identified, 127 markers scored polymorphic between cultivars. Of these, 67 scored present in Janz and 58 present in Kukri. The initial experiment identified significantly less markers than expected with the majority having a low call rate, suggesting that there was a technical issue that interfered with scoring, generating markers with low call rates that were not scored consistently across replicates. All markers in the duplicated experiment scored call rates of 70 or higher. Comparison of the 15 markers identified in the initial experiment and the 127 markers identified in the duplicated experiment showed that 13 of the 15 markers were identified in both experiments.

DArTsoft polymorphism analysis of the parental Janz and Kukri leaf and root samples identified 1,002 high quality markers, with 150 markers polymorphic between cultivars. Of these, 67 markers were scored present in Janz and 83 present in Kukri. From the duplicated experiment, 1,461 high quality markers were identified, with 385 markers polymorphic between cultivars. Of these, 216 markers were scored present in Janz and 169 markers scored present in Kukri. Comparisons of the 150 cultivar specific markers identified in the initial experiment and the 385 markers in the duplicate experiment found 22 markers identified in both experiments. As the results were not as consistent as expected, the data for the root array experiments was excluded from further analysis and is not presented here.

5.2.4 Conclusions

From the data presented, it can be seen that DArT can be used to discriminate between Janz and Kukri cultivars by identifying polymorphic markers that are scored bimodally between replicated samples. Analysis identified 6 high quality reproducible markers that discriminate between Janz and Kukri wheat cultivars in both leaf and root tissue samples. Analysis of leaf and root tissue hybridised to a leaf-extracted array gave higher quality results, measured by comparisons of replicated data, compared to samples hybridised to the root-extracted array that gave variable results

between experiments. Also, it was found that Janz and Kukri leaf samples had a greater sample ratio median divergence compared to root samples, as demonstrated in figures 5.1 and 5.2. This can also be seen from the same samples in figure 5.3.

5.3 Temperature and light stress Janz and Kukri cultivar polymorphisms

5.3.1 Aims

Two bread wheat cultivars, Janz and Kukri, were analysed for polymorphisms between cultivars grown under differing environmental conditions. Samples were grown under 6 conditions and DNA samples analysed using DArT. Samples were originally compared to determine if methylation changes could be detected using DArT (chapter 4), however in this chapter, the germplasm was further analysed to detect cultivar polymorphisms over sample replicates.

5.3.2 Specific Methods

Hexaploid Janz and Kukri wheat seeds were grown at the Research School of Biological Sciences (RSBS) at the Australian National University (ANU) and at the Centre for Molecular Biology to International Agriculture (CAMBIA). All plants were grown in temperature and light controlled growth cabinets. Kukri and Janz cultivars were chosen as two diverse hexaploid wheats. Samples were grown at 10°C, 20°C and 30°C and at high (250 lumens) and low (75 lumens) light levels. Plants were germinated from seed at 20°C and transferred to environmentally controlled growth cabinets. Plants were grown to maturity at the specified temperature and lighting conditions and seed collected. The seed was then germinated on filter paper at room temperature, with seedling leaf and root samples harvested and the remainder of the seedlings planted in soil and returned to the corresponding temperature cabinets. Plants were allowed to grow for 30-40 days where mature leaf samples were taken.

5.3.3 Results

5.3.3.1 Janz and Kukri analysis

DArTsoft polymorphism analysis of 24 Janz and 13 Kukri tissue samples from material grown at differing light and temperature conditions identified 554 high quality polymorphic markers. Of these, 411 markers differentiated between cultivars, with 235 markers scored present in Janz samples and 176 markers in Kukri samples.

Analysis was further performed across temperatures, with 10 Janz samples and 7 Kukri samples analysed for cultivar polymorphisms in plants grown at 10°C. Analysis identified 562 high quality polymorphic markers, with 419 markers scored polymorphic between Janz and Kukri samples. Of these, 243 markers were scored present in Janz and 176 markers present In Kukri.

Analysis was performed for 6 Janz and 6 Kukri samples grown at 20°C, where 562 high quality markers were identified, 427 markers that discriminate between cultivars. Of these, 249 markers were scored present in Janz samples and 178 markers present in Kukri samples.

Analysis of 8 Janz and 4 Kukri samples grown at 30°C where 562 high quality markers were identified, 331 high quality, cultivar specific markers, 166 markers in Janz samples and 165 markers in Kukri samples. This data is summarised in table 5.2.

Comparison of markers identified in each of the 10°C, 20°C and 30°C temperature analyses separately show 326 cultivar specific markers identified in all three analyses, 90 markers identified in 10°C and 20°C and not 30°C, 2 markers in 20°C and 30°C but not 10°C, 7 markers in 10°C alone, 9 markers at 20°C alone and 3 markers at 30°C alone. This shows that scoring for each sample over multiple replicates was relatively accurate, given 326 markers scored in all 3 separate temperature analyses compared to the 411 markers (79.31%) identified from all samples combined. The data is summarised in figure 5.4.

Table 5.2: Number of polymorphisms between Janz and Kukri samples grown at various environmental conditions

Temperature set	Number of high quality polymorphic markers	Number of cultivar specific polymorphic markers	Number of markers scored present in:	
			Janz	Kukri
10°C, 20°C and 30°C	554	411	235	176
10°C	562	419	243	176
20°C	562	427	249	178
30°C	562	331	166	165

Looking at the 326 markers identified from all three temperature analyses further, it can be seen that markers are either scored present or absent with a clear divergence between bimodal clusters. Figure 5.5 shows 10 selected markers and their sample ratio median scores for 13 Kukri and 23 Janz samples. It can be seen that the first 13 Kukri samples are scored bimodally into two distinct clusters, and then scores are inversed for the 23 Janz samples. The closer the scores are for each sample, the more reproducible and reliable the marker is, as all samples are from the same cultivar, except for the varying environmental conditions. Experimental variations can also cause a divergence in clusters.

A better representation of Janz and Kukri cultivar specific polymorphisms is to average the scores for each marker over all replicates/samples. The average of the 13 Kukri cultivars and 23 Janz cultivars for each of the 10 markers is shown in figure 5.6. It can be seen that certain markers show a greater divergence between bimodal clusters than other markers. An example is marker 800904090002_G_4, where the average Kukri sample median ratio score is 0.777852077 and the average Janz score is -2.531042043, thus a calculated divergence of 3.30889412. Inversely, marker 800904090002_H_13 has an average Kukri score of 0.898681462 and an average Janz score of -1.055171826, a lesser divergence of 1.953853288.

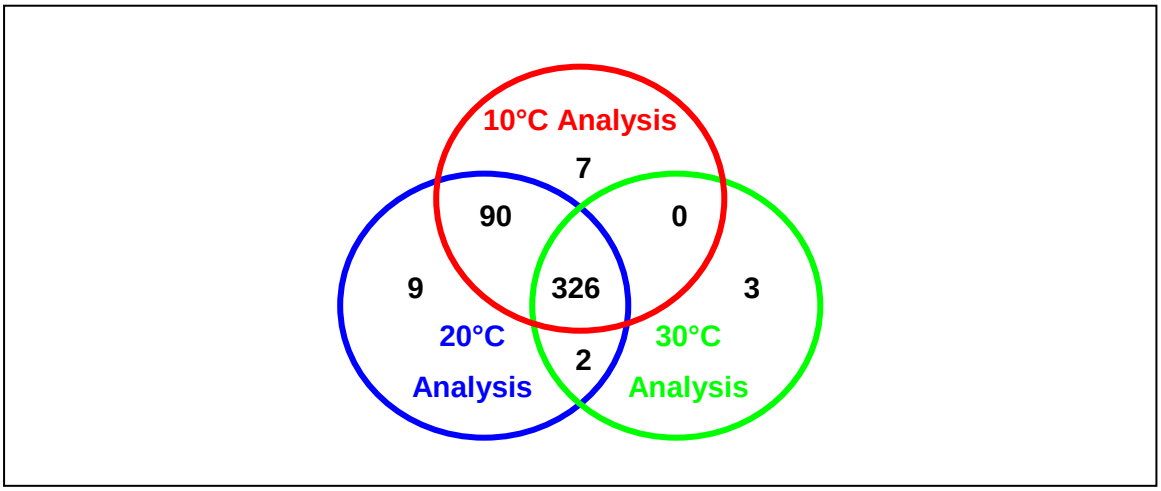


Figure 5.4: Janz and Kukri cultivar polymorphisms analysis performed at 10°C, 20°C and 30°C separately then combined, showing the overlap in markers found in each analysis.

5.3.4 Conclusions

From the analysis, it can be seen that DArT analysis of 37 Janz and Kuri samples identified over 411 cultivar specific polymorphisms within the temperature and light experiment. Separating the samples into temperature specific analyses, 326 cultivar specific polymorphisms were identified in each of the analysis (79.31%) showing that even though the sample size is reduced, the stringent parameters used in the analysis allow for relatively consistent results.

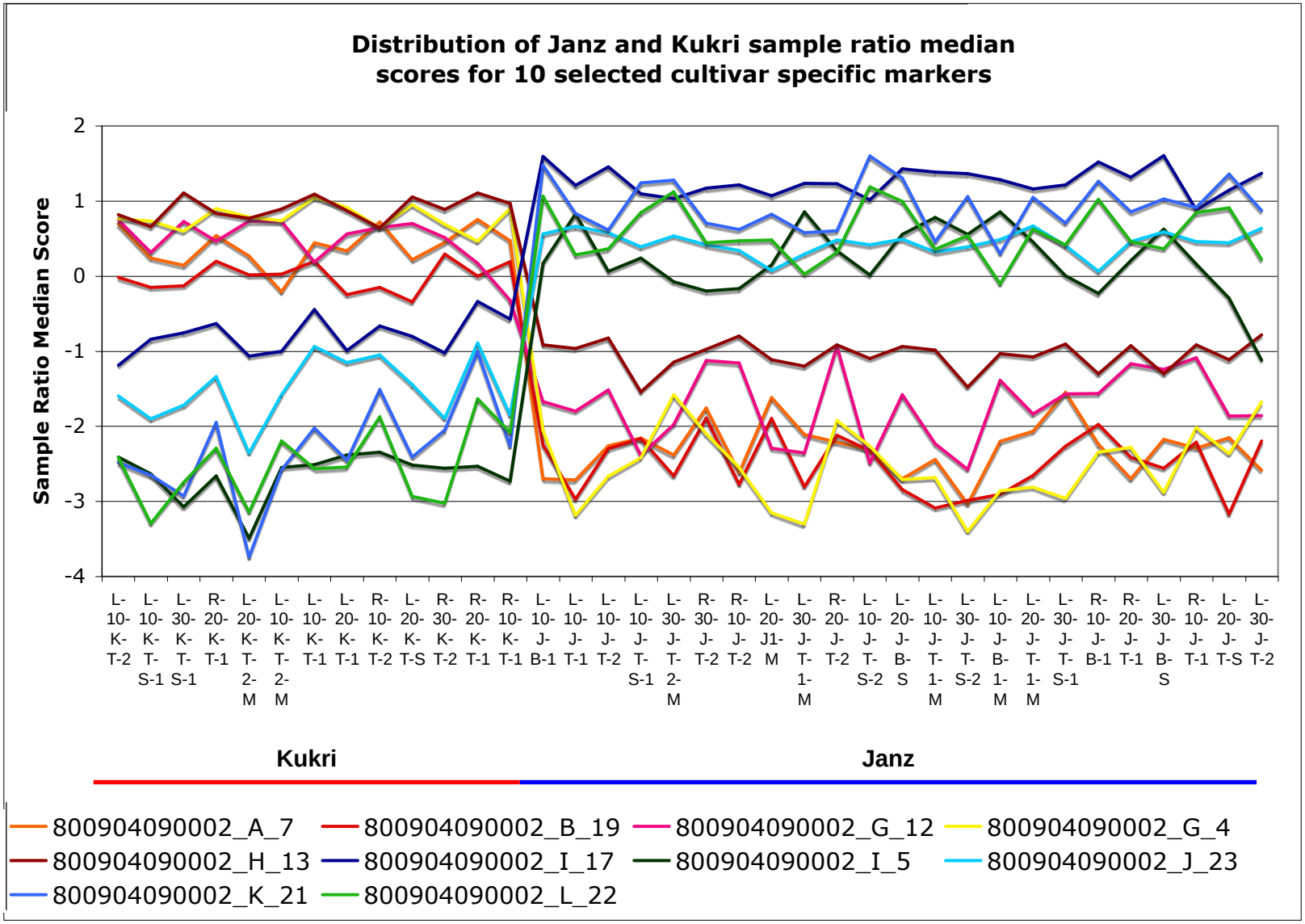


Figure 5.5: Sample ratio median scores for 10 cultivar specific polymorphic markers showing bimodal distribution between 36 Kukri and Kanz cultivars.

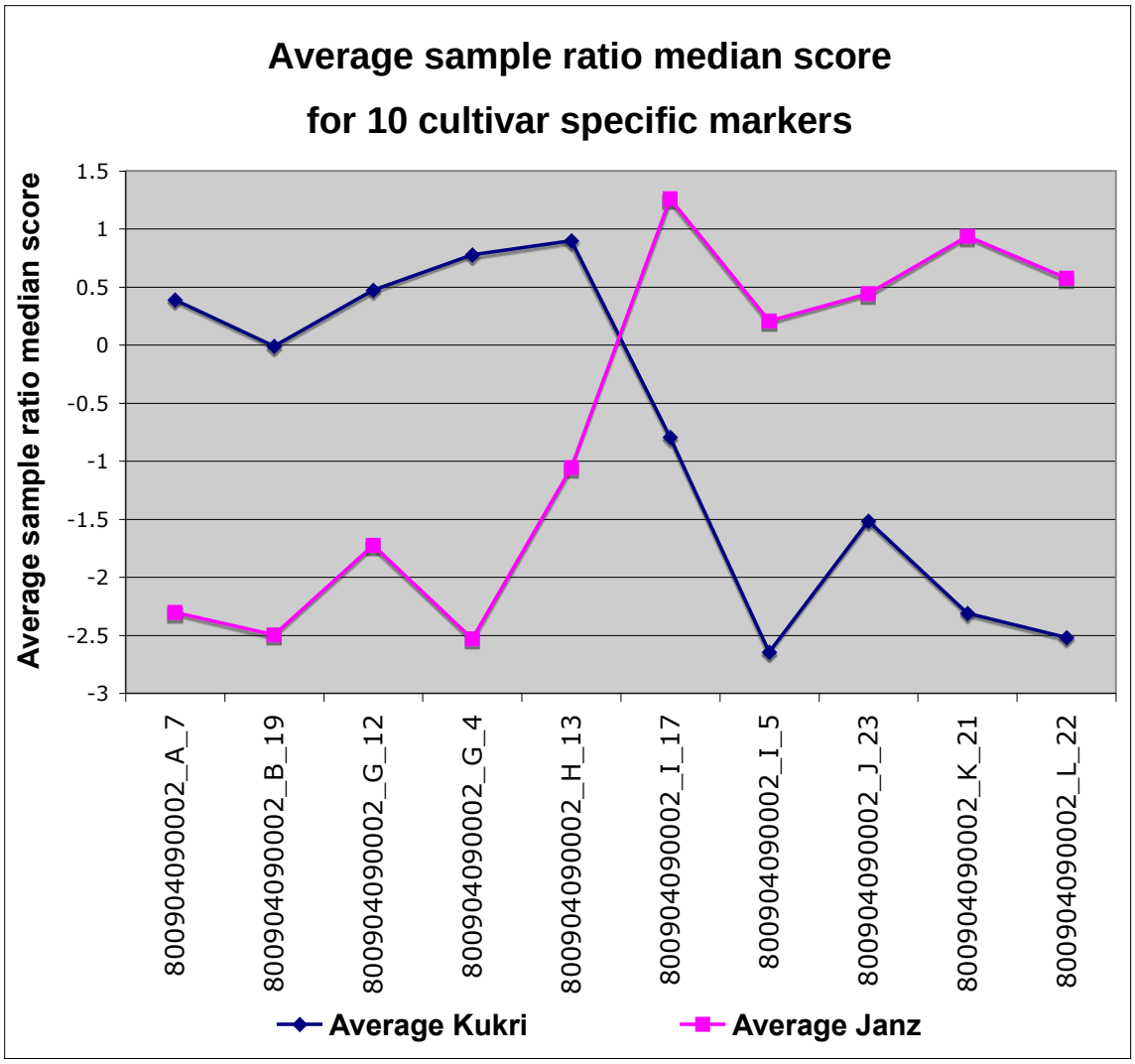


Figure 5.6: Average sample ratio median scores for 10 cultivar specific polymorphisms between Janz and Kukri samples. From the graph, it can be seen that Janz scores diverge away from Kukri scores to form two distinct polymorphic groups for these 10 markers. This pattern of marker scores can be used to distinguish between cultivar samples.

5.4 Janz and Westonia cultivar polymorphisms

5.4.1 Aims

Janz and Westonia cultivated bread wheats were grown at the Centre for Agriculture and Molecular Biology to International Agriculture (CAMBIA) and DNA extracted at DArT P/L. Samples were analysed using DArT across replicates to identify cultivar specific polymorphisms. The aim of this experiment is to identify markers across duplicate experiments that can be potentially used as DNA molecular markers to distinguish bwtween and genotype Janz and Westonia samples.

5.4.2 Specific Methods

Janz and Westonia wheat cultivars were germinated from seed in-house and grown in soil until plants matured. Leaf material was collected from both cultivars approximately 5 weeks after emergence from the soil, taking the entire leaf from the first or second tiller. DNA was extracted from leaf samples and targets generated using PstI and Taq$^{\alpha}$I restriction enzymes and hybridised to a wheat 10 plate array made from various wheat cultivars. Samples were labeled in either Cy3 or Cy5 fluorescent dyes.

5.4.3 Results

Janz and Westonia PstI and Taq$^{\alpha}$I samples were compared across replicate targets and across 4 replicated experiments totaling 20 replicates for each samples. Results from the comparison of Janz and Westonia cultivar scores for the 4 experiments are shown in table 5.3 and figure 5.7, with the number of high quality markers identified and the percentages compared to the number of features on the array. Analyses identified 106 markers scored in experiments 1, 2 and 3 as well as 111 markers scored in experiments 2 and 3. As experiment 2 and 3 showed the highest number of consistently scored markers, they will be examined further. The 222 markers in this group include the 111 markers identified between experiments 2 and 3, the 106 markers identified between experiments 1, 2 and 3 and the 5 markers identified between experiments 2, 3 and 4. Of these, 118 markers were scored present in Janz samples and 104 markers in Westonia samples (table 5.4).

Table 5.3: Janz and Westonia cultivar specific polymorphisms over 4 replicated experiments

Experiment	Number of high quality* polymorphisms (%)	Number of markers scored present in		
		Janz	Westonia	Total (%^)
1: 4 Cy3 replicates	296 (7.71%)	103	95	198 (66.90%)
2: 4 Cy3 replicates	954 (24.84%)	190	211	401 (42.03%)
3: 2 Cy3 + 2 Cy5 replicates	1,283 (33.41%)	158	155	313 (24.40%)
4: 8 Cy3 replicates	600 (15.63%)	197	196	393 (66.50%)

* Markers limited to a Q value of 75 or greater, a reproducibility maximum of 1 scoring discrepancy and a call rate of 80 or greater.

^ Percentage of cultivar specific markers out of totally number of high quality markers identified.

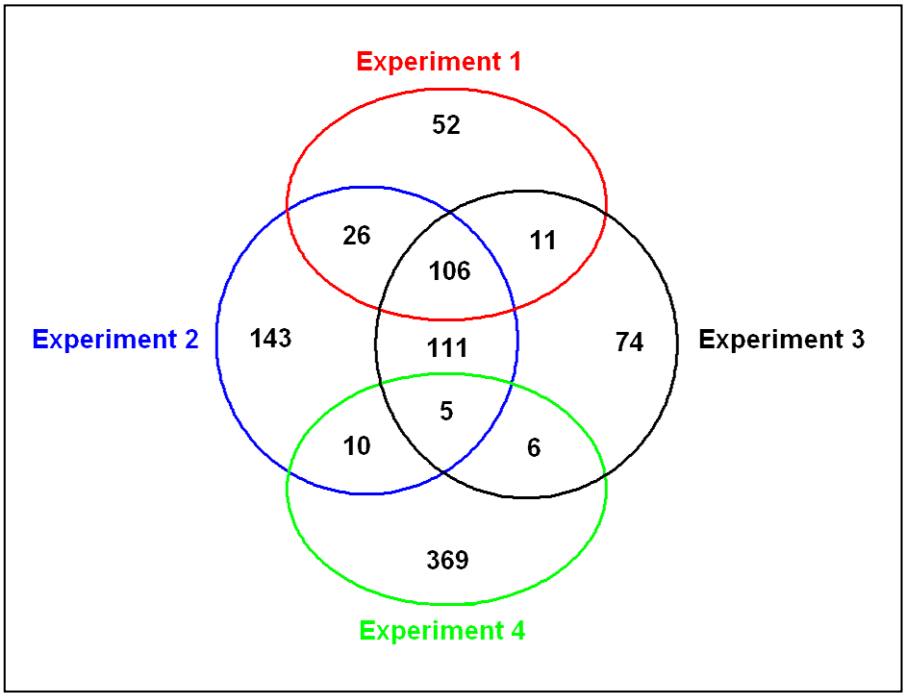


Figure 5.7: Comparison of the number of polymorphic markers identified from analysis of Janz and Westonia cultivars between 4 replicated experiments.

Table 5.4: Janz and Westonia cultivar specific markers identified from 4 replicated experiments (222 markers, 118 Janz and 104 Westonia).

118 Janz Markers			104 Westonia Markers		
800904300001_N_20	801504280004_P_6	801504280003_A_9	800904300001_E_17	801504280003_I_20	801504280002_L_15
801504280003_F_10	801504280005_A_6	801504280003_C_7	800904300001_O_21	801504280003_M_18	801504280002_L_16
801504280004_J_2	801504280005_A_7	801504280003_E_10	800904300002_B_16	801504280004_A_18	801504280003_A_5
801504280004_O_4	801504280005_B_1	801504280003_K_13	800904300002_M_24	801504280004_C_7	801504280003_C_6
800904300001_B_19	801504280005_B_21	801504280003_M_12	800904300002_N_19	801504280004_E_10	801504280003_D_10
800904300001_B_20	801504280005_F_16	801504280003_N_8	800904300002_O_20	801504280004_E_7	801504280003_D_18
800904300001_D_3	801504280005_H_17	801504280003_P_16	800904300002_P_23	801504280004_G_3	801504280003_F_6
800904300001_E_6	801504280005_P_4	801504280003_P_9	800904300003_D_21	801504280004_H_12	801504280003_J_11
800904300001_H_8	801504280006_D_24	801504280004_A_7	800904300003_D_5	801504280004_I_13	801504280004_B_18
800904300001_O_6	800904300001_D_13	801504280004_B_4	800904300003_I_15	801504280004_I_14	801504280004_B_3
800904300002_G_11	800904300001_D_2	801504280004_E_12	800904300004_B_17	801504280004_K_10	801504280004_F_17
800904300003_E_7	800904300001_E_23	801504280004_E_16	800904300004_E_16	801504280004_L_23	801504280004_G_1
800904300003_N_15	800904300001_E_4	801504280004_E_19	800904300004_O_23	801504280004_N_5	801504280004_H_6
800904300004_C_11	800904300001_I_10	801504280004_E_20	801504280001_D_14	801504280005_A_9	801504280004_I_7
800904300004_E_15	800904300001_K_3	801504280004_E_21	801504280001_D_17	801504280005_C_5	801504280004_J_5
800904300004_J_17	800904300001_K_9	801504280004_G_17	801504280001_E_10	801504280005_G_4	801504280004_M_5
800904300004_L_20	800904300001_L_23	801504280004_I_11	801504280001_E_15	801504280005_I_7	801504280004_O_5
801504280001_B_13	800904300001_M_10	801504280004_I_21	801504280001_F_12	801504280005_I_8	801504280005_B_4
801504280001_B_7	800904300001_N_16	801504280004_J_14	801504280001_G_16	801504280005_K_9	801504280005_B_8
801504280001_F_4	800904300002_D_19	801504280004_K_22	801504280001_H_2	801504280005_M_6	801504280005_F_1
801504280001_G_13	800904300002_G_6	801504280004_N_9	801504280001_I_10	801504280006_H_18	801504280005_F_5
801504280001_I_13	800904300002_N_14	801504280004_P_21	801504280001_I_16	800904300001_C_8	801504280005_H_6
801504280001_I_5	800904300002_O_8	801504280005_C_7	801504280001_I_17	800904300001_D_18	801504280006_J_4
801504280001_M_6	800904300003_H_17	801504280005_D_8	801504280001_I_18	800904300001_I_15	801504280006_O_19
801504280001_M_7	800904300004_D_16	801504280005_E_1	801504280001_J_10	800904300001_J_11	801504280001_C_17
801504280001_O_4	800904300004_F_5	801504280005_E_3	801504280001_K_5	800904300001_N_15	801504280001_I_11
801504280002_N_20	800904300004_J_13	801504280005_E_5	801504280001_K_6	800904300002_B_7	801504280001_J_12
801504280003_I_17	800904300004_P_17	801504280005_H_5	801504280001_L_5	800904300002_K_10	801504280001_J_17
801504280003_K_15	801504280001_C_18	801504280005_H_9	801504280002_A_18	800904300002_N_11	801504280001_J_4
801504280003_L_12	801504280001_G_8	801504280005_I_4	801504280002_C_23	800904300003_N_13	801504280003_A_21
801504280003_L_23	801504280001_I_23	801504280005_J_2	801504280002_F_18	800904300003_O_5	801504280003_A_4
801504280003_M_7	801504280001_I_8	801504280005_J_21	801504280002_F_24	800904300003_P_18	801504280003_B_8
801504280004_C_16	801504280001_O_6	801504280006_J_8	801504280002_I_21	800904300004_F_18	801504280003_G_16
801504280004_D_17	801504280002_E_17	801504280006_M_9	801504280002_K_24	800904300004_N_12	801504280003_H_6
801504280004_E_3	801504280002_H_18	801504280006_N_2	801504280002_O_20	801504280001_B_8	
801504280004_F_4	801504280002_I_18	801504280006_N_6			
801504280004_J_16	801504280002_L_12	801504280006_O_16			
801504280004_K_18	801504280002_M_16	801504280006_P_6			
801504280004_K_6	801504280002_O_19				
801504280004_K_7	801504280003_A_22				

The 222 Janz and Westonia cultivar specific markers were further examined by limiting markers to a Q score of 85 or above and a reproducibility of 100 for all replicates within an experiment. The markers were also limited to a reproducibility of 100 in at least 2 of experiments 1, 2 and 3. 12 selected markers from this group that scored positive for Janz are shown in figure 5.8 and 12 markers positive for Westonia in figure 5.9. The 12 markers for Janz and Westonia are shown for their sample ratio median scores from experiment 2 averaged over 4 replicates within the experiment. Sample (average) scores for all 4 replicates for Janz Cy3, Janz Cy5, Westonia Cy3 and Westonia Cy5 samples are shown. Samples labeled with Cy3 were produced using PstI and Taq$^{\alpha}$I restriction enzymes during the complexity reduction step. Samples labeled with Cy5 were produced using PstI, Taq$^{\alpha}$I and the addition of McrBC. The methylation sensitive McrBC can discriminate between methylated DNA fragments as described in Chapter 4. From figure 5.8 and 5.9, it can be seen that there is some variation in the sample ratio median scores for each of the cultivars, as they were produced with or without McrBC. However those scores show less of a divergence within their corresponding cultivar scores than from the apposing cultivar for each marker.

5.4.4 Conclusion

From the data collected, it can be seen that although the same experiment was replicated 4 times, with 4-8 sample replicates in each experiment, results vary from identifying 296 high quality polymorphisms in experiment 1 to 1,283 high quality polymorphisms in experiment 3. All experiments used Cy3 dye for target analysis except experiment 3, which used Cy3 and Cy5 dyes. This could attribute to the variation in marker identification frequencies. However, all of the 222 markers found in experiment 2 were identified in experiment 3, showing that there is some consistency in the results.

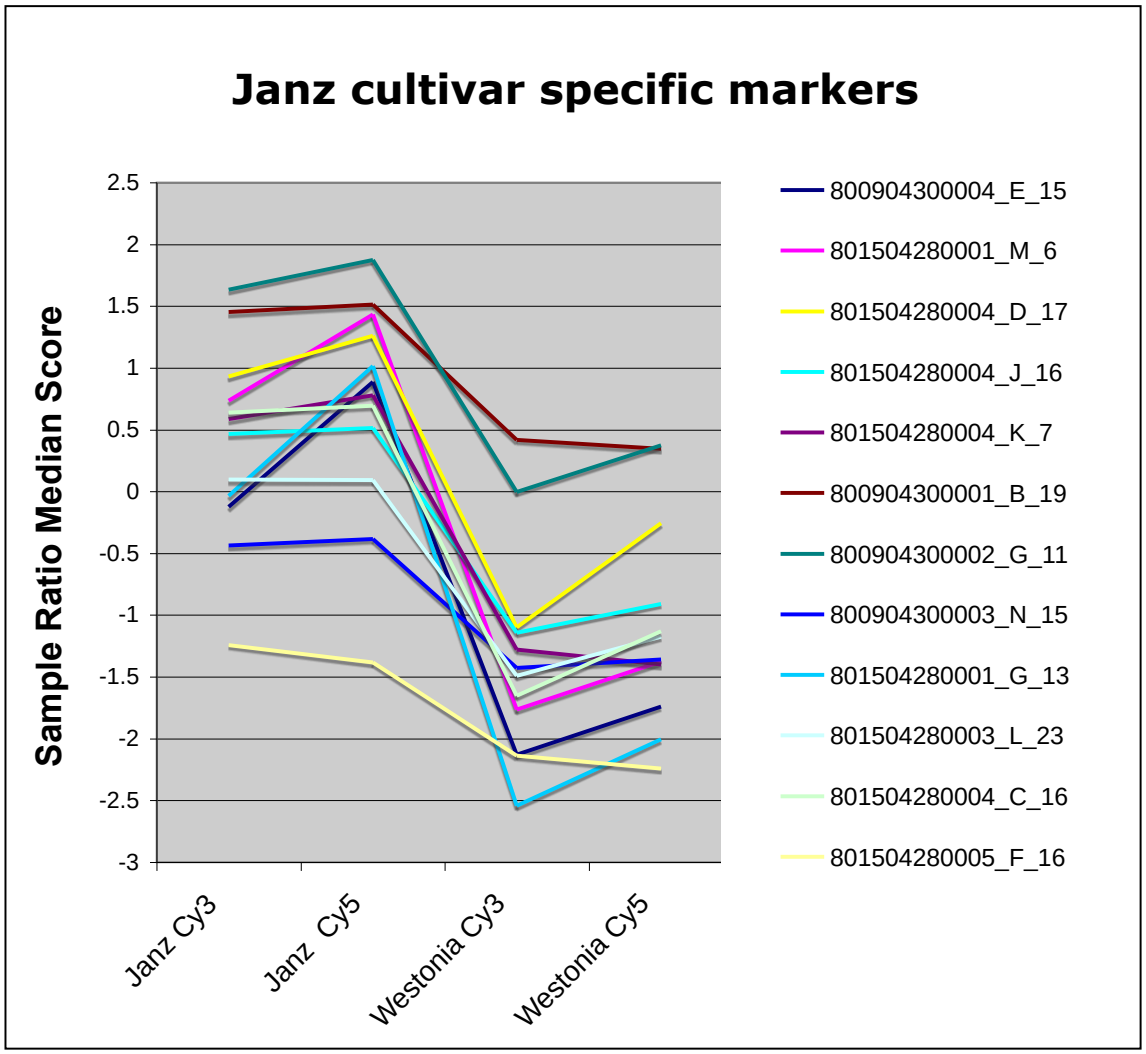


Figure 5.8: Janz specific markers showing 12 markers scored present for Janz and absent for Westonia. There is a clear seperation of marker scores for each cultivar with some markers more divergent than others.

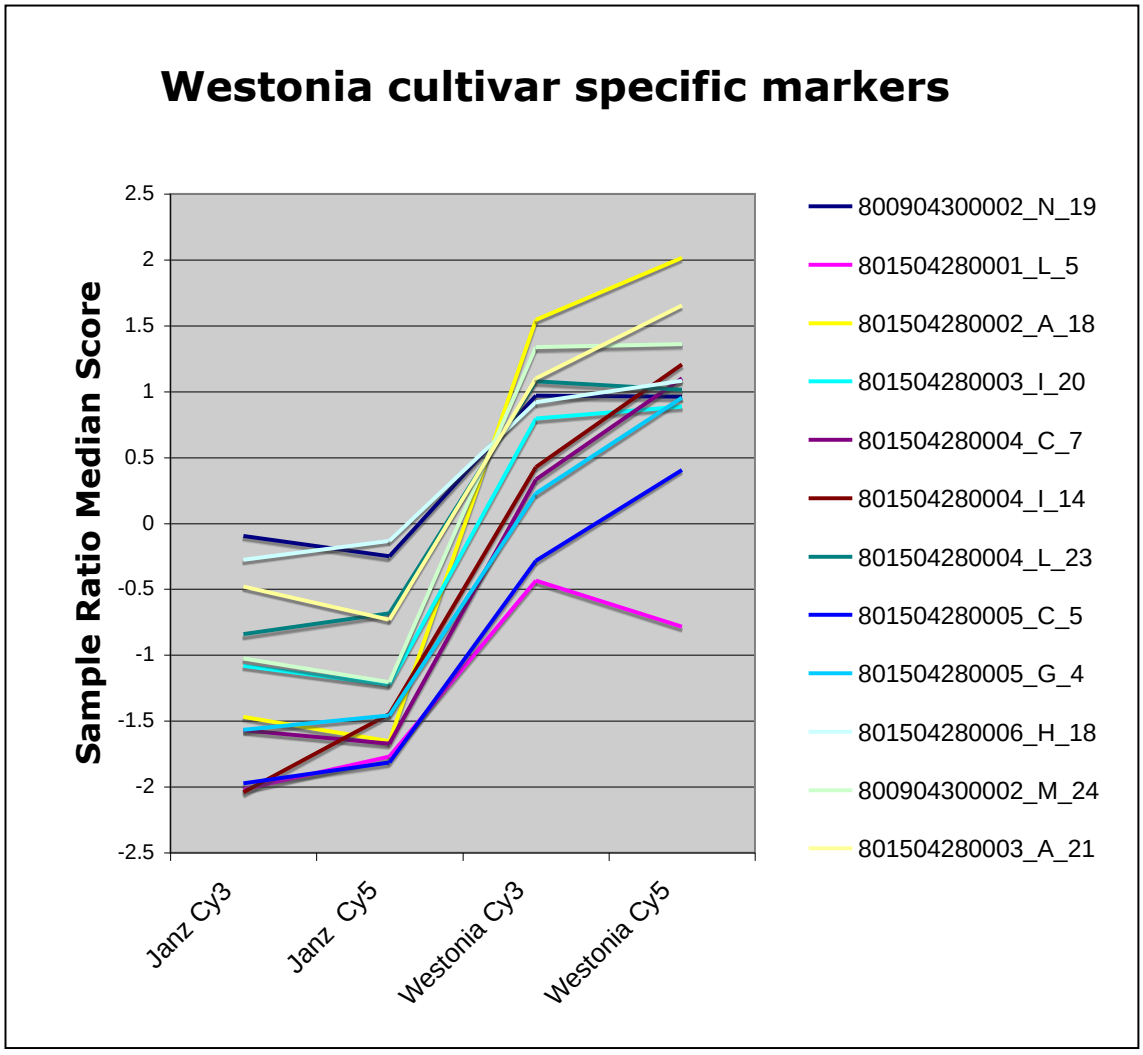


Figure 5.9: Westonia specific markers showing 12 markers scored present for Westonia and absent for Janz. There is a clear seperation of marker scores for each cultivar with some markers more divergent than others.

5.5 Wild and Cultivated Durum Discovery

5.5.1 Aims

To further genetic diversity studies, wild and cultivated durum samples were analysed. Cultivated samples included 8 genotypes used for the salt tolerance experiments, with DNA from root and duplicated leaf samples used. Wild leaf samples from 55 durum cultivars were added and analysed over 4 different microarrays.

5.5.2 Specific Methods

Leaf and root tissue samples were harvested from 8 cultivated durum cultivars grown from seedlings in the laboratory provided from Dr Rana Munns and Dr Richard James from CSIRO Plant Industries, Canberra (table 5.5). The 8 cultivated wheat samples can further be divided into 2 groups, denoted as the 'ssp. durum samples' that include durum samples 149, 39, Wollaroi and Tamaroi. The second group is the 'cultivated durum' samples that include carthlicum, polonicum, turanicum and Turgidum. An additional 55 wild durum wheat cultivars (table 5.5) were included that were provided from several DArT collaborative partners. DNA was extracted from leaf and root material using the standard DNA extraction protocol. Microarrays were printed from libraries created from an in-house DArT diverse wheat collection. A total of 72 x 384-well plates, equaling 27,648 clones were printed over 4 separate microarrays. Clones were printed in duplicate onto glass slides identified as library 1 to library 4.

PstI and Taq$^{\alpha}$I targets were created from the 55 wild (some in duplicate) and 8 leaf (in duplicate) and root cultivated seedling samples and hybridised to all 4 arrays and analysed. The 94 targets and 2 negative controls were hybridised so that each target was hybridised to each of the 4 different arrays. The 4 libraries were hybridised with targets labeled with Cy3 fluorescent dye and scanned then analysed using DArTsoft polymorphism analysis software.

Table 5.5: Wild and cultivated durum samples

55 wild durum samples					
ACMORSE	ACPATHFINDER	ACTAR84	AGHRASS1	ASTRODUR	AWL2/BIT
AZEGHAR2	BELIKH2	BEN	CAPEITI8	CHAM1	CLAUDIO
COLOSSEO	CRESO	DON	PEDRO	DUILIO	GIDARA2
GRAZIA	HAURAN1	IRIDE	JENNAH	KHETIFA-TAMGURT	KORIFLA
KYLE	LAHN	LANGDON	LEVANTE	LLOYD	LOUKOS1
MAIER	MERIDIANO	MESSAPIA	MEXICAL75	NEFER	NEODUR
OFANTO	OMRAB15	OMRUF2	ORJAUNE	OUASSEL1	PLATA16
QUADALETE	RASCON/2TARRO	REVA	SARAGOLLA	SEBAH	SENATORE
CAPPELLI	SIMETO	SVEVO	TRINAKRIA	USA-ACCESSION	VALFORTE
ZEINA1					

8 cultivated durum samples	
T. turgidum L. ssp. carthlicum	*T. turgidum L. ssp. durum 149*
T. turgidum L. ssp. polonicum	*T. turgidum L. ssp. durum 39*
T. turgidum L. ssp. turanicum	*T. turgidum L. ssp. durum Tamaroi*
T. turgidum L. ssp. Turgidum	*T. turgidum L. ssp. durum Wollaroi*

5.5.3 Results

5.5.3.1 Durum Discovery Library 1

The DArT durum discovery library 1 consists of 19 plates totaling 7,296 clones generated from durum samples containing the AB and AG genomes printed in duplicate. The array was hybridised using 55 wild leaf samples and 8 leaf and corresponding root samples from cultivated wheats. Samples were analysed using DArTsoft and polymorphisms identified.

5.5.3.1.1 Analysis of 8 cultivated wheat samples

DArTsoft analysis was used to identify polymorphisms between 8 cultivated durum wheat samples from leaf and root tissue. The 8 samples are shown in table 5.5, where analysis identified 700 high quality markers with a Q value of 75 or greater, reproducibility maximum of 1 scoring discrepancy and limited to a call rate of 80 or

greater. Of these, 99 markers were identified as differentiating between wheat cultivars, having the same scores for leaf and root tissue. For each of the 99 cultivar specific polymorphic markers, the 8 cultivars can be grouped according to the presence or absence of that marker. Comparisons between *ssp. polonicum* (55) and the 7 durum cultivars identified 10 markers were scored absent and 1 marker present only in polonicum. Analysis identified 9 *ssp*. *turanicum* (28) specific markers, 17 *ssp*. *turgidum* (62) specific markers, 9 *ssp. carthlicum* (14) specific markers, 29 *ssp. durum* (149) specific markers, 10 *ssp. durum* 139 (39) specific markers, 6 Tamaroi specific markers and 8 Wollaroi specific markers. The 68 markers identified as cultivar specific and 1 durum specific marker are shown in table 5.6 with an example of each shown in table 5.7.

5.5.3.1.2 Analysis of wild and cultivated wheat samples

DArTsoft polymorphism analysis of 55 wild and 8 cultivated durum samples identified 366 high quality polymorphisms limited to a call rate of 80 or greater. These 366 markers can be clustered into groups based on their bimodal score. Markers 801705322001_H_5, 801705322004_E_21 and 801705322003_N_14 score absent (0) for all wild wheat samples with the exception of AGHRASS1, and score present (1) for all cultivated samples, with the exception of durum sample 139 which scored absent in leaf tissue but present in root tissue. Markers 801705321002_O_11 and 801705322001_P_19 were scored present in all samples except polonicum, where the marker was scored absent. These polonicum specific markers are shown Figure 5.10 where the sample ratio median scores were graphed for all samples. It can be seen that there is a clear distinction between the polonicum scores and all other wild and cultivated durum samples for these two markers in both leaf and root tissue. These two markers were also identified in the analysis of the 8 cultivated durum samples

Table 5.6: Cultivar specific durum wheat markers – Library 1

Designation	Species	Cultivar Specific markers		
14	*T. turgidum L. ssp.* carthlicum	801705321008_D_23	801705321008_H_22	801705321006_O_24
		801705321001_E_8	801705321004_H_1	802906152014_H_24
		802906152017_H_11	801705321005_D_9	801705321001_B_9
149	*T. turgidum L. ssp.* durum 149	802906152017_D_12	801705321003_H_13	801705321006_M_11
		801705321008_N_19	802906152016_N_2	801705321002_G_5
		801705321003_M_2	802906152017_N_12	802906152017_J_4
		801705322008_N_22	801705321001_A_24	801705321008_F_6
		801705321008_F_23	801705321005_O_23	801705321007_C_3
		801705322005_M_22	801705321006_N_23	801705321006_C_12
		801705321002_O_14	801705321001_K_14	801705322001_C_23
		801705321005_P_19	801705321007_F_24	801705321006_P_6
		801705322001_G_12	801705322006_E_6	801705322006_K_13
		801705322005_N_23	802906152014_N_1	
55	*T. turgidum L. ssp.* polonicum	801705321002_O_11	801705321001_C_16	801705322002_F_8
		801705322001_P_19	801705321005_F_17	802906152014_I_5
		801705321002_K_23	801705322005_O_11	801705321006_H_9
		801705321002_A_14	802906152017_L_8	
62	*T. turgidum L. ssp.* turgidum	801705322004_L_19	802906152016_P_8	801705321001_F_9
		801705322005_F_19	801705322002_F_1	801705321001_G_22
		801705321007_G_20	801705321004_G_8	801705322004_J_6
		801705321007_P_10	802906152017_O_12	801705322006_C_23
		801705321001_J_8	801705321005_G_8	801705321005_P_10
		801705322004_J_1	801705321002_M_14	
28	*T. turgidum L. ssp.* turanicum	802906152017_E_3	801705322007_L_3	801705322007_H_20
		801705322004_K_5	802906152014_P_8	801705321003_G_5
		801705321006_B_5	802906152014_J_4	802906152016_J_19
T	*T. turgidum L. ssp. Durum* Tamaroi	801705322006_H_22	801705322007_O_16	801705322005_C_3
		801705321007_K_18	801705321003_B_18	801705321007_H_24
W	*T. turgidum L. ssp. Durum* Wolloroi	801705321004_F_16	801705321006_J_5	801705322008_F_7
		802906152016_H_18	801705322004_D_19	801705322003_A_5
		801705321007_O_19	801705321006_G_17	
39	*T. turgidum L. ssp. durum 139*	801705322003_I_3	801705321002_J_16	801705322005_O_8
		801705322005_G_11	801705322006_H_2	802906152014_O_16
		801705321004_F_5	801705322008_A_7	801705321007_D_9
		801705321008_A_8		

Table 5.7: Selected polymorphic cultivated durum wheat samples – Library 1

Marker	Ssp. polonicum (55)		Ssp. turanicum (28)		Ssp. turgidum (62)		Ssp. carthlicum (14)		Ssp. durum (149)		Tamaroi (T)		Wollaroi (W)		Ssp. durum (39)	
Tissue	Root	leaf	root	leaves	root	leaves	Root	leaves	root	leaves	root	leaves	root	leaves	root	leaves
801705321008_F_14	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
801705321002_O_11	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
801705321006_H_9	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
802906152017_E_3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1
801705322004_L_19	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
802906152017_H_11	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1
801705321005_D_9	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
801705321007_F_24	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1
801705322005_C_3	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1
801705322004_D_19	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
801705322008_A_7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0

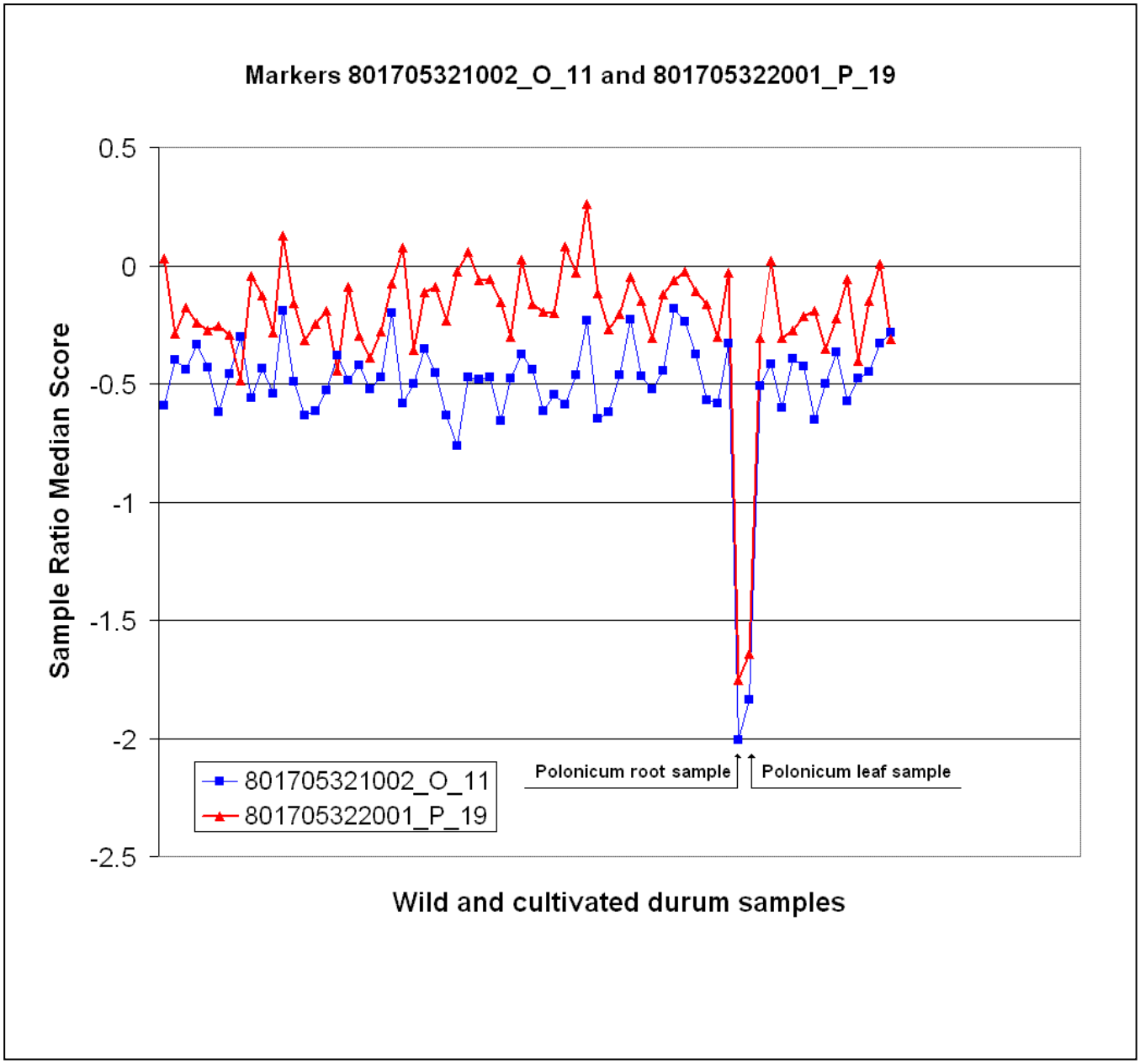


Figure 5.10: Polonicum specific markers 801705321002_O_11 and 801705322001_P_19 as identified by analysis of 52 wild and 8 cultivated durum samples. There is a clear divergence of the polonicum marker in both replicated leaf and root tissue to the other samples analysed, allowing for this marker to be used as a positive identification marker for Polonicum.

5.5.3.1.3 Analysis of carthlicum and Wollaroi cultivars

5.5.3.1.3.1 Analysis using 8 cultivated durum samples

DArTsoft polymorphism analysis of leaf and root tissue from the 8 cultivated wheat samples identified 200 high quality markers. Of these, 39 markers were identified as polymorphic between carthlicum and Wollaroi, with 18 markers scored present in carthlicum and 21 markers in Wollaroi.

5.5.3.1.3.2 Analysis using all 55 wild and cultivated durum samples

DArTsoft polymorphism analysis of leaf and root tissue from the 55 wild and cultivated wheat samples identified 366 high quality markers. Of these, 74 markers were identified as polymorphic between carthlicum and Wollaroi in both leaf and root tissue samples. Analysis identified 50 markers that were scored present in carthlicum and 24 markers in Wollaroi. A further 30 tissue specific markers were identified as being polymorphic between leaf and root samples for either cultivar.

5.5.3.1.3.3 Comparison of markers identified in both analyses

A comparison was made between markers that were identified from the experiment that included all wild and cultivated durum samples and markers found in the analysis of the 8 cultivated samples alone. Of the 39 markers identified in the cultivated sample analysis and the 74 markers in the wild and cultivated analysis, 11 markers were identified in both analyses, with 3 markers in carthlicum and 8 markers scored present in Wollaroi.

5.5.3.2 Durum Discovery Library 2

The DArT durum discovery library 2 consists of 12 plates totaling 6,528 clones generated from durum samples containing the A and B plus A and D genomes as well as the A, B and D genomes. Each clone was printed in duplicate on the array. The array was hybridised using 55 wild leaf samples and 8 leaf and corresponding root samples from cultivated wheats. Samples were analysed using DArTsoft and polymorphisms identified.

5.5.3.2.1 Analysis of 8 cultivated wheat samples

DArTsoft analysis was used to identify polymorphisms between 8 cultivated durum wheat samples from leaf and root tissue. Analysis identified 1,113 high quality markers, limited to a call rate of 80 or greater. Of these, 177 markers were identified as differentiating between wheat cultivars, having the same scores across replicates and leaf and root tissues. For each of the 177 cultivar specific polymorphic markers, the 8 cultivars can be grouped according to the presence or absence of that marker. From the 177 cultivar specific markers, analysis scored 20 carthlicum, 16 polonicum, 10 turanicum, 33 turgidum, 18 durum 139, 54 durum 149, 14 Tamaroi and 12 Wollaroi specific markers. Table 5.8 shows the cultivar specific markers that were identified from durum library 2.

Table 5.8: Cultivar specific durum wheat markers – Library 2

Designation	Species	Cultivar Specific markers		
14	*T. turgidum L. ssp.* carthlicum	800906058008_O_15	800906058008_B_23	800906058002_C_17
		800906058014_H_8	800906058008_M_15	800906058016_K_22
		800906058003_O_11	800906058011_K_12	800906058001_N_18
		800906058004_E_22	800906058003_E_13	800906058006_H_1
		800906058014_M_4	800906058014_C_15	802906247001_G_14
		800906058016_F_14	800906058014_J_18	800906058001_B_10
		800906058008_K_5	800906058007_H_11	
149	*T. turgidum L. ssp.* durum 149	800906058004_C_21	800906058001_D_9	800906058003_L_15
		802906152002_J_12	802906247001_F_12	800906058003_P_9
		800906058014_J_2	800906058007_G_11	800906058015_G_2
		800906058010_N_19	800906058001_M_1	800906058005_H_24
		800906058004_H_7	800906058008_B_21	800906058001_G_15
		800906058007_C_5	802906247001_H_22	802906152002_K_9
		800906058005_F_16	800906058004_J_12	800906058004_D_6
		800906058013_N_2	800906058008_E_4	800906058001_M_10
		802906152002_C_19	800906058016_J_12	800906058001_P_13
		800906058011_J_16	800906058007_E_13	800906058014_J_4
		802906247001_P_13	800906058002_F_3	800906058007_O_14
		800906058013_M_7	800906058001_L_4	800906058004_B_3
		800906058004_L_13	800906058007_F_15	800906058007_P_13
		800906058007_C_11	800906058001_M_4	800906058012_N_2
		800906058004_M_15	800906058013_C_2	800906058012_M_2
		800906058015_E_3	800906058005_F_1	800906058005_E_23
		800906058014_L_24	800906058012_L_3	800906058003_L_21
		800906058005_K_18	800906058005_O_20	800906058016_G_24

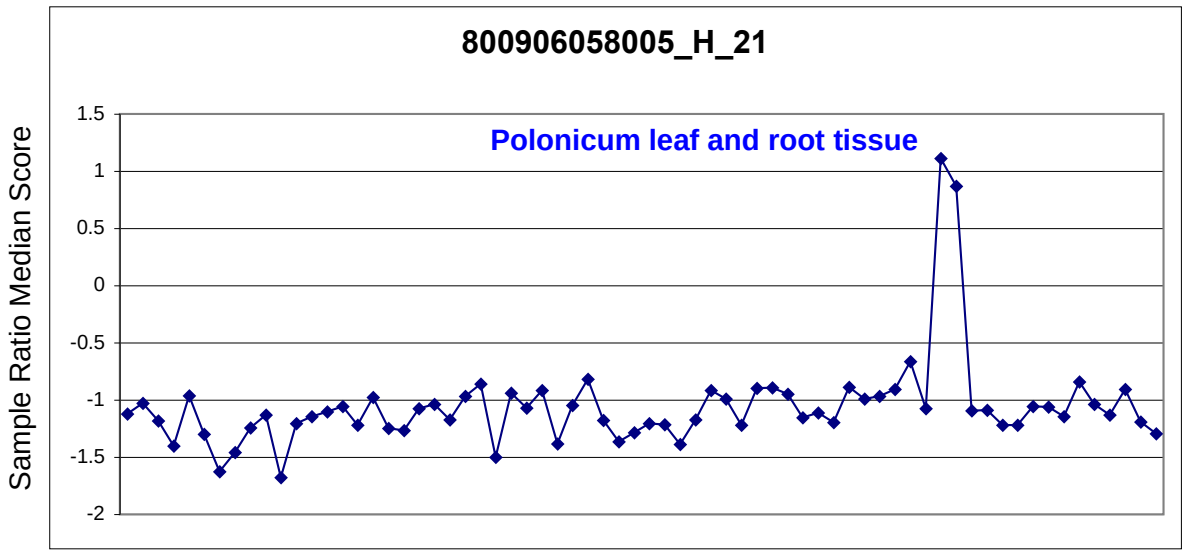
55	*T. turgidum L. ssp.* polonicum	800906058015_M_16	802906247001_F_14	802906247001_I_4
		800906058014_H_4	800906058002_I_19	800906058005_H_21
		800906058004_O_16	800906058006_I_20	800906058016_P_13
		800906058003_J_1	800906058008_I_22	802906152002_O_20
		800906058015_A_19	800906058009_C_11	800906058002_G_19
		800906058013_N_18		
62	*T. turgidum L. ssp.* turgidum	800906058012_O_4	800906058006_C_8	800906058002_C_14
		800906058010_P_9	800906058005_N_12	800906058006_E_1
		800906058007_N_19	802906152002_I_6	800906058002_H_3
		800906058005_B_22	800906058006_C_12	800906058001_J_11
		800906058016_H_6	800906058014_H_3	802906152002_M_8
		800906058010_A_11	800906058001_O_17	800906058001_O_12
		800906058014_I_20	800906058005_J_18	800906058006_A_24
		800906058016_A_6	800906058016_G_8	
28	*T. turgidum L. ssp.* turanicum	802906152002_N_3	800906058011_O_11	800906058008_G_17
		800906058004_O_13	800906058006_F_22	800906058006_A_15
		800906058016_O_23	802906152002_O_15	800906058007_G_4
		800906058007_P_21		
139	*T. turgidum L. ssp.* durum 139	800906058001_O_17	800906058001_O_12	800906058004_B_13
		800906058002_C_14	800906058010_O_11	802906247001_E_11
		800906058006_E_1	802906152002_I_2	800906058008_P_15
		800906058002_H_3	800906058003_P_17	800906058004_O_24
		800906058001_J_11	800906058005_O_15	800906058009_N_16
		802906152002_M_8	800906058012_A_7	800906058011_I_9
T	*T. turgidum L. ssp.* Durum Tamaroi	800906058005_F_18	800906058008_M_19	800906058004_D_23
		800906058005_J_20	800906058015_G_13	800906058016_I_13
		802906247001_N_2	800906058003_F_8	800906058016_H_10
		800906058005_O_19	802906247001_A_15	800906058015_M_8
		800906058007_O_22	800906058008_N_10	
W	*T. turgidum L. ssp.* Durum Wolloroi	800906058005_J_14	800906058002_F_11	802906152002_M_24
		800906058007_F_17	802906152002_J_6	800906058001_B_4
		800906058002_G_20	800906058015_A_17	800906058006_P_23
		800906058014_C_16	800906058008_I_4	800906058004_B_4

5.5.3.2.2 Analysis of wild and cultivated wheat

DArTsoft polymorphism analysis of 55 wild and 8 cultivated durum samples identified 552 high quality polymorphisms limited to a call rate of 80 or greater. These 522 markers can be clustered into groups based on their bimodal score. Marker 800906058005_J_18 is scored absent in all cultivars except in the leaf and root tissue of *ssp. turgidum*. Marker 800906058015_G_2 is scored absent in durum 149 leaf and root tissue, but present in all other cultivars. Similarly, marker 800906058005_H_21 is scored present only in polonicum, and marker

802906152002_I_2 only in durum 139. The sample ratio median scores for all samples for these 4 markers are show in Figure 5.11, where the polymorphic cultivar is clearly defined.

Further analysis identified 11 markers that are scored polymorphic between the 4 *ssp. durum* samples and the 4 cultivated durum samples. Of these, 6 markers score present for carthlicum, polonicum, turanicum and turgidum cultivars and 5 markers present for Wollaroi, Tamaroi, durum 139 and durum 149 cultivars in both leaf and root tissue. Comparison of the sample median ratio scores showed that marker 800906058002_N_1 showed the greatest divergence between bimodal clusters, as shown in figure 5.12. Comparison of the other 10 markers showed that even though there was a division between bimodal sample scores, the distinction wasn't as clearly defined when referenced to the sample ratio median values, indicating that they don't discriminate as effectively between the cultivated durum groups and ssp. durum cultivars as well as marker 800906058002_N_1.



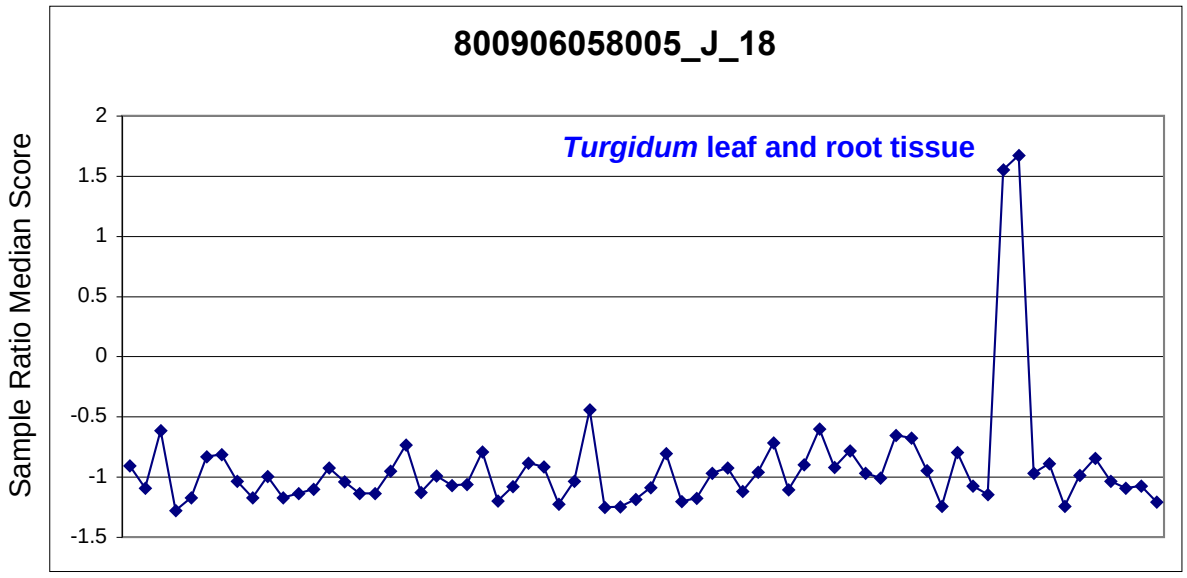
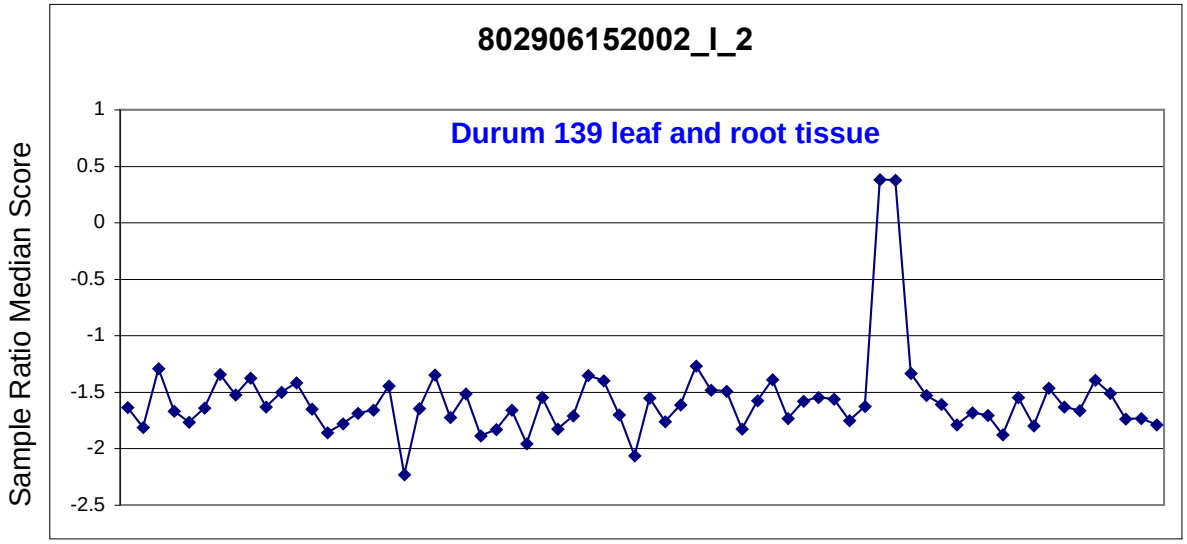


Figure 5.11 (a): Sample Ratio Median scores for selected markers showing cultivar specificity in Durum Discovery Library 2 analysis.



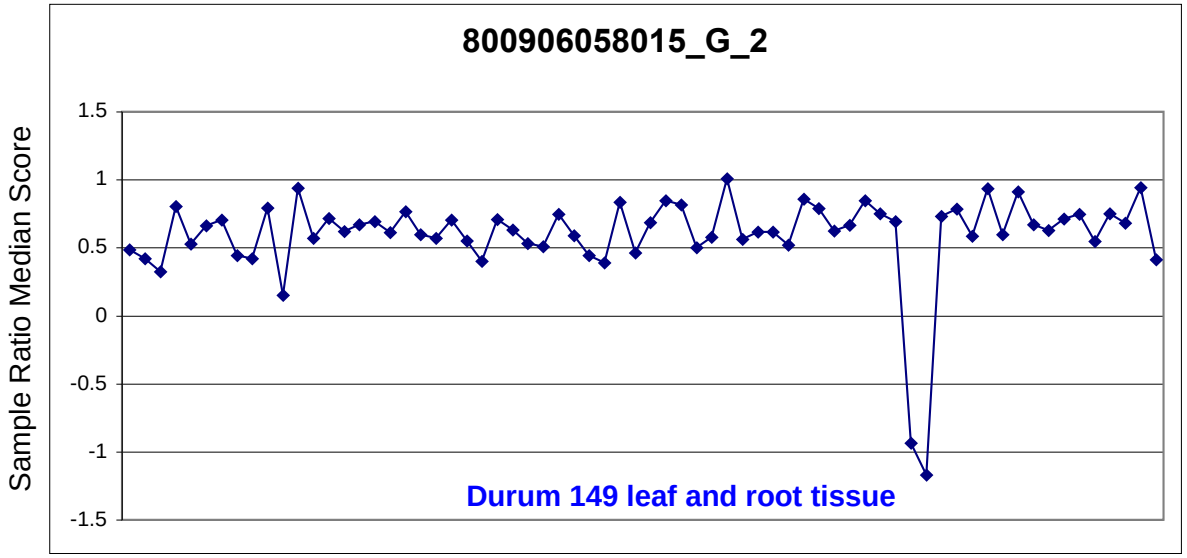


Figure 5.11 (b): Sample Ratio Median scores for selected markers showing cultivar specificity in Durum Discovery Library 2 analysis.

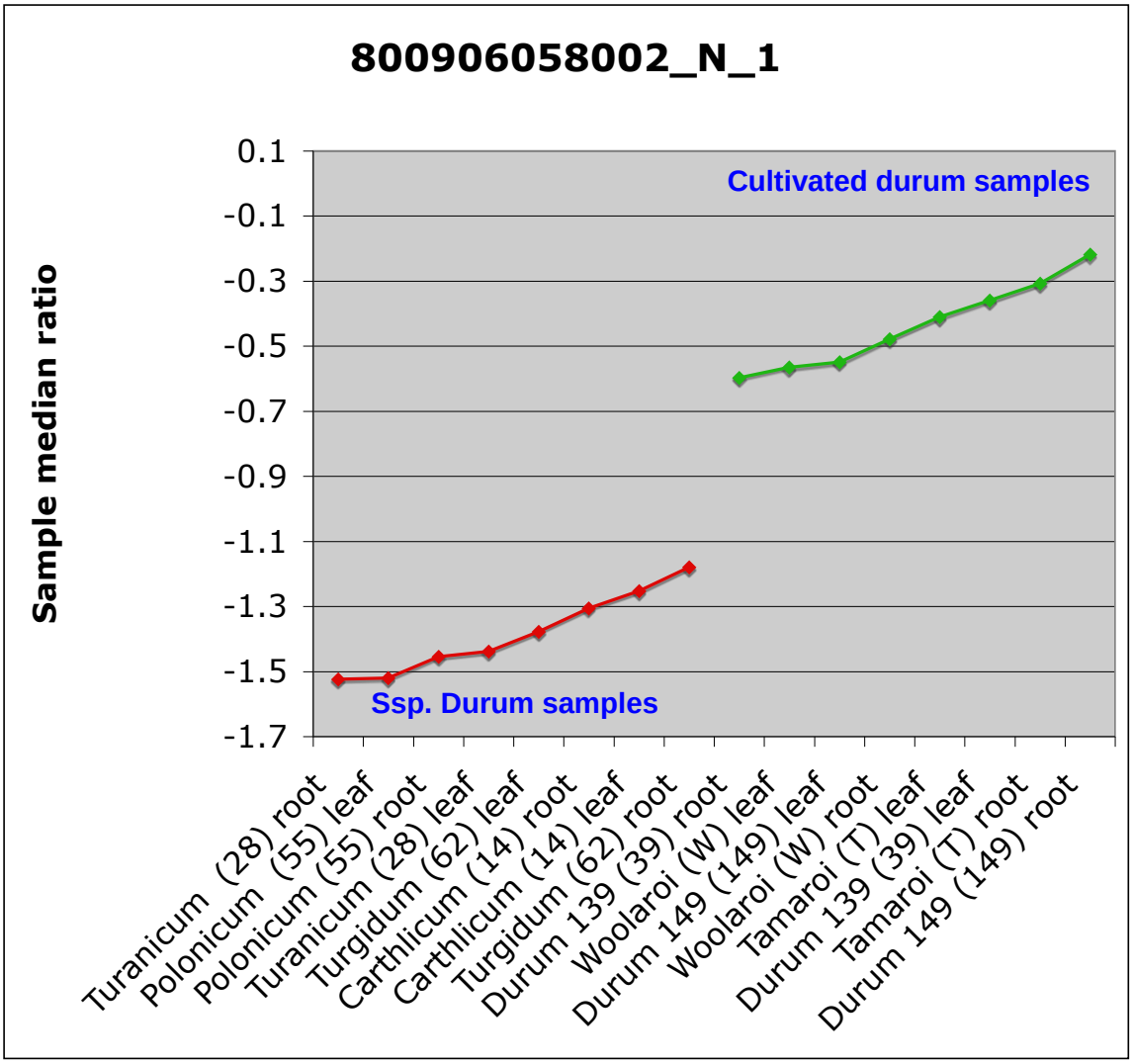


Figure 5.12: Sample Ratio Median scores for marker 800906058002_N_1 showing a clear divergence between ssp. durum and cultivated samples from Durum Discovery Library 2 analysis

5.5.3.2.3 Analysis of carthlicum and Wollaroi cultivars

5.5.3.2.3.1 Analysis using 8 cultivated durum samples

DArTsoft polymorphism analysis of 8 cultivated durum samples identified 737 polymorphisms, with 290 markers differentiating between carthlicum and Wollaroi samples. Of these, 124 markers were scored present in carthlicum and 166 markers in Wollaroi.

5.5.3.2.3.2 Analysis using all 55 wild and cultivated durum samples

DArTsoft polymorphism analysis of 55 wild and 8 cultivated durum samples identified 522 polymorphic markers, with 194 markers polymorphic between carthlicum and Wollaroi samples. Of these, 75 markers were present in carthlicum and 119 markers in Wollaroi.

5.5.3.2.3.3 Comparison of markers identified in both analyses

A comparison was made between markers identified from the overall experiment which included all 55 wild and cultivated durum samples and markers found in the analysis of the 8 cultivated samples alone. Of the 290 markers identified in the cultivated sample analysis and the 194 markers in the wild and cultivated analysis, 184 markers were identified in both analyses, with 70 markers in carthlicum and 114 markers scored present in Wollaroi.

5.5.3.3 Durum Discovery Library 3

The DArT durum discovery library 3 consists of 18 plates totaling 6,912 clones generated from various durum samples printed in duplicate. The array was hybridised using 55 wild leaf samples and 8 leaf and corresponding root samples from cultivated wheats. Samples were analysed using DArTsoft and polymorphisms identified.

5.5.3.3.1 Analysis of 8 cultivated wheat samples

DArTsoft analysis was used to identify polymorphisms between 8 cultivated durum wheat samples from leaf and root tissue. Analysis identified 535 high quality markers limited to a call rate of 80 or greater. Of these, 76 markers were identified as able to differentiate between cultivars. Analysis identified 5 polonicum specific polymorphic markers, 9 turanicum, 12 turgidum, 8 carthlicum, 24 durum, 7 Tamaroi and 3 Wollaroi specific markers. The 76 markers identified as cultivar specific and an additional 2 ssp. durum specific markers are shown in Table 5.9, with an example of the bimodal distribution from each cultivar specific marker shown in table 5.10.

Further analysis identified markers 802906152012_H_20 and 802906152022_D_19 as polymorphic between the spp. durum leaf and root samples and other cultivated durum samples. These markers were scored absent in ssp. durum samples 149, 39, Wollaroi and Tamaroi and present in Carthlicum, Polonicum, Turgidum and Turancium samples. The bimodal scores are shown for marker 802906152012_H_20 in table 5.10 and graphically using the sample ratio median scores in figure 5.13. It can be clearly seen that samples scored absent from the analysis cluster below a sample ratio median score of 0 and samples scored present are scored above 0. This holds true for both leaf and root samples across all 8 cultivars. Marker 802906152022_D_19 follows a similar distribution, with durum samples clustering below a sample ratio median score of 0.5 and non-durum samples above 0.5 as shown in figure 5.14.

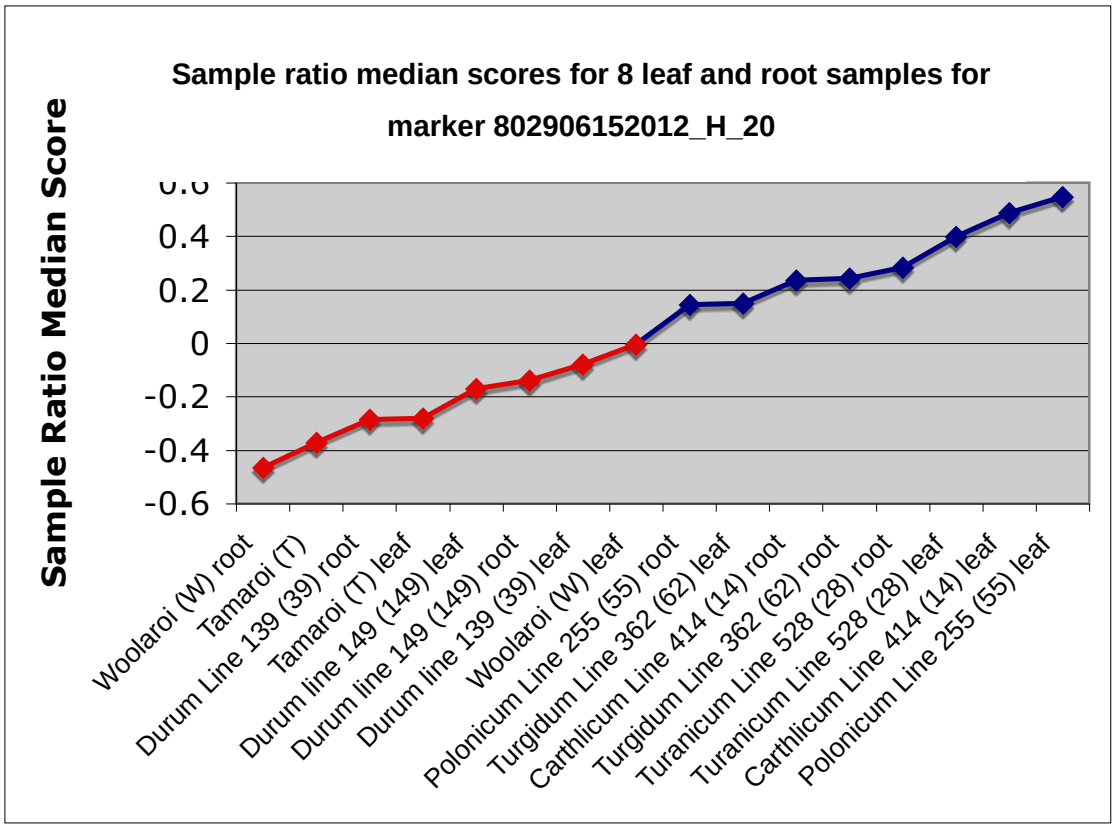


Figure 5.13: Distributon of sample ratio median scores for marker 802906152012_H_20, showing ssp. durum cultivars (red) cluster below 0 for leaf and root samples and samples scored above 0 for durum cultivars (blue).

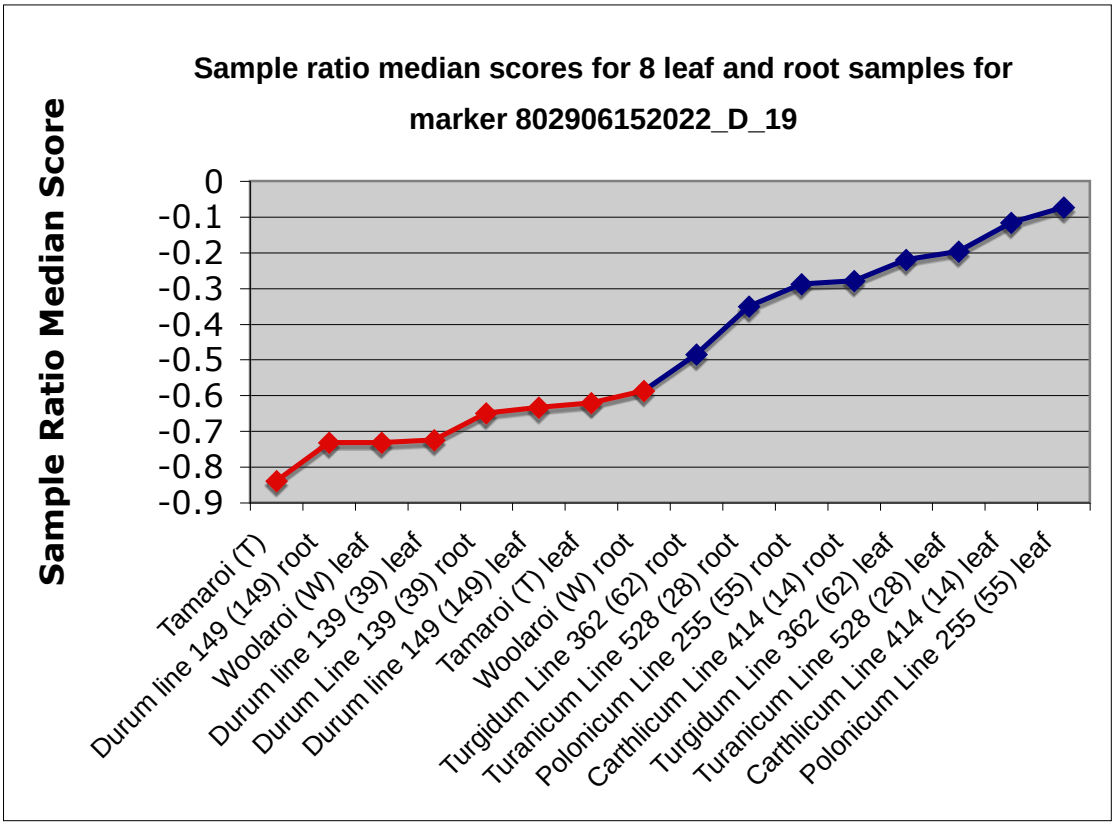


Figure 5.14: Distributon of sample ratio median scores for marker 802906152022_D_19, showing ssp. durum cultivars (red) cluster below 0.5 for leaf and root samples and samples scored above 0.5 for durum cultivars (blue).

Table 5.9: Cultivar specific durum wheat markers – Library 3

Designation	Species	Cultivar Specific markers		
14	*T. turgidum L. ssp.* carthlicum	802906152013_C_6	802906152025_N_6	802906152009_D_23
		802906152006_E_6	802906152019_N_15	802906152024_M_19
		802906152012_M_6	802906152007_O_24	
149	*T. turgidum L. ssp.* durum 149	802906152020_D_8	802906152013_A_17	802906152021_L_16
		802906152012_J_21	802906152011_F_11	802906152025_F_2
		802906152012_G_3	802906152022_E_6	802906152005_A_20
		802906152010_G_20	802906152010_O_18	802906152011_C_22
		802906152013_H_10	802906152020_K_3	802906152022_B_10
		802906152013_D_18	802906152018_B_17	802906152003_G_21
		802906152013_J_6	802906152005_P_18	802906152021_K_20
		802906152007_G_24	802906152005_G_18	802906152005_L_7
55	*T. turgidum L. ssp.* polonicum	802906152025_L_8	802906152019_J_18	802906152021_M_1
		802906152021_F_18	802906152011_H_24	
62	*T. turgidum L. ssp.* turgidum	802906152009_D_9	802906152021_H_12	802906152007_H_11
		802906152020_I_16	802906152011_E_23	802906152009_I_10
		802906152020_M_16	802906152013_M_8	802906152012_G_10
		802906152013_G_17	802906152020_F_1	802906152019_D_20
28	*T. turgidum L. ssp.* turanicum	802906152018_O_9	802906152010_E_24	802906152020_C_19
		802906152025_M_13	802906152012_O_20	802906152012_L_8
		802906152020_G_7	802906152006_O_24	802906152021_H_5
139	*T. turgidum L. ssp.* durum 139	802906152008_E_5	802906152022_A_2	802906152012_E_15
		802906152020_H_23	802906152020_P_12	802906152019_B_18
		802906152018_N_24	802906152010_N_20	
T	*T. turgidum L. ssp.* durum Tamaroi	802906152007_P_21	802906152012_F_23	802906152025_A_10
		802906152018_A_11	802906152018_I_15	802906152020_I_20
		802906152020_C_16		
W	T. turgidum L. ssp. durum Wolloroi	802906152025_D_6	802906152013_A_3	802906152003_A_7
	ssp. durum specific markers	802906152012_H_20	802906152022_D_19	

Table 5.10: Polymorphic cultivated durum wheat samples – Library 3

Marker	Ssp. polonicum (55)		Ssp. turanicum (28)		Ssp. turgidum (62)		Ssp. carthlicum (14)		Ssp. durum (149)		Tamaroi (T)		Wollaroi (W)		Ssp. durum (39)	
Tissue	root	leaf	root	leaves	root	leaves	root	leaves	root	leaves	root	leaves	root	leaves	root	leaves
802906152012_H_20	1	1	1	1	1	1	1	X	0	0	0	0	0	0	0	0
802906152011_H_24	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
802906152021_M_1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
802906152025_M_13	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
802906152006_O_24	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1
802906152021_H_12	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
802906152011_E_23	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
802906152012_M_6	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
802906152019_N_15	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1
802906152011_F_11	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
802906152020_K_3	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1
802906152020_I_20	0	0	0	0	0	0	0	0	0	0	X	1	0	0	0	0
802906152007_P_21	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1
802906152025_D_6	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1
802906152003_A_7	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
802906152008_E_5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0

5.5.3.3.2 Analysis of wild and cultivated wheat

DArTsoft polymorphism analysis of 55 wild and 8 cultivated durum samples identified 166 high quality polymorphisms limited to a call rate of 80 or greater. Of these markers, 25 were identified as cultivar specific. These included 3 markers that were carthlicum specific, 1 marker polonicum specific, 2 markers turanicum specific, 3 markers turgidum specific, 4 markers durum 149 specific, 7 markers Tamaroi specific, 2 markers Wollaroi specific and no markers identified in durum 139. Figure 5.15 shows 9 selected markers within this group that clearly show cultivar specific markers, where the 8 cultivated leaf and root wheat samples are graphed and cluster together with the exception of the cultivar that has a specific polymorphism. The leaf and root sample for that cultivar shows the sample ratio median score clearly divergent to the main cluster. Markers 802906152012_G_3, 802906152012_J_21, 802906152005_L_7 and 802906152020_D_8 show a divergence of Durum line 149 leaf and root samples from the cluster of 8 cultivar samples, indicating a cultivar specific polymorphism. Marker 802906152003_A_7 shows a Wollaroi specific marker, 802906152007_P_21 and 802906152018_A_11 Tamaroi markers, 802906152025_L_8 a Polonicum marker and 802906152020_M_16 a Turgidum marker.

5.5.3.3.3 Analysis of ssp. carthlicum and Wollaroi cultivars

5.5.3.3.3.1 Analysis using 8 cultivated durum samples

DArTsoft polymorphism analysis of 8 cultivated durum samples identified 535 polymorphisms, with 57 markers differentiating between carthlicum and Wollaroi leaf and root samples. Of these, 29 markers were scored present in carthlicum and 28 markers in Wollaroi.

5.5.3.3.3.2 Analysis using all 55 wild and cultivated durum samples

DArTsoft polymorphism analysis of 55 wild and cultivated durum samples identified 522 polymorphic markers, with 118 markers polymorphic between carthlicum and Wollaroi leaf and root samples. Of these, 65 markers were present in carthlicum and 53 markers in Wollaroi.

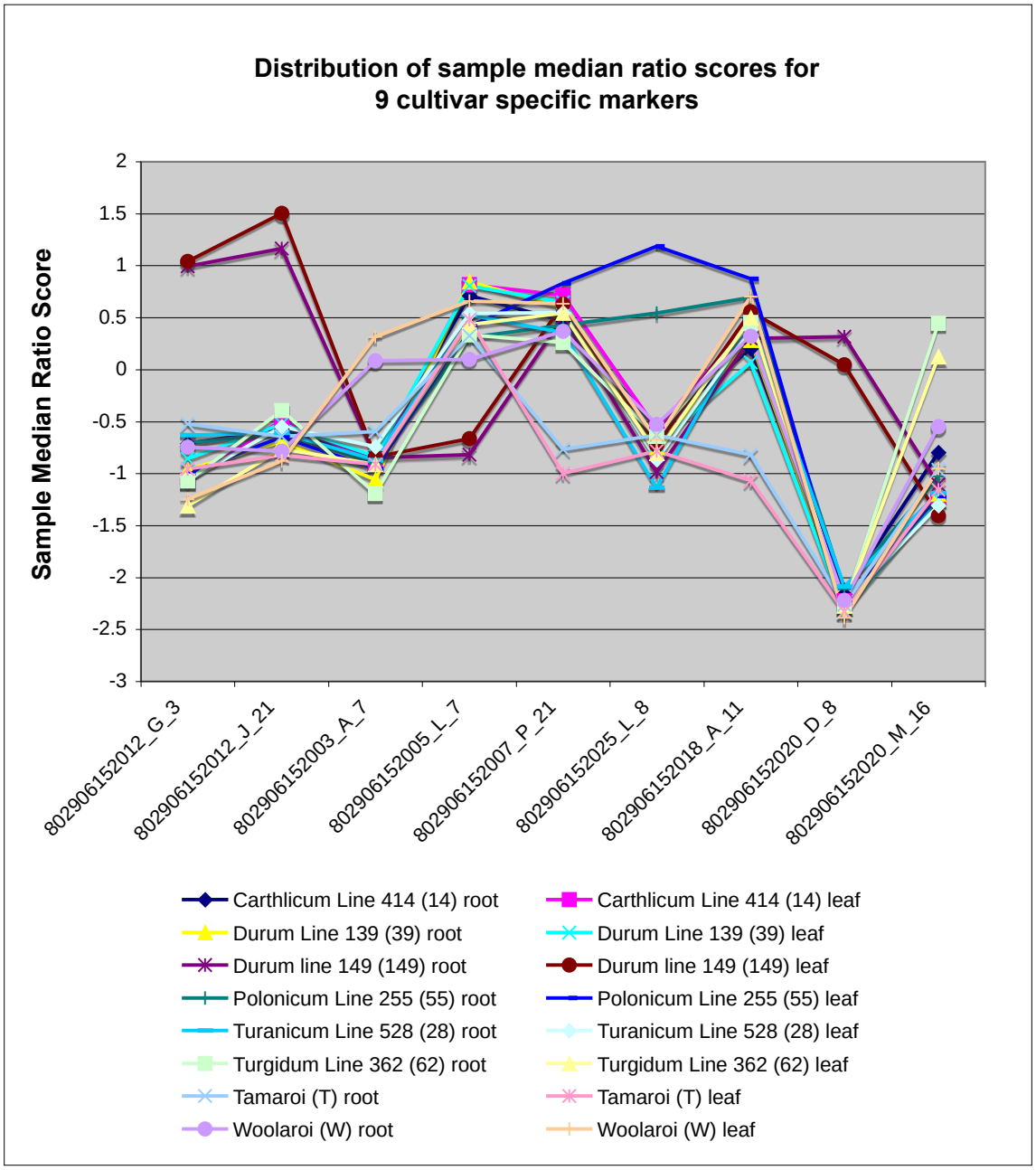


Figure 5.15: Distributon of sample median ratio scores for 9 markers, showing the cluster of 7 cultivars with the polymorphic leaf and root sample diverging away from the cluster, forming a cultivar specific polymorphic marker.

Marker 802906152012_G_3 shows the divergence of Durum Line 149 leaf and root samples forming a distinct and separate cluster when the sample median ratios are graphed.

5.5.3.3.3.3 Comparison of markers identified in both analyses

A comparison was made between markers identified from the 8 cultivated samples and those identified using the 55 wild samples. Of the 57 markers identified in the cultivated sample analysis and the 118 markers in the wild and cultivated analysis, 48 markers were identified in both analyses, 29 markers scored present in carthlicum and 28 markers in Wollaroi.

5.5.3.4 Durum Discovery Library 4

The DArT durum discovery library 4 consists of 17 plates totaling 6,528 clones generated from various durum samples printed in duplicate. The array was hybridised using 55 wild leaf samples and 8 leaf and corresponding root samples from 8 cultivated wheats. Samples were analysed using DArTsoft and polymorphisms identified.

5.5.3.4.1 Analysis of 8 cultivated wheat samples

DArTsoft analysis was used to identify polymorphisms between 8 cultivated durum wheat samples from leaf and root tissue. Analysis identified 1,144 high quality markers limited to a call rate of 80 or greater. Of these, 123 markers were identified as differentiating between wheat cultivars, having the same scores for leaf and root tissue. For each of the 123 cultivar specific polymorphic markers, the 8 cultivars can be grouped according to the presence or absence of that marker. Analysis identified 23 markers specific to carthlicum, 8 markers specific to polonicum, 20 markers specific to turanicum, 31 markers specific to turgidum, 1 marker specific to durum 139, 8 markers specific to durum 149, and 16 markers in each of Wollaroi and Tamaroi leaf and root samples. A list of markers is shown in table 5.11.

Table 5.11: Cultivar specific durum wheat markers – Library 4

Designation	Species	Cultivar Specific markers		
14	*T. turgidum L. ssp.* carthlicum	800904161004_P_7	800904300006_H_9	802906152015_L_7
		800904300003_L_5	800904300004_K_22	801504280004_B_5
		800904161004_O_7	802906152015_D_10	801504280002_O_21
		800904300004_D_1	801504280005_F_8	800904300005_J_17
		800904300004_C_11	801504280002_P_19	801504280005_A_8
		800904300002_D_4	802906152001_O_13	801504280004_B_6
		801504280005_O_5	800904161004_M_19	801504280003_L_23
		802906152004_O_9	800904300006_E_18	
149	*T. turgidum L. ssp.* durum 149	801504280004_C_1	800904300004_O_11	801504280001_O_3
		800904300003_N_2	801504280001_A_13	801504280006_D_23
		800904300003_D_6	800904161004_E_19	
55	*T. turgidum L. ssp.* polonicum	801504280001_K_20	801504280004_N_13	800904161004_M_15
		800904300004_L_23	801504280004_D_23	800904090002_A_23
		800904300004_K_24	801504280004_F_10	
62	*T. turgidum L. ssp.* turgidum	800904090002_H_8	801504280005_P_5	800904300003_H_20
		802906152001_O_19	801504280002_F_14	801504280003_O_12
		801504280001_N_3	800904090002_F_10	802906152015_J_10
		802906152001_J_23	802906152004_C_13	801504280006_H_5
		801504280005_J_6	801504280001_M_1	801504280005_E_3
		802906152001_F_4	800904090002_A_14	801504280006_L_6
		801504280004_I_20	801504280004_N_15	800904300006_D_15
		802906152001_G_13	801504280003_M_18	800904300004_M_21
		800904300001_O_6	801504280003_P_12	800904300005_M_2
		800904300006_J_15	802906152015_E_21	801504280006_H_18
		802906152001_H_7		
28	*T. turgidum L. ssp.* turanicum	800904300006_A_17	801504280004_F_13	801504280001_O_6
		800904300003_N_7	801504280001_E_16	802906152015_D_24
		801504280004_C_23	802906152001_A_3	800904161004_M_5
		800904300005_K_19	800904300005_P_24	800904300003_F_11
		801504280004_G_13	800904300005_O_24	800904300001_L_6
		801504280004_K_11	801504280003_A_22	800904300004_N_20
		800904300004_D_20	801504280004_E_4	
139	*T. turgidum L. ssp.* durum 139	800904300001_I_3		
T	*T. turgidum L. ssp.* Durum Tamaroi	800904300001_I_15	800904300004_H_22	802906152015_L_11
		801504280004_E_16	800904300006_J_21	800904161004_K_23
		802906152015_D_11	801504280005_M_1	801504280006_K_10
		800904090002_B_9	801504280004_I_13	801504280004_N_20
		800904300001_D_18	801504280004_K_10	800904090002_G_16
		800904300001_H_9		

W	*T. turgidum L. ssp.* Durum Wolloroi	800904161004_O_8	802906152015_J_15	801504280001_L_5
		801504280003_N_21	801504280002_E_13	801504280003_P_6
		800904090002_B_19	800904090002_M_16	800904090002_B_10
		800904300002_K_15	802906152004_A_10	801504280002_I_23
		801504280004_L_15	801504280004_A_11	800904300006_C_4
		800904300005_C_5		

Further to the cultivar specific markers, 4 ssp. durum specific markers were identified. Markers 801504280001_C_14, 801504280001_O_10, 801504280004_P_19 and 800904300002_F_9 were found to discriminate between ssp. durum cultivars Wollaroi, Tamaroi, Durum lines 139 and 149 as apposed to durum culitvars, carthlicum, polonicum, turanicum and turgidum cultivars. The sample ratio median scores are graphed in figure 5.16 where the divergence in scores can clearly be seen. Samples scored higher than 0 are scored present in the analysis and markers scored below 0 are absent from analysis.

5.5.3.4.2 Analysis of wild and cultivated wheat

Analysis of 55 wild and 8 cultivated durum wheat samples identified 538 high quality markers, limited to a call rate of 80 or greater. Within this group, analysis identified 7 carthlicum specific markers, 4 polonicum specific markers, 5 turancium specific markers, 8 turgidum specific markers, 12 durum line 149 markers, 1 durum line 139 marker, 14 Tamaroi specific markers and 16 Wollaroi specific markers. Analysis of the same 4 ssp. durum specific markers shown in figure 5.16 follow the same pattern of bimodal segregation for all 68 samples, as shown in figure 5.17. It can be seen that the 8 cultivated wheat samples follow the same segregation, with the 55 wild samples falling into one of the two clusters representing the present (1) and absent (0) groups. The cluster allocation for each sample changes for each marker, such as in Wollaroi where markers 801504280001_C_4, 800904300002_F_9 and 801504280004_P_19 are scored present and marker 801504280001_O_10 scored absent (figure 5.18). In comparison, the Reva samples show that markers 801504280001_C_4 and 800904300002_F_9 are also scored present, however marker 801504280004_P_19 is now scored absent with marker 801504280001_O_10 (figure 5.18). This shows that not all markers are scored the same way across multiple cultivars.

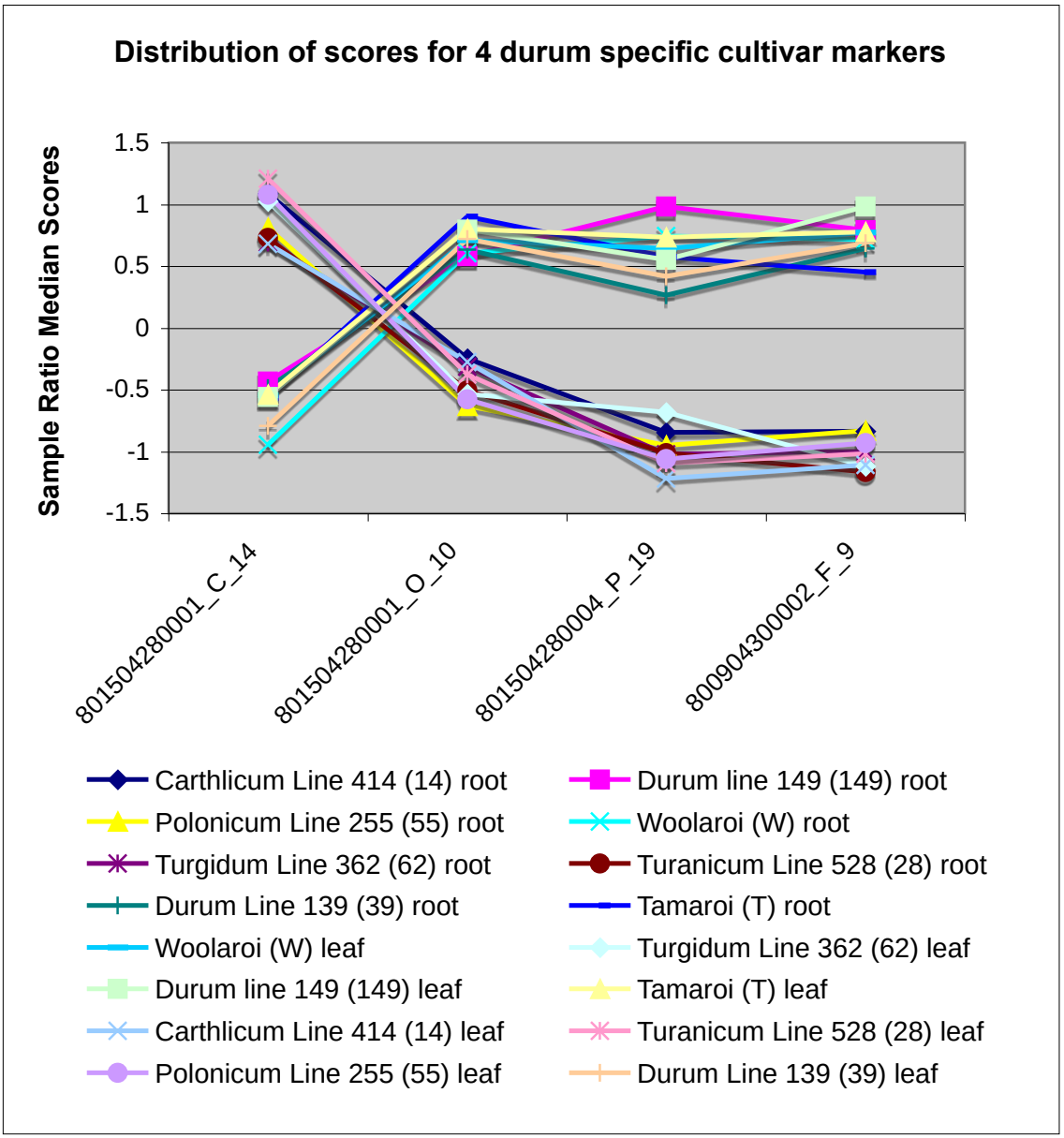


Figure 5.16: Distributon of sample ratio median scores for 4 durum specific markers, showing clear biomodal divergence, the ssp. durum samples (Wollaroi, Tamaroi, Durum lines 139 and 149) and the durum samples (carthlicum, polonicum, turanicum and turgidum cultivars).

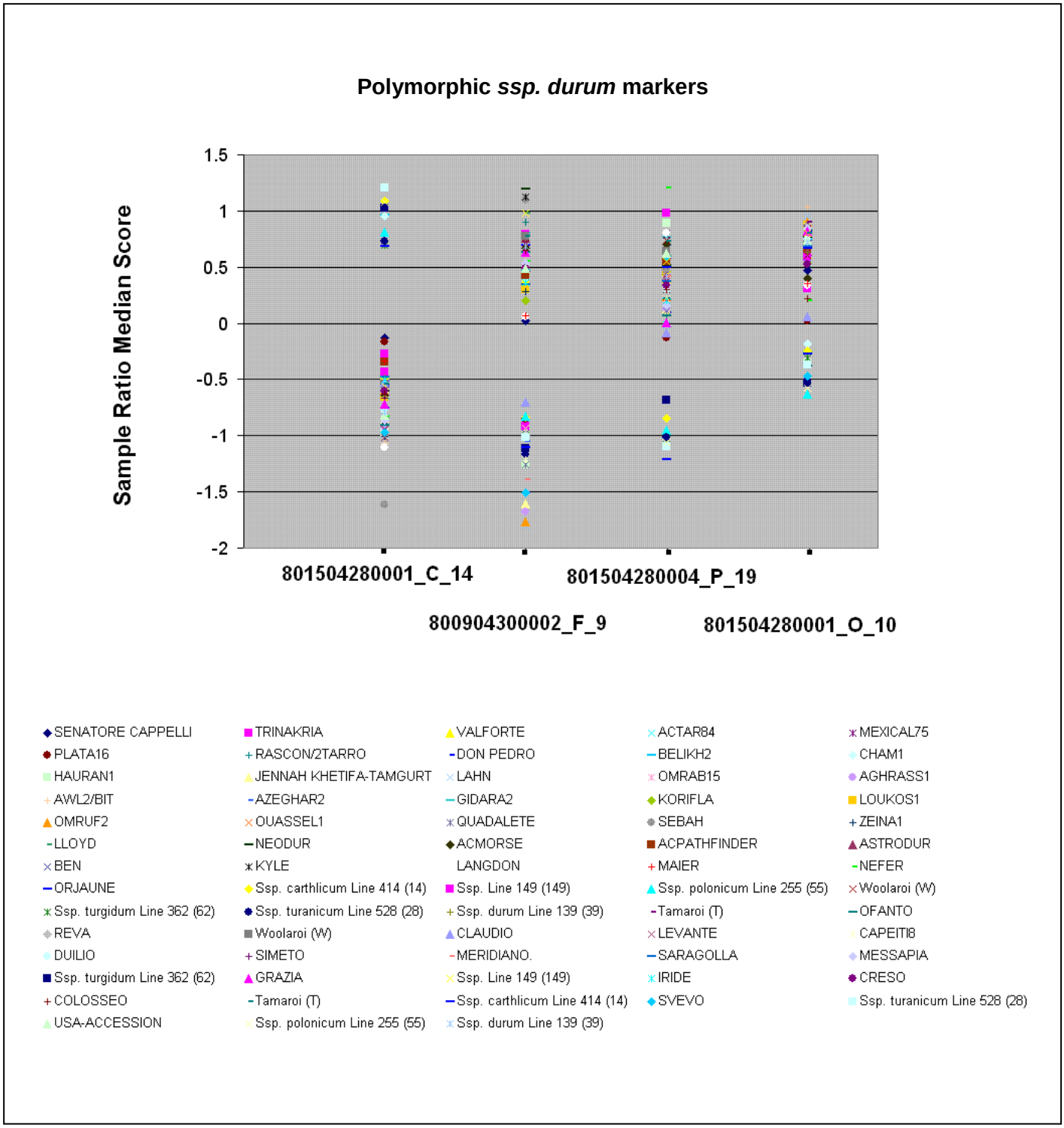
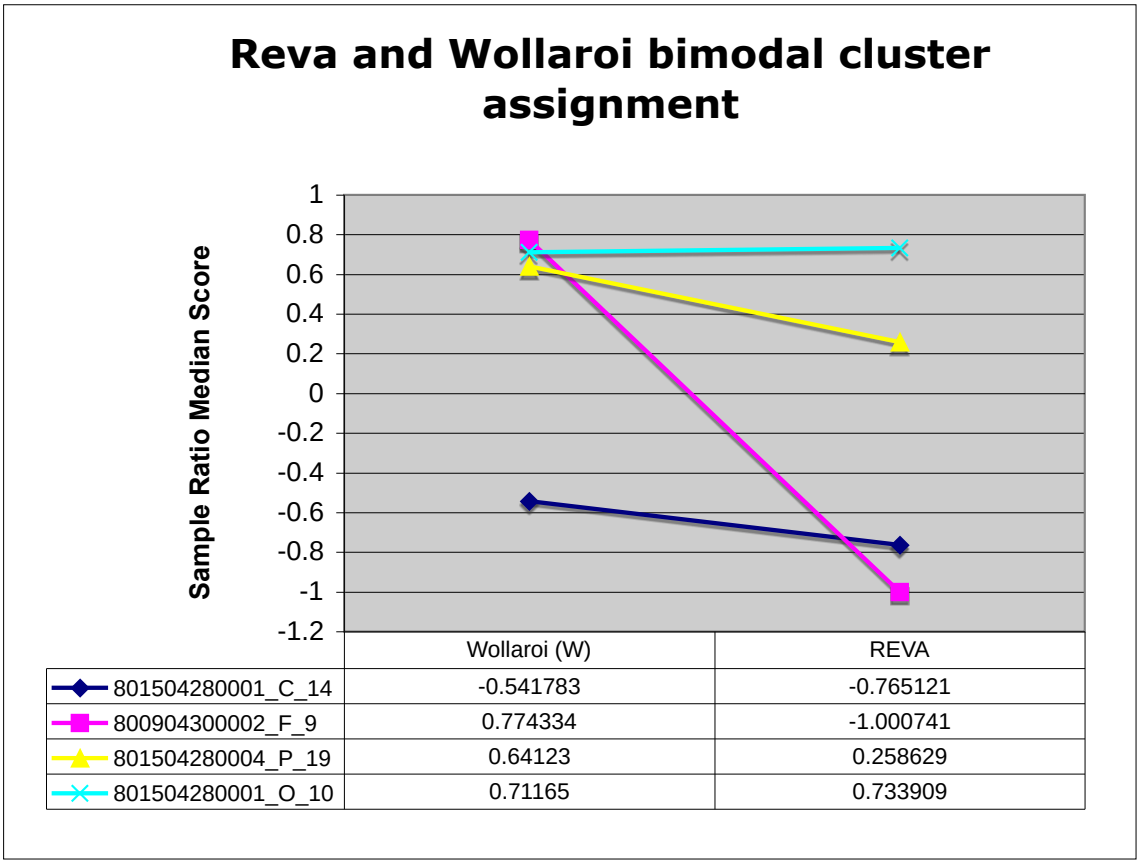


Figure 5.17: Clustering of sample ratio median scores for 55 wild and 8 cultivated samples for 4 markers identified as polymorphic between ssp. durum cultivars and other durum clutivars from durum library 4



	Wollaroi (W)	REVA
801504280001_C_14	-0.541783	-0.765121
800904300002_F_9	0.774334	-1.000741
801504280004_P_19	0.64123	0.258629
801504280001_O_10	0.71165	0.733909

Figure 5.18: Reva and Wollaroi bimodal cluster assignment for 4 durum markers, showing that not all markers are scored in the same cluster across cultivars.

5.5.3.4.3 Analysis of carthlicum and Wollaroi cultivars

5.5.3.4.3.1 Analysis using 8 cultivated durum samples

DArTsoft polymorphismd analysis of the 8 cultivated durum samples identified 717 high quality markers, with 384 markers polymorphic between carthlicum and Wollaroi. Of these, 216 were scored present in carthlicum and 168 markers in scored present in Wollaroi.

5.5.3.4.3.2 Analysis using all 55 wild and cultivated durum samples

DArTsoft polymorphism analysis of 55 wild and 8 cultivated durum samples identified 385 high quality markers, with 199 markers polymorphic between carthlicum and Wollaroi. Of these, 110 markers were scored present in carthlicum and 89 markers present in Wollaroi.

5.5.3.4.3.3 Comparison of markers identified in both analyses

A comparison was made between markers that were identified from the overall experiment which included all 55 wild and 8 cultivated durum samples and markers found in the analysis of the 8 cultivated samples alone. Of the 384 markers identified in the cultivated sample analysis and the 199 markers in the wild and cultivated analysis, 190 markers were identified in both analyses, with 103 markers in carthlicum and 87 markers scored present in Wollaroi. These included the 4 selected markers in figures 5.15 and 5.16.

5.5.4 Conclusions

From the results described, it can clearly be seen that potential polymorphic markers have been identified between all 8 cultivars. These markers can be used for identifying and genotyping cultivar samples. Further, these markers can be linked to traits of interest, and followed through breeding experiments.

From the 4 durum libraries, 475 markers were identified that differentiated between one of the 8 cultivated samples when analysed using only the 8 cultivated samples (table 5.12). This equates to 475 cultivar specific polymorphic markers out of 27,648 total markers (clones) over all 4 arrays, or 1.72%. On average, 119 cultivar specific markers were identified on each of the arrays and 59 markers were specific for each cultivar. Analysis of Wollaroi and carthlicum samples using DArTsoft analysis of 8 cultivated samples, 55 wild and 8 cultivated samples and the number of markers found in both analyses, found 387, 300 and 199 carthlicum specific markers and 383, 285 and 234 Wollaroi specific markers. In total, 770, 585 and 433 markers were identified that could differentiate between carthlicum and Wollaroi samples from the samples analysed (table 5.13). The large variance in marker numbers from analyses of 8 and 8/55 samples is due to differences in samples sizes.

Table 5.12: Summary of cultivar specific markers identified in durum libraries 1 – 4 from analysis of 8 cultivated durum leaf and root samples.

Cultivar	Library 1	Library 2	Library 3	Library 4	Total
Ssp. polonicum (55)	11	16	5	8	40
Ssp. turanicum (28)	9	10	9	20	48
Ssp. turgidum (62)	17	33	12	31	93
Ssp. carthlicum (14)	9	20	8	23	60
Ssp. durum 149	29	54	24	8	115
Ssp. durum Wollaroi (W)	8	12	3	16	39
Ssp. durum Tamaroi (T)	6	14	7	16	43
Ssp. durum 139	10	18	8	1	37
Total	99	177	76	123	475

Table 5.13: Comparison of carthlicum and Wollaroi markers analysed with (a) 8 cultivated samples, (b) 55 wild and 8 cultivated samples, (c) and markers found in both analyses.

Cultivar	Library 1			Library 2			Library 3			Library 4			Total		
	(a)	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**
Ssp. carthlicum (14)	18	50	3	124	75	70	29	65	23	216	110	103	387	300	199
Wollaroi (W)	21	24	8	166	119	114	28	53	25	168	89	87	383	285	234
Total	39	74	11	290	194	184	57	118	48	384	199	190	770	585	433

5.6 Diversity in durum and polonicum

5.6.1 Aims

Durum line 39 (139) and polonicum (55) leaf, growing region and root samples were analysed using DArT to detect cultivar specific polymorphisms with the aim to identify salt tolerance molecular markers. Improving the salt tolerance of crop and pasture species requires access to new genetic diversity (either natural or transgenic) and efficient techniques for identifying salt-tolerance. Francois *et al*. (1986)[193] and Gorham *et al*. (1987)[194] describe that genetic differences in Na+ exclusion are highly correlated with differences in salinity tolerance between tetraploid and hexaploid wheat. Durum and polonicum were selected as two tetraploid cultivars that show differing tolerance to salinity, as described in Munns *et al* (2003) who looked at Na+ accumulation. Durum line 39 gave a total leaf Na+ per % dead leaf concentration of 42 µmol compared to polonicum that gave 93 µmol. Skiff, a barley cultivar, was included as a control, as barley is naturally salt tolerant, and gave a reading of 107. Thus for this analysis, durum 39 will be considered salt in-tolerant and polonicum as salt tolerant. Other screening methods are available to evaluate genetic diversity including methods based on growth or yield, damage, tolerance to very high salinity levels or physiological mechanisms.

Molecular markers technology can reduce the work involved in phenotypic screens. Once a locus (QTL) or gene associated with a specific trait is identified, a PCR based molecular marker can be developed. Markers can be tested on seeds or seedlings, and provide a cost effective way of screening large numbers of individuals in a segregating population. Molecular marker analysis is non-destructive and does not require controls or salt treatments. DArT was employed to evaluate diversity between durum and polonicum tetraploid cultivars that have been previously shown to have varying levels of salt tolerance[195].

5.6.2 Specific Methods

Ssp. *polonicum* (55) and *T. turgidum L. ssp. durum* (39) tetraploid wheat was grown in a glass house at the Crown Scientific and Industrial Research Organisation (CSIRO), Black Mountain, ACT. Plants were grown in quadruplicate under control, incrementally increased salt and all-at-once shock salt conditions. Seedling DNA

samples were also used, that were grown in a Petri dish at room temperature for 5-7 days from seed, where the first leaf was harvested. The DArT method uses a combination of restriction enzymes during the complexity reduction steps. The standard PstI and Taq$^{\alpha}$I digestion was used and McrBC was combined with PstI/Taq$^{\alpha}$I for further methylation discrimination (see Chapter 4). PstI/ Taq$^{\alpha}$I and PstI/ Taq$^{\alpha}$I /McrBC samples were analysed separately to identify cultivar specific polymorphic markers between durum and polonicum samples. In addition, PstI/MseI and PfiMI/MseI restriction enzymes were used to further methylation studies (see Chapter 4) and the cultivar specific polymorphisms reported here. Samples were hybridised to the DArT Wheat 8 plate array (V2.2 May 06) containing a combination of polymorphic rich plates from various wheat libraries. Specifically, wheat 2.1.1 (plates1-4), wheat 2.1.3 (plates 6,7), wheat 2.1.4 (plates 1-4) and wheat 2.1.5 (plates 5,8) were used.

5.6.3 Results

A series of experiments were performed and results summarised in table 5.14. Analysis of Polonium and Durum 39 samples showed that DArT was able to discriminate between the two cultivars providing Polonium-like and Durum 39-like candidate molecular markers. These polymorphisms may not be cultivar specific, as analysis was only performed between the two samples, but the markers can discriminate between the two samples.

5.6.3.1 Restriction Enzyme Digests

Analysis was performed using various restriction enzymes combinations during the complexity reduction steps of the DArT protocol. PstI/Taq$^{\alpha}$I is a routine enzyme combination where PstI is the primary 6 base pair cutter (not methylation sensitive) and Taq$^{\alpha}$I the secondary 4 base pair cutter (dam methylation sensitive). Analysis of a PstI/Taq$^{\alpha}$I experiment found that from a 24 slide, duplicate target experiment using seedling, incremental and control samples that 495 of the 3072 triple replicated spots on the array were scored polymorphic in Cy3 and 476 in Cy5. Comparison of both sets showed that 358 markers were scored polymorphic in both Cy3 and Cy5 analysis. A second experiment was performed using control and incremental samples with 4 replicate targets over the same 3072 triple replicated array. Analysis

identified 387 markers in Cy3 and 405 markers in Cy5 that were polymorphic between polonicum and durum 39 cultivars over the 7 total replicates.

In contrast, analysis of a 32 slide experiment with 768 spots printed in quadruplicate using PfiMI and MseI with 4 targets per sample identified 18 markers in Cy3 and 33 markers in Cy5, considerable less than the PstI/Taq$^{\alpha}$I library. Comparison of polymorphic markers identified in both analyses identified 13 polymorphic markers common to both. PfiMI restriction enzyme is an 11 base pair cutter (blocked by dcm methylation) while MseI is a 4 base pair cutter (not methylation sensitive).

Similarly, analysis of a 24 slide experiment with 768 spots printed in quadruplicate using PstI and MseI with 4 targets per sample identified 65 markers in Cy3 and 75 markers in Cy5. Comparison of polymorphic markers identified in both analyses identified 36 markers common to both. PstI is a 6 base pair cutter that is not methylation sensitive.

Overall, the PstI/Taq$^{\alpha}$I digestion gave a higher proportion of markers that were identified as polymorphic between polonicum and durum 39 cultivars.

5.6.3.2 Seedling cultivar polymorphic markers

Figure 5.19 shows 55 randomly selected seedling cultivar specific polymorphisms that were bimodally scored the same across replicates and tissue types in Cy3 targets. When the sample ratio median scores are graphed for each marker (figure 5.19), it can be seen that the polonicum leaf and root scores are bimodally divergent compared to the durum 39 leaf and root scores. Markers are sorted according to the difference between the average polonicum score and the average durum 39 score.

Table 5.14: Summary of durum and polonicum 39 specific polymorphisms using differing complexity reduction methods

Sample	Representation	High Quality Polymorphisms	Cultivar polymorphisms	Durum specific	Polonicum specific
Seedling Cy3	PstI /Taq$^{\alpha}$I	1,036	639	308	331
Seedling Cy5	PstI /Taq$^{\alpha}$I	1,024	635	320	315
Seedling Cy3 vs. Cy5	PstI /Taq$^{\alpha}$I	-	441	192	249
Control Cy3	PstI /Taq$^{\alpha}$I	822	523	232	268
Control Cy5	PstI /Taq$^{\alpha}$I	918	501	251	250
Control Cy3 vs. Cy5	PstI /Taq$^{\alpha}$I	-	404	220	184
Seedling and Control Cy3	PstI /Taq$^{\alpha}$I	822	399	172	227
Seedling and Control Cy5	PstI /Taq$^{\alpha}$I	918	393	185	208
Seedling and Control Cy3 vs. Cy5	PstI /Taq$^{\alpha}$I	-	319	138	181
Incremental Cy3	PstI /Taq$^{\alpha}$I	594	481	224	257
Incremental Cy5	PstI /Taq$^{\alpha}$I	523	503	259	244
Incremental Cy3 and Cy5	PstI /Taq$^{\alpha}$I	-	383	182	201
Control and Incremental Cy3	PstI /Taq$^{\alpha}$I	594	411	178	223
Control and Incremental Cy5	PstI /Taq$^{\alpha}$I	523	334	154	180
Control and Incremental Cy3 and Cy3	PstI /Taq$^{\alpha}$I		313	139	174
Incremental Cy3	PflMI / MseI	31	15	9	6
Incremental Cy5	PflMI / MseI	41	28	12	16
Incremental Cy3 vs. Cy5	PflMI / MseI	-	12	6	6
Control and Incremental Cy3	PflMI / MseI	31	12	6	6
Control and Incremental Cy5	PflMI / MseI	41	19	11	8
Control and Incremental Cy3 and Cy5	PflMI / MseI	-	12	6	6
Control, Incremental and Shock Cy3	PstI / MseI	39	31	14	17
Control, Incremental and Shock Cy5	PstI / MseI	47	31	14	17
Control, Incremental and Shock Cy3 and Cy5	PstI / MseI	-	28	13	15
Control, Incremental and Shock Cy3	PflMI / MseI	67	32	8	24
Control, Incremental and Shock Cy5	PflMI / MseI	19	10	3	7
Control, Incremental and Shock Cy3 and Cy5	PflMI / MseI	-	9	2	7
Seedling, Control, Incremental and Shock Cy3	PstI / MseI	72	50	17	33
Seedling Control, Incremental and Shock Cy5	PstI / MseI	93	51	17	34
Seedling, Control, Incremental and Shock Cy3 + Cy5	PstI / MseI	-	35	10	25

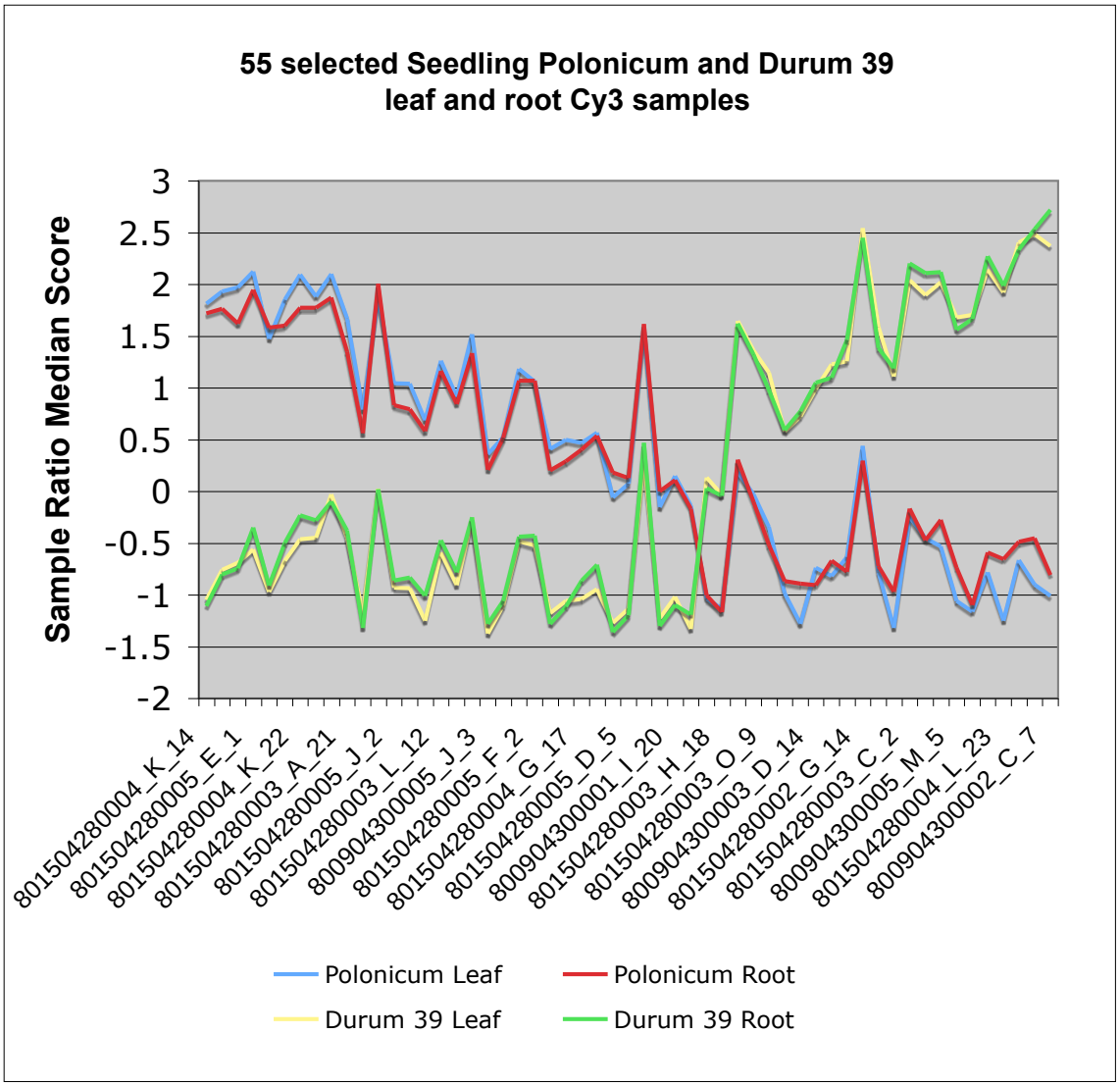


Figure 5.19: 55 selected markers showing cultivar specific polymorphisms between polonicum and durum 39 seedling Cy3 samples, sorted by decreasing divergence.

Marker 801504280004_K_14 (wPt-3378) scored 100 for reproducibility and call rate, P=99.55 and Q=91.25 and shows a polymorphic divergence of 2.85 when the average sample ratio median score for polonicum (leaf 1.81 and root 1.71) and durum (leaf -1.05 and root -1.11) are calculated. This marker is scored present for polonicum and absent for durum.

Similarly, marker 800904300002_C_7 (wPT-0280) scored 100 for reproducibility and call rate, P=99.20 and Q=90.94, shows a divergence of 3.45 in sample ratio median scores. The polonicum leaf score was -1.00 and root -0.81 and durum leaf 2.37 and root 2.72, showing that the marker was scored present in durum samples and absent in polonicum samples.

When replicate Cy5 samples were analysed, the same two markers were scored polymorphic between cultivars. Marker 801504280004_K_14 scored 100 for reproducibility and call rate, P=96.82 and Q=88.75, and was shown to have a divergence of 3.01 between polonicum (leaf 2.11 and root 1.45) compared to durum (leaf -1.26 and root -1.33) samples. This marker was thus scored present in polonicum samples and absent in durum samples (as in Cy3 analysis). Marker 800904300002_C_7 also scored 100 for reproducibility and call rate, P=99.58 and Q=91.28, and was shown to have a divergence of 4.61 between polonicum (leaf -1.20 and root -1.01) compared to durum (leaf 3.35 and root 3.66) samples. This marker was thus scored present in durum samples and absent in polonicum samples (as in Cy3 analysis). Both polonicum and durum markers are shown in figure 5.20 for Cy3 and Cy5 scores for leaf and root replicated samples. It can be seen that the scores for Cy3 and for Cy5 are similar, being on average only 0.23 sample ratio median score values different, with the exception of marker 800904300002_C_7 in durum samples that show a difference of 0.96 sample ratio median score values between Cy3 and Cy5 analyses.

Similarly, analysis can be performed for all samples from seedling, control, incremental and shock polonicum and durum samples. This includes Cy3 and Cy5 targets as well as targets generated using different restriction enzymes during the complexity reduction protocol, that is, PstI/Taq$^{\alpha}$I, PflMI/MseI and PstI/MseI. These varying conditions are explored in chapter 3 and the cultivar polymorphisms are further analysed here.

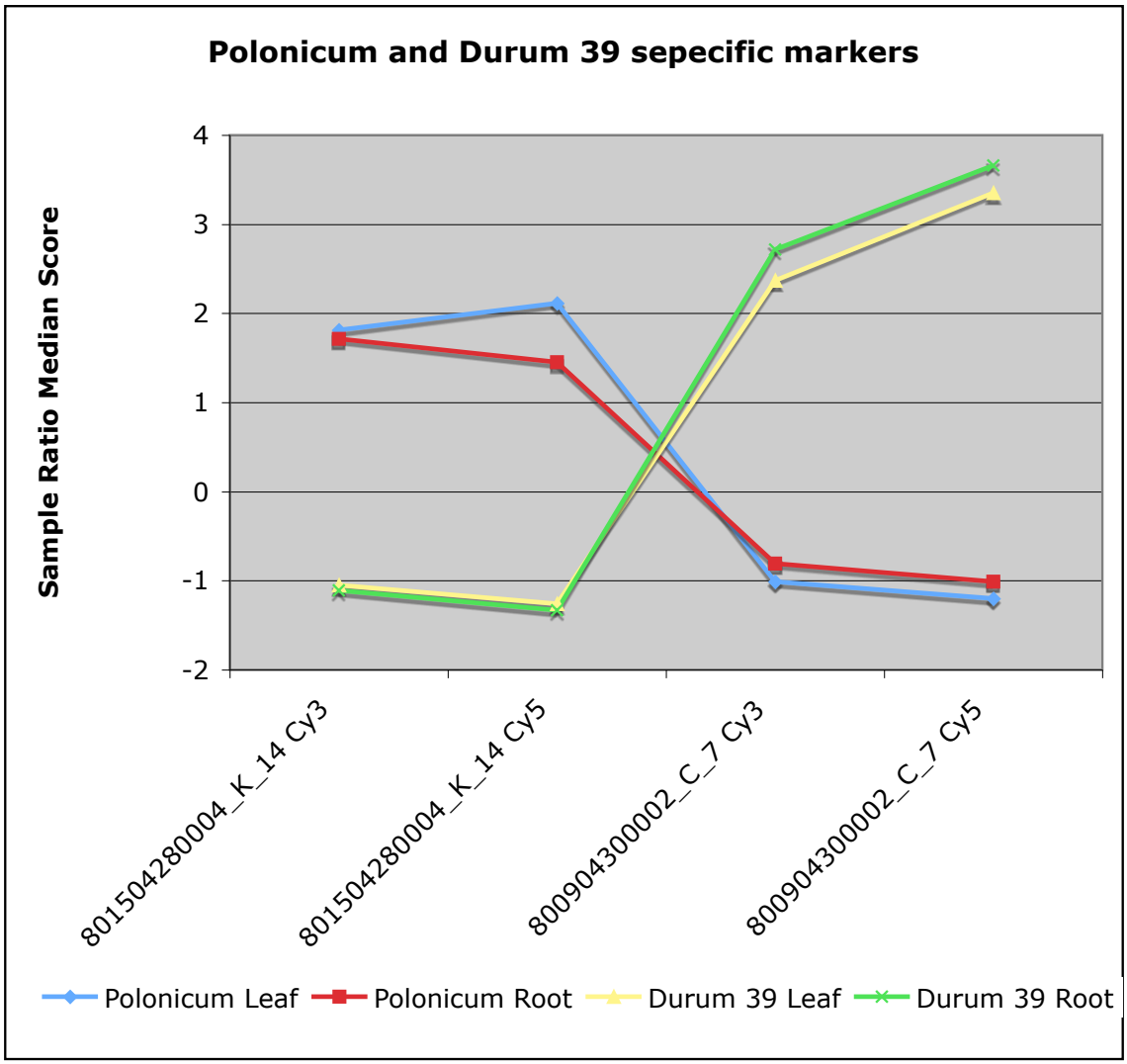


Figure 5.20: Polonicum-like and durum-like markers scored for sample ratio median values from Cy3 and Cy5 analysis. There is aclear seperation of markers between Polonicumand Durum 39 samples in both root and leaf samples.

5.6.4 Conclusions

From the analyses performed and results shown, DArT is a powerful technique for discriminating between cultivars samples. Replicates are scored reproducibility over target and printed spots and can be labeled with varying dyes. DArT identified markers that can distinguish between durum line 39 (139) and polonicum (55) DNA samples and cultivars. These potential markers can be used for genotyping these samples for a durum line 39-like or a polonicum-like result, and can be linked to traits of interest for plant breeding experiments.

5.7 Bulk Segregant Analysis

5.7.1 Aims

The aim of this experiment is to perform and analyse a breeding experiment between Wollaroi, a cultivated durum wheat and *ssp. carthlicum*, a wild durum line, with 99 progeny was performed using bulk segregant analysis (BSA). BSA groups plants according to phenotypic expression of a trait and tests the aim is to measure the allele frequency between the population bulks to determine if salt treatment has any effect.[196] Wollaroi and carthlicum plants were crossed and 99 progeny produced that were grown in a salt enriched environment, introduced via irrigation before DNA extraction. Samples were phenotyped by ranking them according to mean soil plant analytical development (SPAD) chlorophyll readings from 3 leaf samples. SPAD readings provide an indication of the chlorophyll content of plant leaves without damaging the plant. This provides an indication of photosynthesis and the biological activity of the leaf. Samples with a low SPAD reading show reduced chlorophyll content, suggesting reduced photosynthesis activity and a lower tolerance to the salt enriched environmental conditions. Samples with higher SPAD readings indicate greater chlorophyll content, suggesting higher photosynthesis activity and a higher tolerance to an enriched salt environment.

5.7.2 Specific Methods

5.7.2.1 BSA Analysis settings

Bulk segregant analysis was performed using BSArT 1.3, an in-house software package that compares bulks, random bulks and parental samples. Analysis filtered

spot intensity and target ratio median data, removing spots with a with a CV of signal pixels greater than 100% in both channels, spots with 20% or greater pixels saturated in either channel, spots with a signal-to-noise ratio of less than 3 in both channels, spots with a SD of background pixels 10 times the slide median or greater and spots smaller than 30% of the slide median. Clone quality was performed, with 3% of clones with the largest across-replicate-spot standard deviation of logRatio scores removed, as well as clones with 25% or less replicate spots remaining removed. Analysis of between-slide normalisation used the central 90% of spot logRatio values for normalisation of means and scaling of standard deviations. Polymorphic clones were selected with a logRatio difference between parents of greater than 0.59 for array 1, 0.43 for array 2, 0.68 for array 3 and 0.70 for array 4, using P values of less than 0.2 in the self comparison. Clones were rejected that had a logRatio difference of greater than 0.29 for array 1, 0.29 for array 2, 0.39 for array 3 and 0.35 for array 4, with a self comparison of P less than 0.2. Rejected clones beyond +/- 18% in array 1, +/- 24% in array 2, +/- 26% in array 3 or +/-19% in array 4 of the average relative abundance in the self comparison using P of less than 0.2.

5.7.2.2 Sample production

DNA was extracted and a digestion/ligation reaction prepared using the standard DArT protocol. Samples were PCR amplified in duplicate from the single digestion/ ligation reaction, with each duplicated 50 µl PCR reaction mixed and then separated back into 2 reactions. Replicates were used for BSA experimentation and the duplicate for Recombinant Inbred Line analysis. From each of the bulks that were produced, 50 µl of the PCR mixture from each sample (progeny line) were mixed, so that 15 samples comprised each bulk. This was performed for bulk A, bulk B, random bulk A and random Bulk B. From each of the bulks, multiple replicate aliquots of 50 µl were taken. Multiple PCR amplifications were performed using each parental sample, using the same digestion / ligation protocol, with the PCR's mixed and divided into 12 aliquots of 50 µl each. A mixture of parent A and parent B was also prepared by mixing equal quantities of each parent, then dividing the mix into 50 µl aliquots. Bulks and parental samples were precipitated, washed and labeled as described in table 5.15. Samples were then hybridised to each of the 4 arrays.

Table 5.15: BSA Target preparation

Slide	Cy3 Target	Cy5 Target
1	Bulk A replicate 1	Bulk B replicate 1
2	Bulk A replicate 2	Bulk B replicate 2
3	Bulk A replicate 3	Bulk B replicate 3
4	Bulk B replicate 1	Bulk A replicate 1
5	Bulk B replicate 2	Bulk A replicate 2
6	Bulk B replicate 3	Bulk A replicate 3
7	Parent A replicate 1	Parent B replicate 1
8	Parent A replicate 2	Parent B replicate 2
9	Parent A replicate 3	Parent B replicate 3
10	Parent B replicate 1	Parent A replicate 1
11	Parent B replicate 2	Parent A replicate 2
12	Parent B replicate 3	Parent A replicate 3
13	Parent AB replicate 1	Parent AB replicate 1
14	Parent AB replicate 2	Parent AB replicate 2
15	Parent AB replicate 3	Parent AB replicate 3
16	Parent AB replicate 4	Parent AB replicate 4
17	Parent AB replicate 5	Parent AB replicate 5
18	Parent AB replicate 6	Parent AB replicate 6
19	Random A replicate 1	Random B replicate 1
20	Random A replicate 2	Random B replicate 2
21	Random A replicate 3	Random B replicate 3
22	Random B replicate 1	Random A replicate 1
23	Random B replicate 2	Random A replicate 2
24	Random B replicate 3	Random A replicate 3

5.7.3 Results

5.7.3.1 SPAD Analysis

Samples were grouped or 'bulked' according to SPAD readings, taking into account the standard error, with samples greater than 7.0 rejected (table 5.16). Based on initial readings, 15 samples with low SPAD readings under 18 were bulked to represent the salt in-tolerant phenotype, termed 'bulk A' (table 5.15). These samples were also selected to have a low standard error and a similar SPAD reading in the duplicated samples. A further 15 samples with SPAD readings above 30 termed 'bulk B' were selected to represent the salt tolerant phenotype (table 5.16). Random samples were selected via a random number generator, excluding samples in bulks A and B and the parents, with 15 samples bulked as 'random A' and 15 samples bulked as 'random B' (table 5.16).

Parent A, *ssp. carthlicum* was phenotyped as salt tolerant, scoring 31.0 in mean SPAD readings with a standard error of 2.2. Parent B, Wollaroi, was phenotype as salt in-tolerant, scoring 13.3 in mean SPAD readings. Mean SPAD data is shown in table 5.14 and graphed in figure 5.21.

Table 5.16: Ranked BSA phenotypic SPAD data

Line	SPAD Mean*	Standard Error
136	4.6	1.5
103	9.1	3.2
117	9.3	2.0
89	9.4	3.9
110	9.4	5.5
155	10.6	5.4
118	11.7	6.3
Wollaroi	13.3	1.9
95	13.4	5.3
92	13.6	7.8
37	13.9	2.2
6	14.1	2.7
153	14.2	3.0
124	14.3	9.8
5	15.0	1.6
128	15.3	6.2
60	16.0	2.1
147	16.2	3.2
24	16.3	2.0
20	16.5	2.2
84	16.5	7.2
132	16.7	3.1
57	17.2	2.5
16	17.7	2.8
74	17.8	7.8
146	17.8	2.5
22	18.0	4.2
68	18.3	4.3
14	18.5	2.4
101	18.7	3.7
52	19.5	2.7
69	19.5	1.1
114	19.7	4.1
1	20.0	3.7
94	20.0	5.5
88	20.3	4.8
125	21.0	4.9
45	21.5	2.0

Line	SPAD Mean*	Standard Error
120	24.9	4.4
19	25.1	1.7
97	25.9	4.2
62	26.9	3.3
47	27.3	3.0
102	27.4	8.4
134	27.4	4.7
40	27.5	3.0
43	27.8	3.6
82	28.0	4.6
115	28.3	5.2
142	28.6	5.7
3	28.7	3.6
140	28.9	7.5
154	28.9	5.8
151	29.0	4.1
54	29.7	3.2
145	29.8	5.4
27	30.0	5.2
9	30.3	2.5
Carthlicum	31.0	2.2
13	31.1	3.1
112	31.1	6.1
123	31.5	4.5
38	31.7	7.8
86	32.1	4.0
121	33.2	6.7
149	33.3	4.7
87	33.4	7.0
99	33.5	3.2
129	34.0	6.9
131	34.1	3.7
135	35.1	2.1
130	35.3	4.6
109	35.4	3.5
127	35.7	4.9
148	36.0	2.7
91	36.5	1.6

Line	SPAD Mean*	Standard Error
23	21.6	4.1
56	21.8	3.8
21	21.9	2.5
32	22.0	1.9
144	22.1	6.3
108	22.4	4.2
80	22.5	3.9
10	22.7	2.8
126	23.0	4.2
138	23.1	4.2
104	23.3	7.9
42	24.4	3.8
85	24.6	4.8

Line	SPAD Mean*	Standard Error
50	36.6	4.5
100	36.8	3.4
106	37.4	3.0
113	37.4	7.5
75	37.5	6.1
64	37.8	1.2
98	38.3	2.8
150	38.4	7.0
141	39.0	5.5
17	39.5	0.9
152	40.5	1.3
73	43.1	0.3

* SPAD mean readings of 3 leaf samples

Green - Salt in-tolerant phenotype, low SPAD readings

Red - Salt tolerant phenotype, high SPAD readings

Yellow - Random bulk A

Blue - Random bulk B

Purple - Parental line

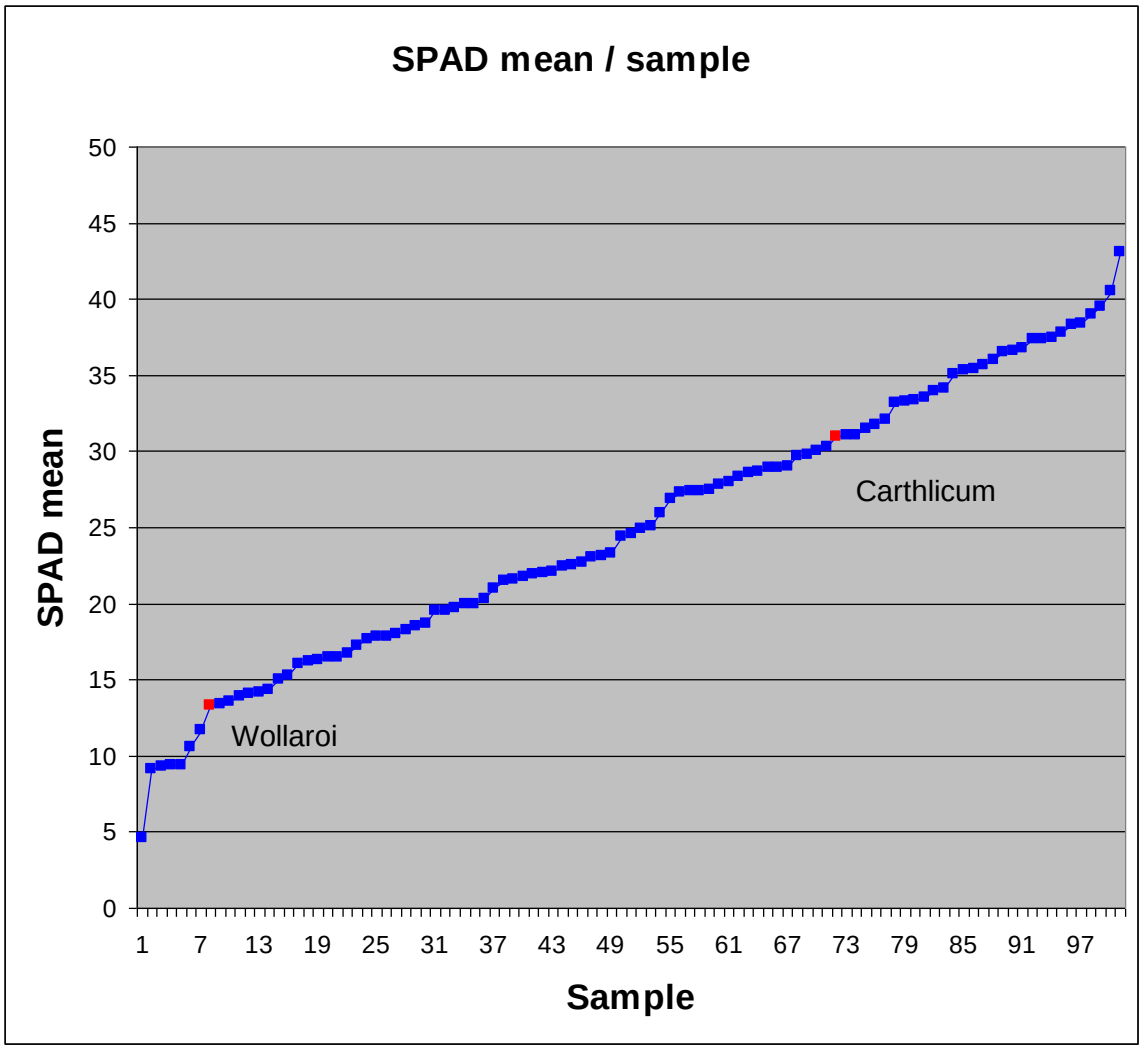


Figure 5.21: Breeding experiment between Wollaroi and Carthlicum, showing the SPAD scores for both prarents and 99 progeny lines. There is a clear seperation of Wollaroi and Carthlicum samples indicating low and high salt tolerance.

5.7.3.2 Bulk Ratios

DArT analysis of the BSA experiment identified 1,153 high quality polymorphic markers, with 312 markers identified on array 1, 362 on array 2, 109 on array 3 and 370 markers identified Array 4. These markers were compared and the LD scores (%) for bulk A compared to bulk B, parent A+B and parent A+B controls, random bulk A and random bulk B. Figure 5.22 shows the comparison of these three bulks, with the three curves showing the relative abundance (%) against the markers in decreasing order.

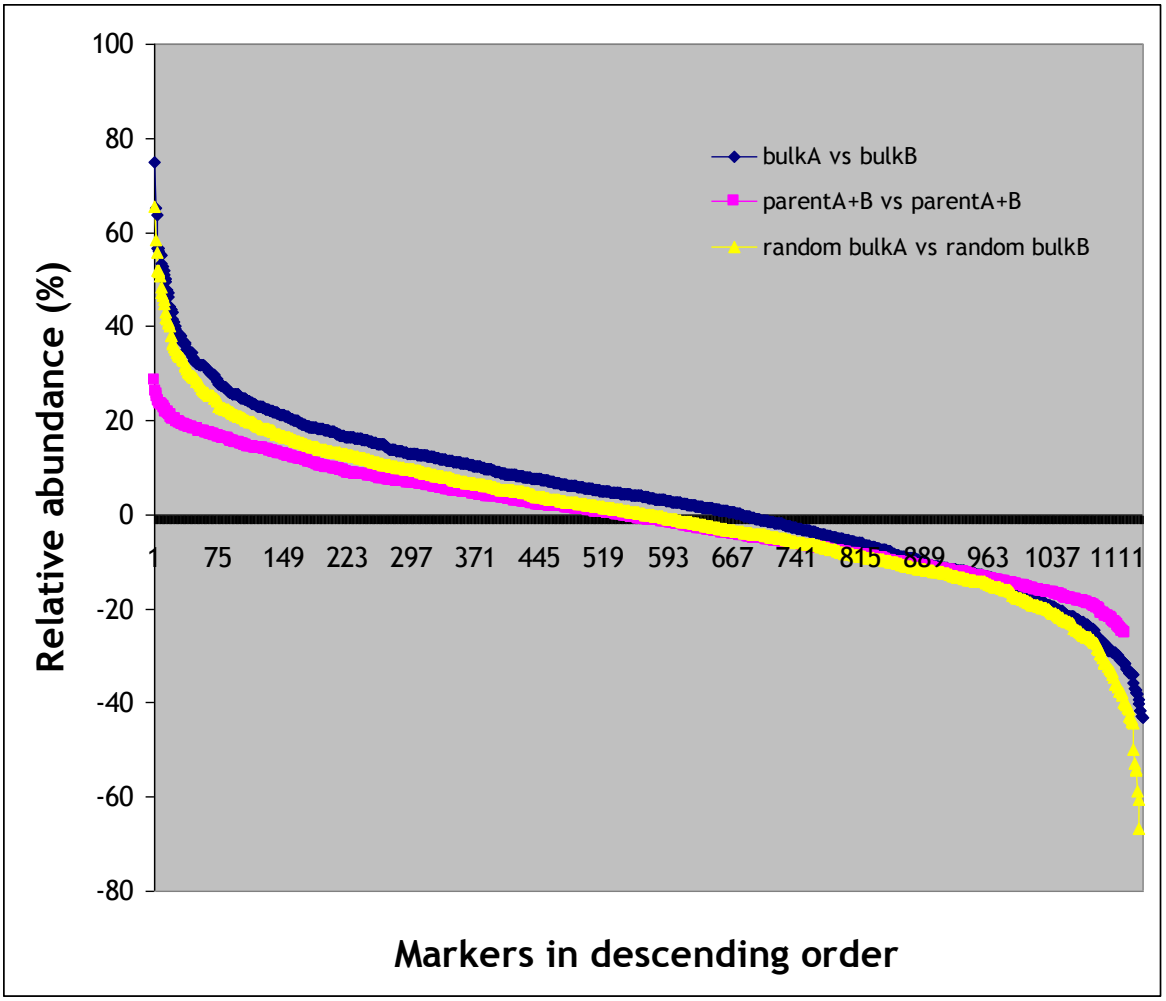


Figure 5.22: BSA bulk sample ratios showing the relative abundance of bulked samples, parent A+B samples and random A and B bulked samples. There is minimal divergence of samples indicating minimal polymorphic divergence of bulked samples.

5.7.4 Conclusion

From the comparison of 'bulk A versus bulk B' and 'random bulk A versus random bulk B' it can be seen that theis is no significant difference in curves, that is tere is no significant deviation in allele frequencies (figure 5.22). There is no obvious segregation of salt in-tollerant or tolerant groups, indicating that the SPAD readings may not have been a good indicator of salt tolerance or that there is minimal tolerance difference's between Wollaroi and carthlicum cultivars.

5.8 Recombinant Inbred Line analysis

5.8.1 Aims

A cross between Wollaroi (W) and *T. turgidum L. ssp. carthlicum* (14) was performed and 96 progeny self-fertilised until the 6th filial generation (F6). All plants were grown at CSIRO Plant Industries by Dr Rana Munns and Dr Richard James in a climate controlled glass house. Leaf material was harvested from the first leaf of the second tiller where 5cm of the leaf tip was removed from each genotype in quadruplicate. The aim is to identify polymorphisms between the 94 inbred Wollaroi and carthlicum samples.

5.8.2 Specific Methods

DNA from leaf samples was extracted from 96 F6 recombinant inbred lines generated from a cross between Wollaroi and carthlicum cultivars. Samples were PCR amplified using the DArT protocol to generate targets for microarray hybridisation.

5.5.3 Results

5.8.3.1 Analysis of Wollaroi, carthlicum and 94 progeny

DArTsoft analysis of the 96 recombinant inbred lines identified 1,064 high quality polymorphic markers. Of these, 836 markers were scored polymorphic between the 2 parents with 418 markers scoring present in Wollaroi and 418 markers in carthlicum. Of the remaining markers, 117 markers were scored the same in both cultivars, 71 markers scored present and 46 markers scored absent but scored polymorphic across progeny. This phenomenon occurs when one (or both) parents contain heterogeneous germplasm, in that, not all alleles within the plant are the same on each replicate chromosome, leading to polymorphic behavior of markers scored the same in both parents. A further 111 markers were scored with some degree of discrepancy between spot replicates, 31 markers in Wollaroi, 73 markers in carthlicum and 7 markers in both cultivars.

The distribution of scores across the 94 progeny for the 836 polymorphic markers between the parents is varied, with markers grouping according to their bimodal scores. Marker 801504280004_B_5 segregates 50% present and 50% absent, with 48 samples scored absent (0) including Wollaroi and 48 samples scored present (1) including carthlicum. This marker has been sequence previously and is located on wheat chromosomes 6A and 6B, designated wPt-7599. These markers all score a PIC value of 0.5, that is, 50% of the samples are scored present and 50% scored absent. There are 33 markers identified that segregate in this fashion.

A PIC value of 0.499945 indicates 1 sample scoring discrepancy that is not scored either 1 or 0, but represented with an X in the data tables. This score did not fit with the bimodal distribution clusters of 1 or 0, usually caused by experimental conditions such as dust or debris or a spot that was not printed on the slide. As the array contained duplicate spots, one spot may be been scored one way and the other the opposite also scored with an X. Analysis identified a further 8 markers that scored a PIC value of 0.499945 and 16 markers with two scoring discrepancies and a PIC value of 0.499942. Overall, 823 markers (98.5%) were scored with a PIC value of 0.4 or greater. These markers follow Mendelian segregation pattern of 1:1 ratios from parents to offspring. Marker 801504280001_O_11, located on chromosome 1A, designated wPt-3698, has a PIC value of 0.262716, where 76 samples including carthlicum were scored absent, 14 samples including Wollaroi scored present and 6 markers scored with a discrepancy (X). The ratios of absent to present is almost 80:15, not taking into account the 6 samples that were not scored perfectly. This is very close to the Mendelian ratio of 4:1, with 75% of markers segregating with one parent and 25% with the other parent.

At the other extreme, 3 markers were identified with a PIC value of 0.040799, having 2 samples scored present and 94 samples scored absent. These markers, 801504280004_K_22, 801504280001_I_13 and 802906152015_H_24 all are scored present only in Wollaroi and sample leaf extract 113. These markers do not follow Mendelian segregation patterns, and may be attributed to a methylation polymorphism detected by PstI. Markers 801504280004_K_22 and 801504280001_I_13 have both been sequenced previously and are located on wheat chromosome 1A, with markers names wPt-8644 and wPt-4709. Figure 5.23

shows marker 801504280004_B_5, representing 98.5% of all markers with a 1:1 Mendelian distribution of markers scored present or absent.

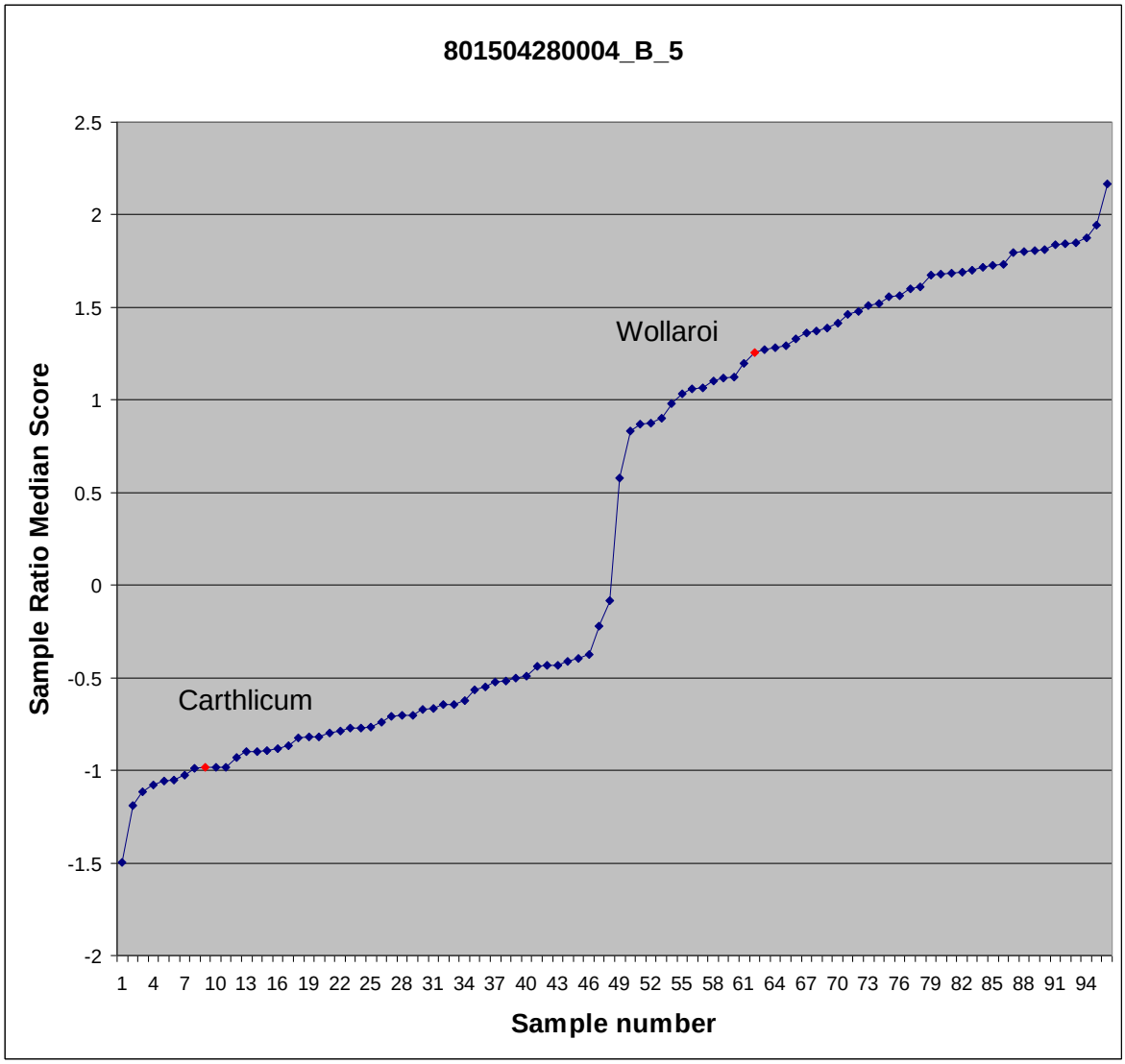


Figure 5.23: 1:1 Mendealian 1:1 distributon of samples for marker 801504280004_B_5 showing a 50% segregation between progeny samples.

5.8.3.2 Analysis of libraries within the array

The 96 sample Wollaroi and carthlicum experiment was hybridised to a PstI and Taq$^{\alpha}$I array comprising 16x 384-well microtitre plates from 4 libraries. Comparisons of the number of clones identified in each library from the 836 high quality markers in the analysis as well as the percentage of high quality polymorphic markers found within each library is shown in table 5.17. As expected, the rearrayed libraries generated from the 4 arrays analysed in section 5.3 (a) and the DArT durum rearray (d) show the highest proportion of markers identified, as previous analyses pre-selected these clones for use. The discovery durum and polonicum (b) and DArT 45 cultivar wheat array 2.3 (c) libraries that contain random clones both show a lower polymorphism rate, as to be expected from non-selected markers.

Table 5.17: Summary of Wollaroi and carthlicum markers over arrays 1-4

Library Name	Number of plates / markers	Number of high quality polymorphic markers	Percentage of high quality polymorphic markers	Percentage of high quality polymorphic markers within the library
(a) Durum rearray (section 5.3)	1 / 384	107	12.80%	27.86%
(b) Durum and polonicum	4 / 1536	121	14.47%	7.88%
(c) DArT Wheat 2.3	5 / 1920	299	35.77%	15.57%
(d) DArT Durum rearray	5 / 1920	537	64.23%	27.97%

5.8.4 Conclusions

From the data presrnted, it can be seen that most polymorphic markers segregate in a Mendelian 1:1 ratio. This is shown in figure 5.23 with the parental samples present in each polymorphic cluster. Suspected methylation polymoprhis can also be identified where 2 markers were scored present and 94 scored absent. Using selected arrays will also affect the percentage of potential high quality polymorphic markers from a library, as markers that are polymorphic are selected for.

5.9 Genetic linkage map

5.9.1 Aims

A genetic linkage map is a map produced from an experimental population showing the position of known genetic markers relative to each other in terms of recombination frequency, rather than as specific physical distance along each chromosome. A genetic map is produced using the frequencies of recombination between markers during crossover of homologous chromosomes. The greater the frequency of recombination (segregation) between two genetic markers, the farther apart they are assumed to be. Conversely, the lower the frequency of recombination between the markers, the smaller the physical distance between them.[197] The aim of this experiment is to use a recombinant inbred cross between Wollaroi and carthlicum and analyse the 94 F6 progeny. The relative positions of each markers with then be mapped.

5.9.2 Specific Methods

Targets were produced from the 96 recombinant inbred lines (RIL) for the F6 cross between Wollaroi and carthlicum. Data was analysed and quality checked, with 7 samples removed. All 607 markers were analysed and individual maps constructed for all RIL populations using EasyMap. EasyMap is a program developed at Diversity Arrays P/L for high-throughput mapping of DH and RIL populations. EasyMap automates the distribution of markers into linkage groups, the ordering of markers within linkage groups (based on the RECORD algorithm), the detection of potential genotyping errors, the re-optimisation of marker orders after replacing potential errors with unknown genotype calls and the estimation of map distances. Linkage groups were then assigned to chromosome / linkage groups based on a comparison across populations and the existing chromosome assignments of markers printed on the array. EasyMap function ‘Kosambi’ and linkage evaluation setting ‘Self RI’ were used.

At the current level of marker coverage, most chromosomes are represented by more than a single linkage group. The order and orientation of linkage groups within chromosomes was established (were possible) by comparing each linkage group against the Synthetic/Opata and Cranbrook/Halberd linkage groups for which the

orientation was known. Due to the lack of sufficient numbers of marker bridging among populations this was not always possible, so some degree of uncertainty about linkage group order and orientation remains, particularly for D-genome chromosomes. In addition, there was a number of loci that were excluded because they were not sufficiently linked to any other linkage group. The ordering and orienting of linkage group also allowed the identification of a number of multi-locus markers that map to two loci within a single chromosome. Not all of these multi-locus markers may have been recognised at this stage.

5.9.3 Results

5.9.3.1 Analysis of Carthlicum vs. Wollaroi with 85 progeny

866 markers were distributed into 61 linkage groups based on scored A or B on which is more-like a certain parent, Wollaroi or carthlicum. The P vaule was used at different thresholds to estimate the number of groups and un-linked markers that remain. This is shown in table 5.18, with P=0.0001 resulting in 47 linkage groups with 15 un-linked markers. The markers assigned to these groups are shown in table 5.19 and these are also sorted by chromosome A and B number in table 5.20.

Table 5.18: Distribution of groups and markers at different thresholds

Linkage Criterion:	Number of groups	Number of unlinked markers
P = 0.01	21	11
P = 0.001	42	8
P = 0.0001	47	15
P = 0.00001	58	17

Table 5.19: Linkage groups

Group	# markers	Chromosome markers (number)
1	18	6B (4)
2	26	5B (9) 2A (1)
3	36	6A (11) 6B (2) 6A/6B (4) 1A (1) 1B3B (1)
4	12	4B (4)
5	26	3A (6) 1A (1) 1A3/A (1)
6	16	4B (2) 2A (1)
7	15	3B (3)
8	2	1A (1)
9	33	3B (11) 3D (2)
10	49	6B(13) 6A/6B (1)
11	14	3B (4) 3A (1)
12	30	1B (5)
13	9	7B (5)
14	2	3A (1)
15	15	2B (7)
16	4	2D (2)
17	4	1A (3)
18	8	2A (1) 2A/2D (1)
19	30	7B (8) 3B/7D (1) 3D (2)
20	43	4A (15)
21	10	5A (3)
22	24	5B (9) 5B/5D (1) 7A (1)
23	9	1A (2)
24	2	5A (1)
25	11	7A (1)
26	7	1B (3)
27	10	2A (3)
28	17	2B (5)
29	13	1A (1) 6A (1)
30	8	7A (6)
31	6	6A/7A (1)
32	3	2A (2)
33	3	1B (2)
34	3	-
35	3	7A (2)
36	3	-
37	3	-
38	5	2B (1)
39	3	-
40	2	7B (2)
41	6	-
42	4	5A (1)
43	2	-
44	3	4B (3)
45	2	1B (2)
46	4	1B (1)
47	2	3B (1)
Unlinked	15	1A/1B (1) 6A (1) 6B (1)

Table 5.20: Linkage groups

Chromosome Genome A	Linkage groups	Chromosome Genome B	Linkage groups
1A	23, 29	1B	12, 26
2A	18, 27, 32	2B	15, 28
3A	5,	3B	7, 9, 11
4A	20	4B	4, 6, 44
5A	21, 42	5B	2, 22
6A	3, 29, 31	6B	1, 10
7A	25, 30, 31	7B	13, 19

5.9.3.2 Linkage group maps

Linkage groups maps are generated using MapChart software described by Voorrips (2002)[198] and available at http://www.biometris.wur.nl/uk/Software/MapChart/. Figure 5.24(a)-(g) shows the results of each linkage map for each of the 14 chromosomes (1-7 A and 1-7 B) with larger images shown in 7.4 Appendix D, Linkage maps.

5.9.4 Conclusion

From the linkage maps, it can be seen that the 866 markers can be distributed based on linkage groups and known genetic map positions over the 7A and 7B wheat chromosomes. Some of the markets do not map and will need further analysis to determine their location. Sequencing these markers will allow them to be compared to other linkage maps that have been generated using DArT and other techniques. Traits of interest can also be associated with individual or groups of markers for use in plant breeding experiments.

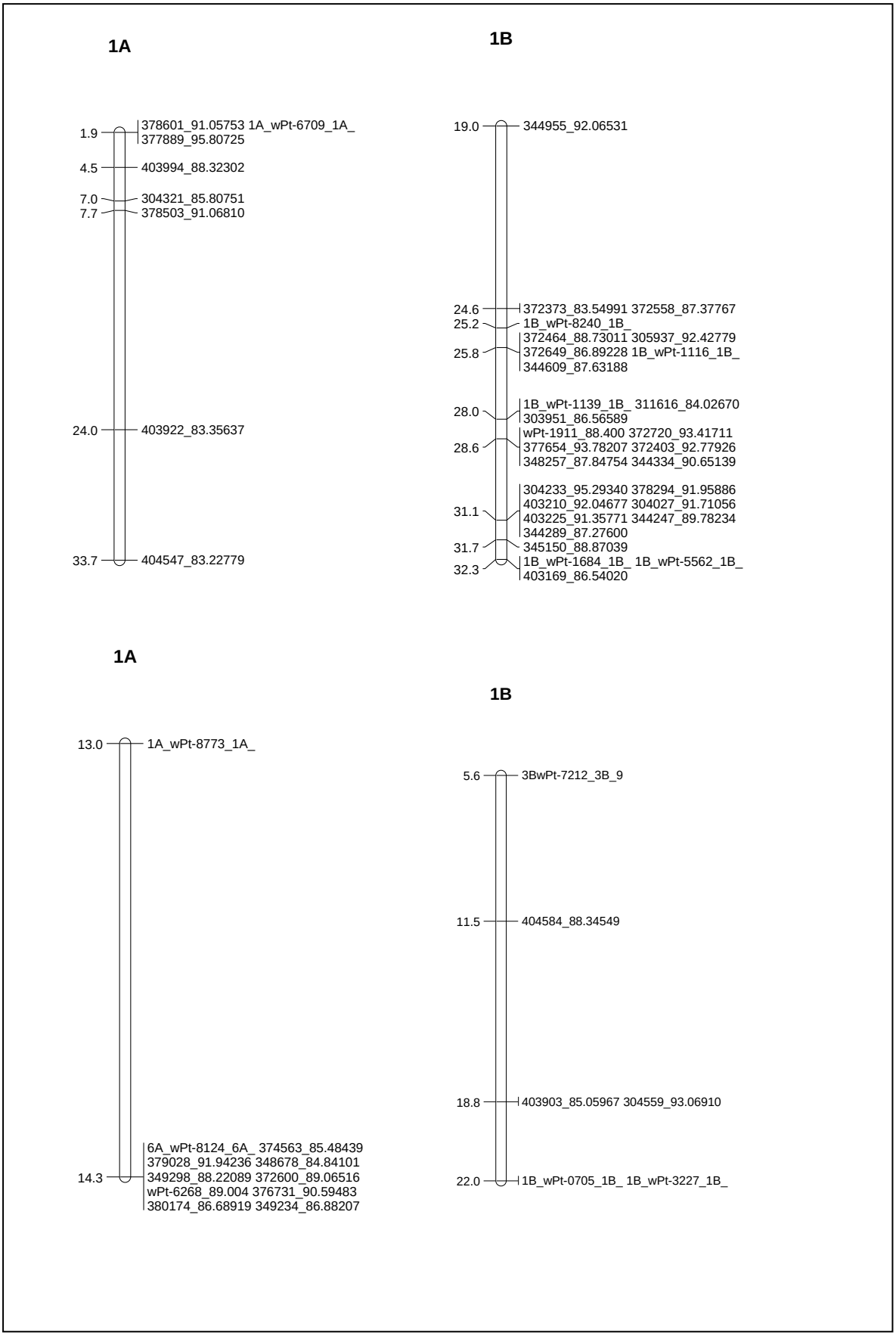


Figure 5.24(a): Linkage map for chromosomes 1A and 1B

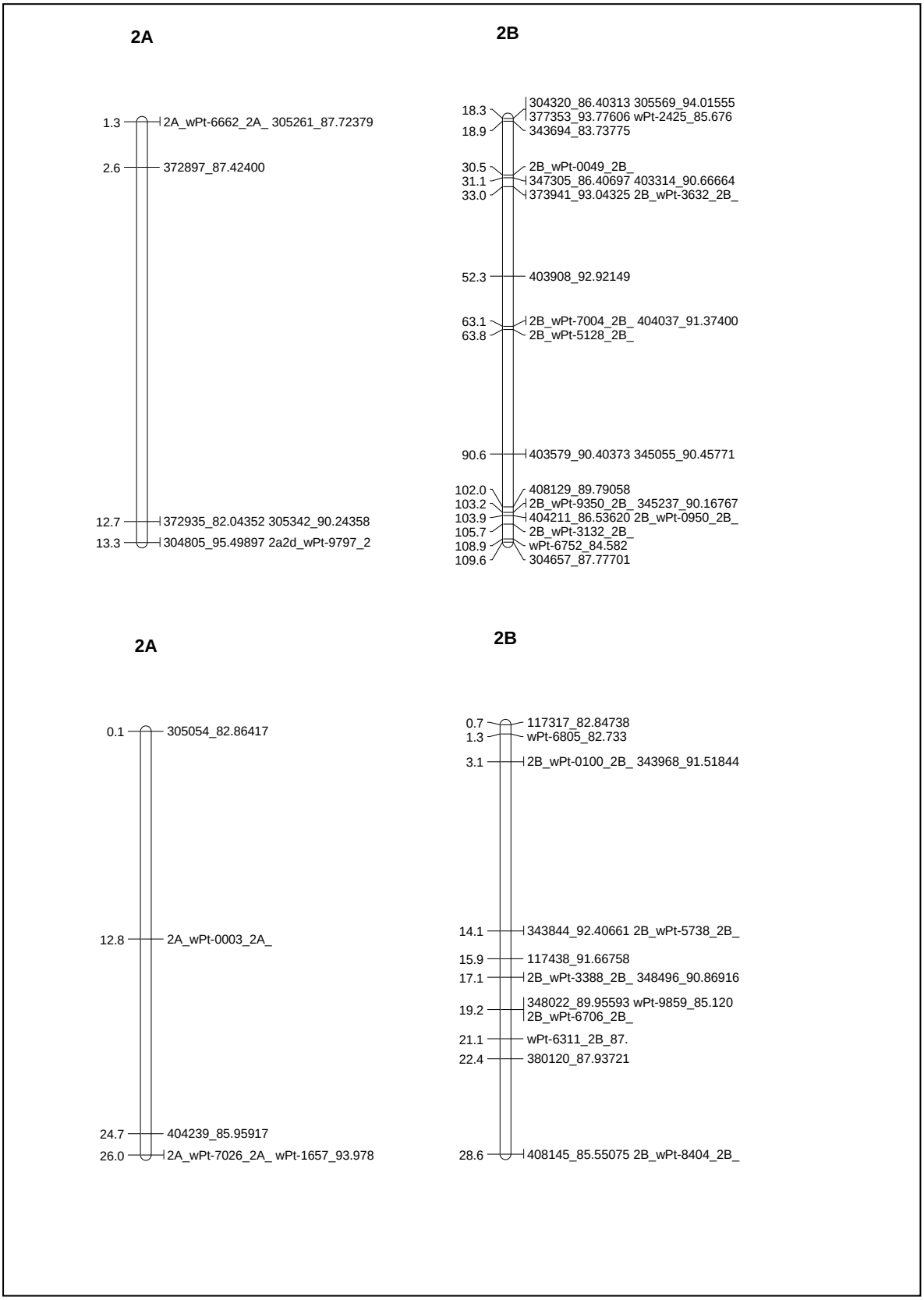


Figure 5.24(b): Linkage map for chromosomes 2A and 2B

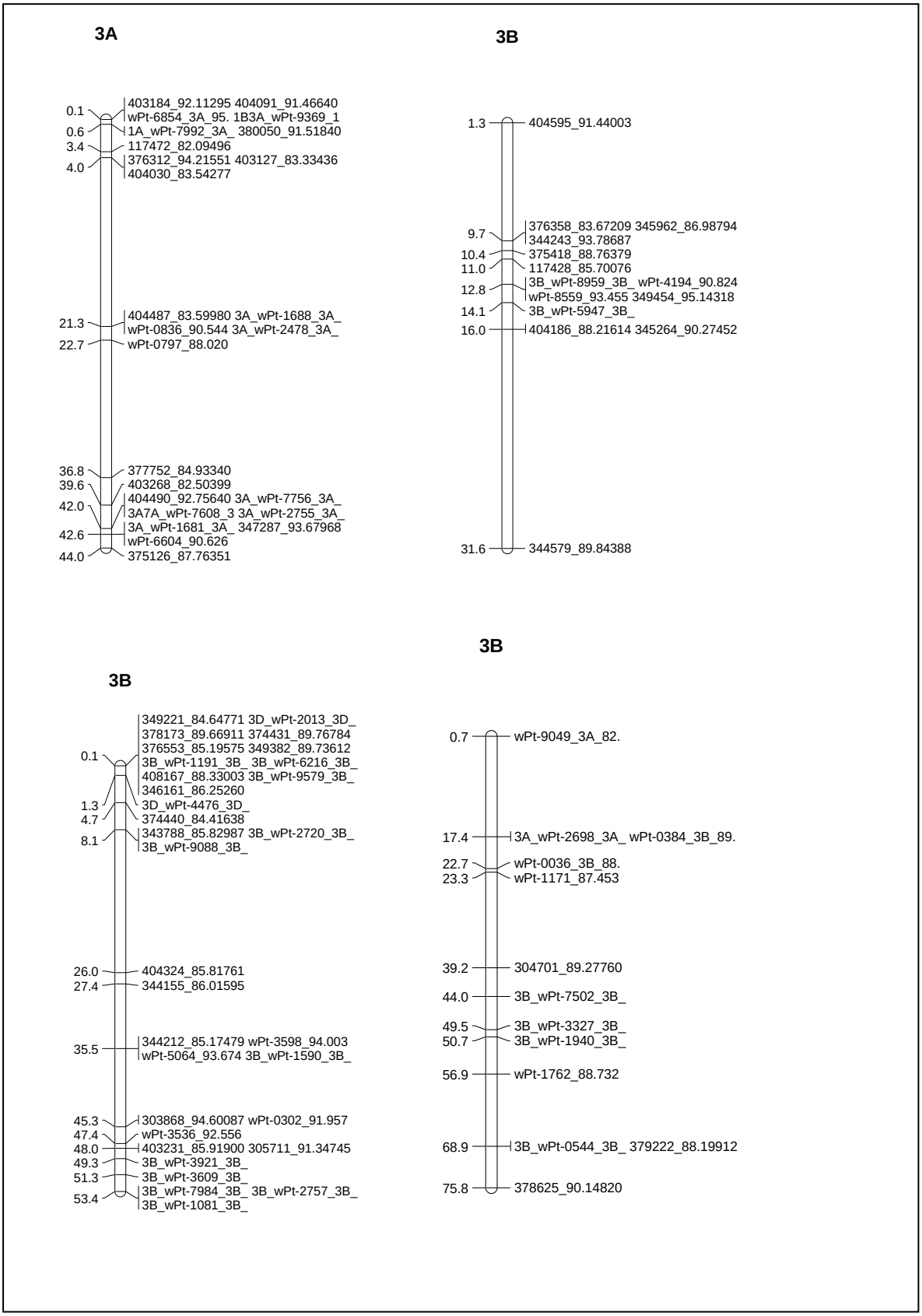


Figure 5.24(c): Linkage map for chromosomes 3A and 3B

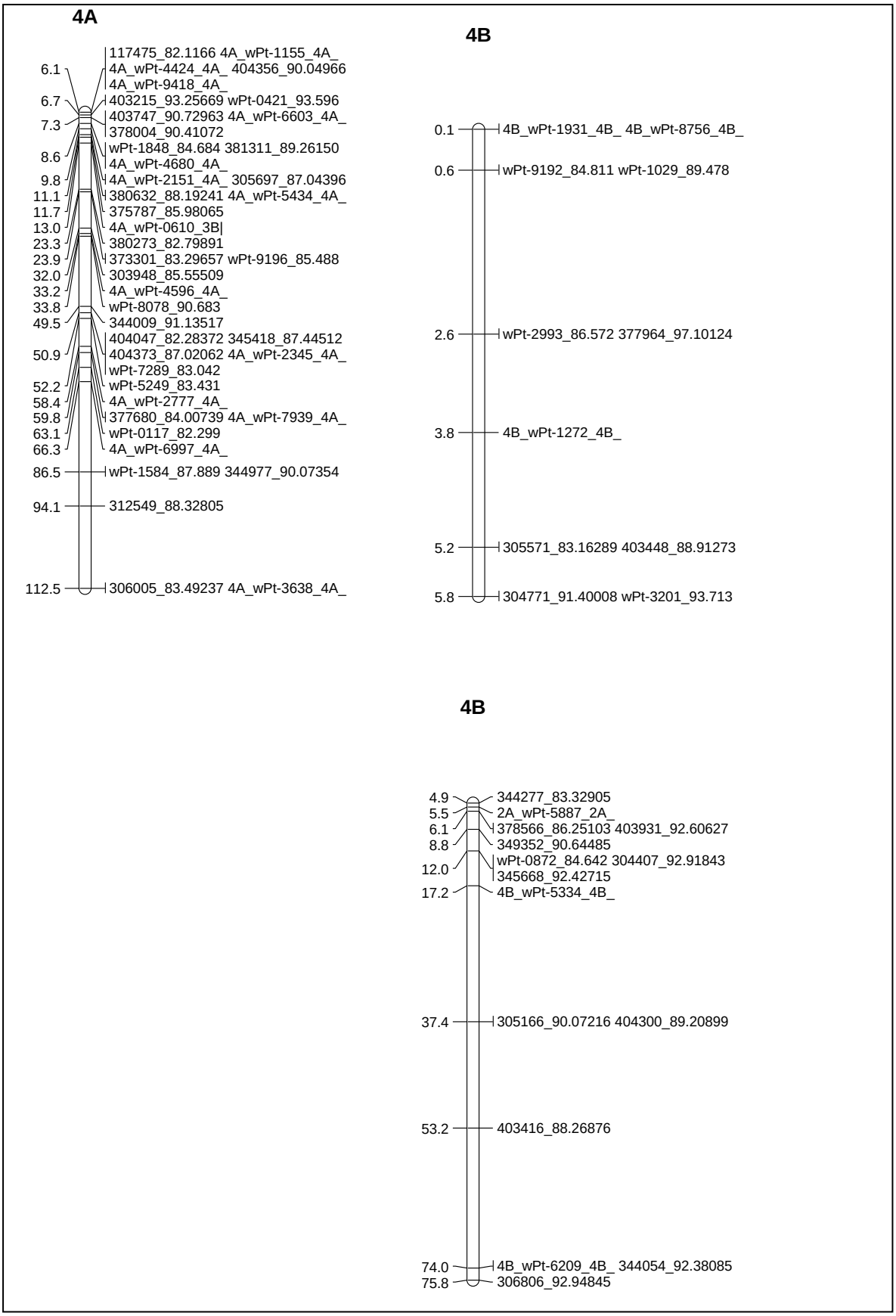


Figure 5.24(d): Linkage map for chromosomes 4A and 4B

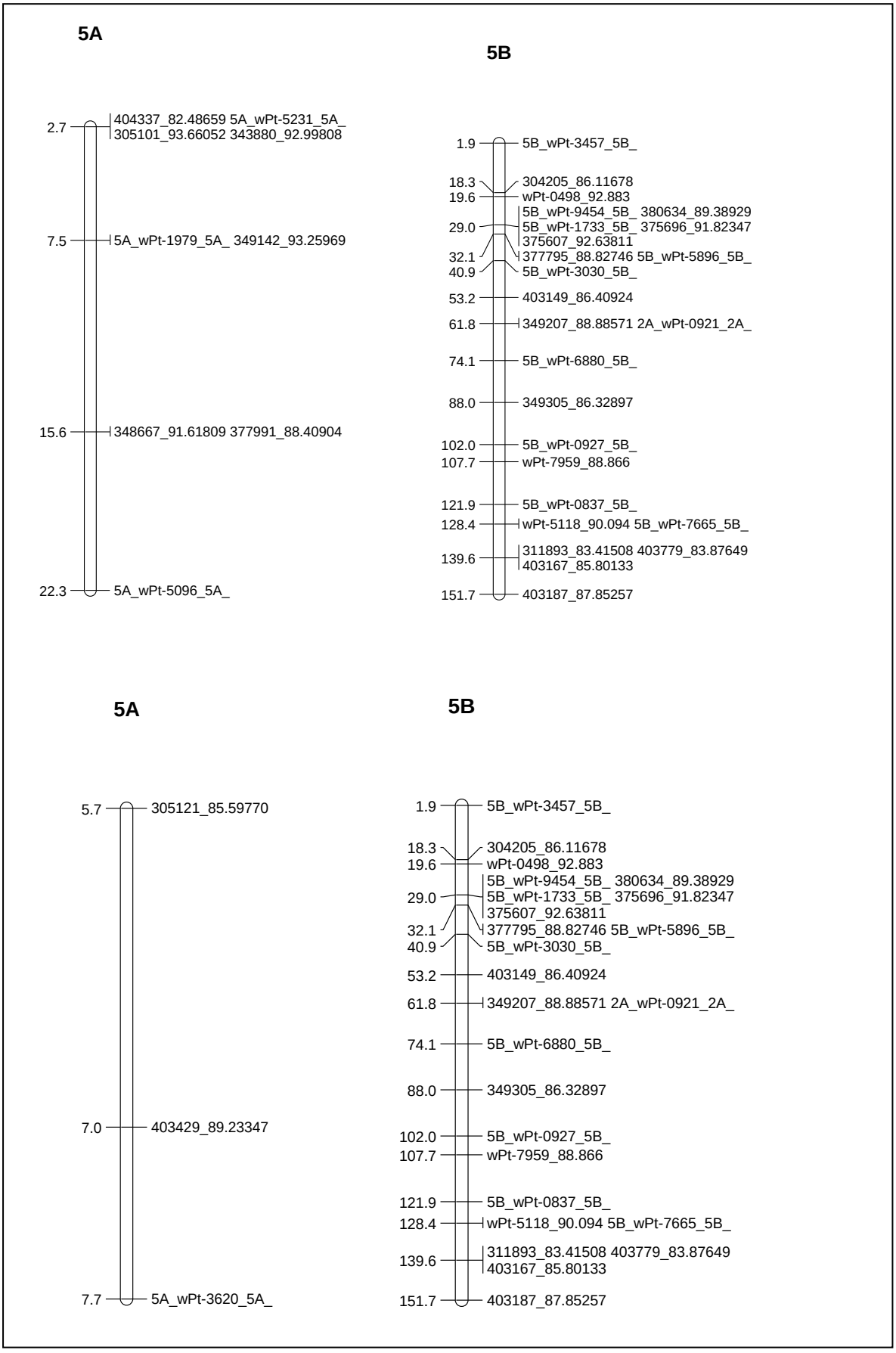


Figure 5.24(e): Linkage map for chromosomes 5A and 5B

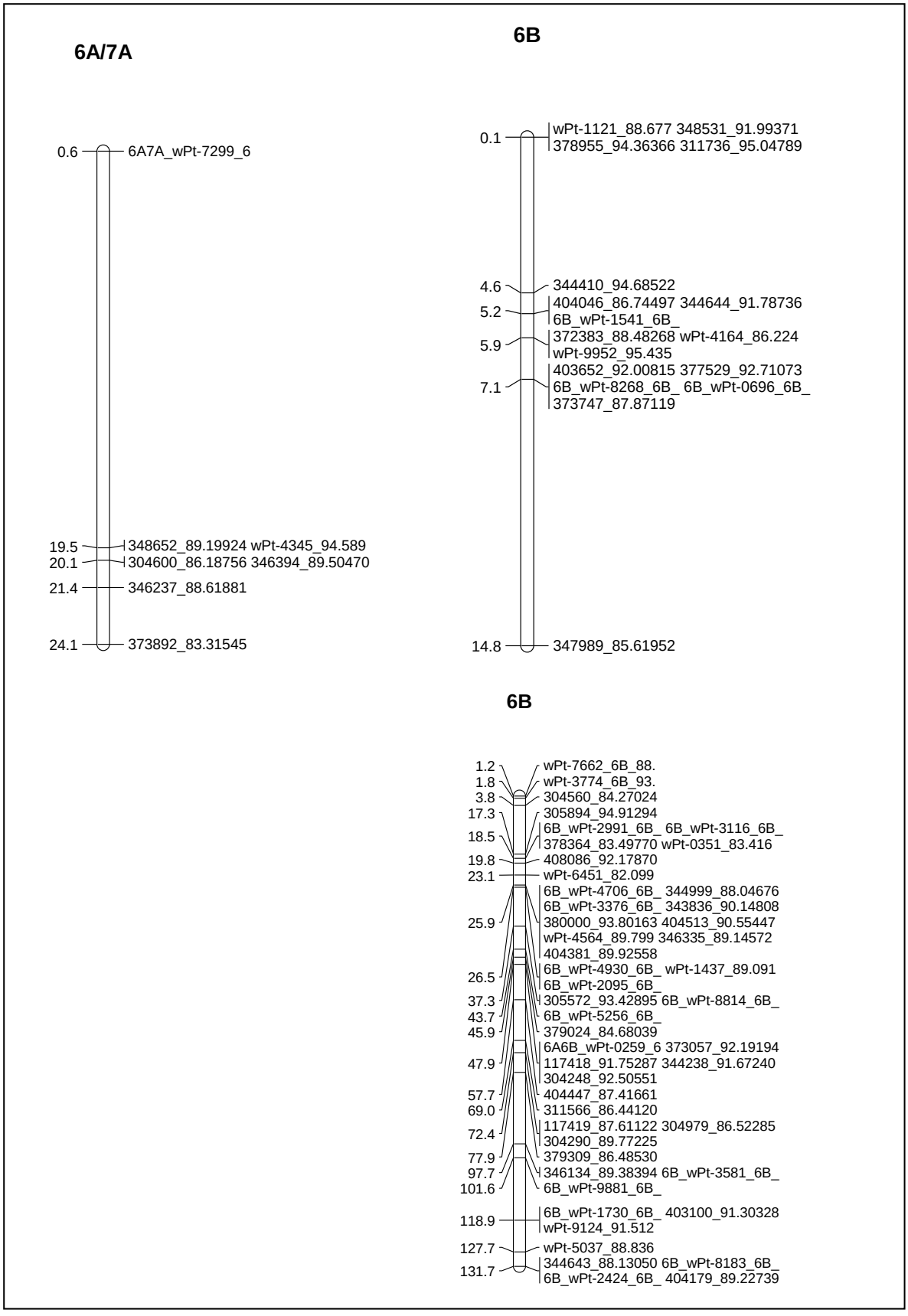


Figure 5.24(f): Linkage map for chromosomes 6A/7A and 6B

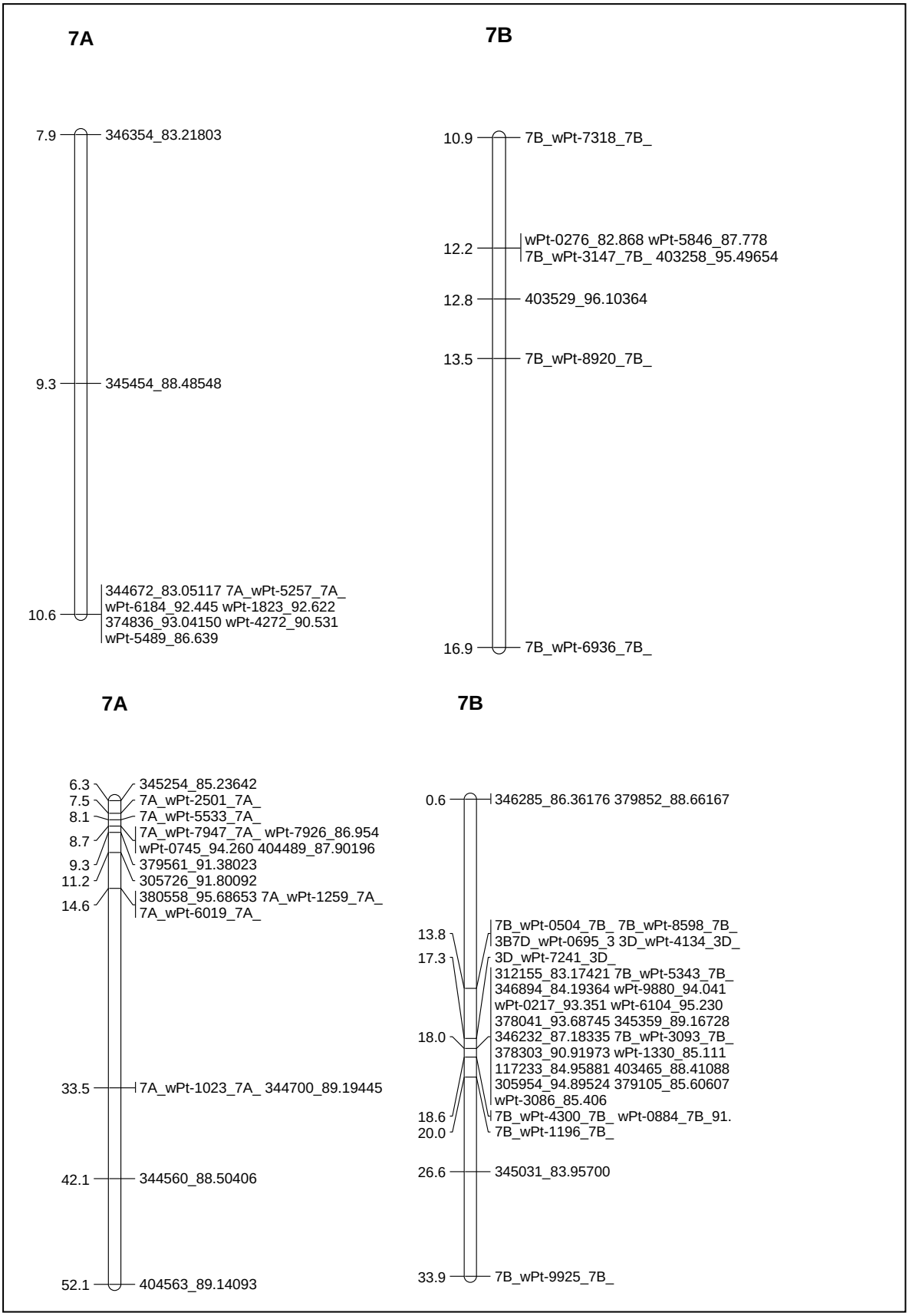


Figure 5.24(g): Linkage map for chromosomes 7A and 7B

Chapter 6

Conclusion and Discussion

"Our greatest glory is not in never falling

but in rising every time we fall"

Confucius

6.0 Conclusions and Discussion

6.1 Conclusions

From the data presented it can be seen that DArT can be used to detect polymorphic sequence differences between wheat cultivars. Specifically, hexaploid and tetraploid cultivated and wild wheats were analysed using DArT to find potential molecular markers. The markers identified can be used to geneotype wheat samples aiding in germplasm identification. In addition, polymorphic markers can be linked to traits of interest and used in plant breeding programs to select or avoid the trait. An example is salt tolerance, where the presence of a known marker increases the tolerance of a plant in high salt soils. This thesis identified many such markers that have been shown to aid plant growth in varying light, temperature and salt conditions. This relatively fast and cheap molecular analysis can be used in the laboratory in preparation of a plant breeding trial to better characterise germplasm without the need for extensive field trials.

Further, epigenetic analysis was explored with the aim to identify DNA methylation polymorphisms using a methyl-sensitive restriction enzyme (McrBc). Success of these experimented in comparison to DNA sdequence analysis was limited due to the variation in marker results. Although expected variation exists, the robustness of the DArT procedure is not as well defined for methylation detection. Further experimentation is required to identify methylation sites with in the genome, using bisulphide conversion or methyl next generation sequencing. Similar to the DArT methodologies, AFLP-based detection of DNA methylation can be used such as the method described by Xu *et al* (2000)[199]. The methylation sensitive restriction enzymes HpaII and MspI were used in addition to EcoRI to investigate DNA methyltion in apple shoots and adult field leaves[200]. Up to 25% difference in the methylation states of the 2 tissue samples were identified[201].

6.2 Discussion and limitations

The recent development of new high-throughput genotyping methods, mainly based on detection of SNPs, will have a large impact on both fundamental and applied research. The large numbers of markers these methods can generate are increasingly used in new breeding programmes for faster introduction of new traits in cultivars.[202] [203] Existing marker technologies (e.g. RFLP, RAPD, SSR and AFLP) have proven their value in the construction of genetic linkage maps, identification of quantitative trait loci, population genetics, biodiversity studies, map-based cloning strategies and marker-assisted selection.

Although very successful these methods have a relative low throughput and high costs. Many new marker methods that use a variety of platforms have a much higher throughput and are often based on the detection of known SNPs. One of the disadvantages of these SNP-based methods is that most require prior DNA sequence information. DArT was developed as an alternative low-cost, high-throughput, open-platform, marker method that does not require prior sequence information.[204] [205]

6.2.1 DArT Advantages

DArT was developed to overcome some of the limitations of existing marker technologies. Although some of these limitations can be alleviated by equipment (e.g. highly parallel capillary electrophoresis), most of them are inherently linked to the sequential nature, low reproducibility or high assay costs of these marker technologies. The hybridisation-based technology DArT, allows the parallel detection and screening of polymorphisms, in a high-throughput manner. One of the advantages of DArT is that no prior DNA sequence information is required. DArT therefore is of special interest for species with limited amounts of genetic resources and for large complex polyploid genomes for which whole genome sequencing may not be amenable or affordable in the near future. Further, DArT markers can be sequenced easily as they do not have to be extracted and purified from gels. Sequenced DArT markers that are positioned on genetic linkage maps can be of high value for genome assembly and can serve as a starting point for map-based cloning approaches of genes. Specific software developed in-house (www.diversityarrays.com/software) identifies and scores markers. High quality and

reproducible markers are selected on the basis of a range of quality thresholds that can be specified by the user.

The complexity reduction method is the most critical step in the DArT technology, determining the efficiency, and therefore the cost of the data generated. Development of methods that enrich for the presence of unique fragments like suppression subtractive hybridization (SSH) can be highly effective as has been shown in sugarcane[206] and the Dendrobium species.[207] Also methods that reduce the complexity in a specific way, like the use of a modified adapter or the amplification of conserved regions in the genome (e.g. transposon display) can be used to tailor genomic representations. By choosing the appropriate genome complexity reduction method, DArT can be used in haploid, diploid and in polyploid genomes.

6.2.2 DArT Disadvantages

Although high-throughput and cost effective, DArT also has some disadvantages. The cloning of the genomic representation is quite laborious and can be biased towards fragments that are PCR amplified more effectively than others or towards relatively small fragments. Optimisation of the complexity reduction method that results in a homogenous size distribution (no strong banding pattern) of the genomic representation and that can reveal a high number of polymorphic clones is therefore recommended when starting with a new species.

For routine application of DArT, equipment capable of high-throughput printing and scanning is required. Although many laboratories have microarray equipment available and DArT is a robust technology, maintaining a consistent quality can be difficult, especially when such facilities are not routinely operated. A good alternative is to outsource the detection and screening of DArT markers by a service lab, such as Triticarte (www.triticarte.com), which offers low-cost DArT genotyping in barley and wheat.

Cross-hybridization is a problem that is known for all hybridization-based techniques. It can be defined as the binding of a probe to a DNA sequence other than the intended target sequence. This may occur if different amplicons in the hybridization

mix share homology, and anneal both to a common printed spot on the slide. Use of shorter and/or more specific probes and the application of more stringent hybridization conditions reduce but do not exclude cross-hybridization.[208] [209] [210] In DArT, cross-hybridization would reduce the number of polymorphic clones that can be identified in the genomic representation. If enough markers remain this would not be an issue.

Further problems include background segregation of a small number of strongly crosshybridizing fragments present in the genomic representation, resulting in mixed or skewed hybridization patterns and incorrect marker scores. These problems, if occurring, can be detected at the locus level in mapping experiments, but will remain undetected in populations or diversity studies. How serious this problem is will be difficult to assess, a skewed segregation can also be caused due to the presence of sequences at multiple loci or selection (preference for inheritance of certain genomic regions). If allelic DArT fragments that differ in length, are both amplified and anneal, then longer fragments may give a stronger fluorescent signal than smaller fragments, as longer fragments usually contain more fluorescent groups. This results in a signal intensity difference. Such fragments will still be scored reliably as a marker if the difference in hybridization intensity is strong enough and consistent among all genotypes screened.

6.2.3 Future improvements

The development of more efficient complexity reduction methods, improved labeling methods, and new algorithms for the co-dominant scoring of DArT markers are likely to result in the discovery of markers with an increased efficiency, that are more informative and cheaper. Once DArT markers are identified new arrays can be constructed containing only markers desired for specific applications (e.g. following markers tightly linked to a particular trait of interest or characterization of a subset of the gennplasm). These arrays can be formatted in such a way that approximately 1,000 markers can be screened on a single multiarray slide for processing (e.g. 1,000 markers for 96 genotypes on a single array). These arrays or the use of DArT markers on other platforms will enable the rapid screening of markers in large populations. A single DArT assay covering a 'standard' set of agriculturally important loci may soon be more cost-effective than 'mixing and matching' single-marker assays.[211] Higher marker densities, on the other hand, could be achieved for

chromosome landing[212] and map-based cloning approaches [213] by simply pyramiding DArT markers from several genomic representations.

In addition to DArT, other novel SNP-based genotyping methods have been developed. These have been discussed extensively in a large number of reviews such as Grupta *et al* (2001).[214] Recent advantages in sequencing technology enable the development of a next generation of SNP marker technologies, such as the Complexity reduction of Polymorphic sequences (CroPs) technology developed by Keygene in collaboration with 454- Life Sciences.[215] Different array platforms have been developed for the use of expression profiling but also for the detection of markers.[216] Additional array-based genotyping methods (bead arrays developed by Illumina, LYNX, MassARRAY) and platforms (DNA chip, printed and self assembling arrays, MALDI-TOFF mass spectrometry) for detection of SNPs have been developed the last few years.[217] [218] [219] All of these technologies have varying capabilities but are still relatively expensive and often rely on prior sequence information.[220] Their future is dependent on being easier, cheaper and more reliable than current hybridization arrays. DNA chips have become powerful tools not only for monitoring expression of genes, but also for identification of genomic insertions and deletions and copy number changes with Comparative Genome Hybridization (CGH), rapid identification of binding sites of specific DNA-binding proteins with Chromatin Immunoprecipitation (ChIP) arrays and for the detection of polymorphisms (SNPs) in genomic DNA. [221] [222] [223] Other array-based genotyping and SNP-based detection methods are available.

6.3 Competing technologies

6.3.1 Affymetrix

Affymetrix currently dominates the market with respect to the production of high-density arrays (GeneChips) for gene expression analysis (www.affymetrix.com). Affymetrix has developed a unique array design based on a perfect match / mismatch probe strategy.[224] For each probe designed to be perfectly complementary to a target sequence, a second probe is generated that is identical except for a single-base mismatch in its centre. These probe pairs are called the perfect match (PM) probe and the mismatch (MM) probe. Oligonucleotide probes are chosen based on uniqueness criteria and composition design rules and most arrays currently

contain between 15 and 22 probe pairs per gene. The use of the PM minus MM differences averaged across a set of probes greatly reduces the contribution of background and cross-hybridization and increases the quantitative accuracy and reproducibility of the measurement. Probe synthesis occurs in parallel, resulting in the addition of an A, C, T or G nucleotide to multiple growing chains simultaneously. To define which oligonucleotide chains will receive a nucleotide in each step, photolithographic masks, carrying 18-20 square micron windows that correspond to the dimensions of individual features, are placed over the coated wafer. The windows are distributed over the mask based on the desired sequence of each probe. When ultraviolet light is exposed over the mask in the first step of synthesis, the exposed linkers become deprotected and are available for nucleotide coupling. In the following synthesis step, another mask is placed over the wafer to allow the next round of deprotection and coupling. This process is repeated until the probes reach their full length, generally 25 nucleotides. One of the first uses of oligonucleotide arrays for SNP detection was the use of Variation Detection Arrays (VDAs).[225] [226] In this array design, 16 features are routinely synthesized for each locus, comprising coding and non-coding strands for the two alleles with all four combinations of bases for the polymorphic site of each allele. VDAs have been successfully used for large-scale SNP screens,[227] [228] [229] but require amplification of each locus in the genome individually. Further disadvantages include the need to know the sequence composition for the alleles to be queried and the high production costs for the design, optimization and synthesis. The first example in which VDA arrays were hybridized with subsets of total genomic DNA that was amplified with a single primer was shown by Dong et al. (2001)[230]. The method that was able to perform large-scale genotyping using a single primer for amplification was developed by Kennedy et al. (2003).[231] These approaches however had still the disadvantage of the need for specifically designing the arrays for each specific application. Winzeler et al. (1998) were the first to describe the use of existing gene expression arrays from Affymetrix for genome-wide SNP detection, by hybridizing total genomic DNA to these arrays.[232] The genetic variation is identified by measurement of the differential hybridization intensities to the features on these arrays. The polymorphisms that are being discovered in this way are called Single Feature Polymorphisms (SFPs). Direct hybridisation of labeled total genomic DNA to oligonucleotide expression arrays for this SFP detection was initially demonstrated in model organisms with relatively small genomes, such as yeast[233] [234], Zebrafish[235] and Arabidopsis[236]. This approach has successfully been used in high density haplotyping of recombinant inbred lines[237] and in pooled DNA genotyping for association studies in Arabidopsis.[238] [239]

A more recent development is the use of expression data itself (RNA based hybridization) for the detection of markers. These Gene Expression Markers (GEMs) are based on differences in transcript levels that exhibit bimodal distributions in segregating progeny, while SFP markers rely on differences in bybridization to individual oligonucleotide probes. Unlike SFPs, GEMs can be derived from any type of DNA-based expression microarray (long oligonucleotide probes, spotted cDNAs, or short oligonucleotides) because they are based on gene expression differences, not on individual probe hybridization. The first use of these expression-level polymorphisms was shown in S. cerevisiae [240] [241] and later in Arabidopsis.[242]

6.3.2 NimbleGen

The design of the Affymetrix arrays is rather rigid, mainly due to the high costs involved in manufacturing the physical masks. Recent developments in oligo nucleotide array synthesis technology, such as the Maskless Array Synthesis (MAS) technology allow for flexible design and corresponding lower costs.[243] [244] The MAS method, employed by NimbleGen (www.nimblegen.com), uses a digital micromirror system to direct light at specific elements during each round of synthesis, thus allowing for quick turnaround in array design and optimization. FlexGen B.V., a technological spin-off from Dutch Space B.V., is currently developing a bench top instrument (FlexArrayer) that uses a laser-guided, spot-by-spot photochemical oligonucleotide synthesis process based on the virtual masking technology capable of synthesizing 'Arrays-on-Demand'.

6.3.3 Bead arrays

Another application in which arrays are used for high-throughput genotyping is the bead-array technology, as employed by Illumina (www.illumina.com). The technology produces bead arrays in either a 96-well format (Sentrix Array Matrix) or on a silicon-based single slide format (Sentrix BeadChip).[245] [246] Several hundred thousand copies of unique 50-mer oligonucleotide are covalently attached to a 3 µm silica bead. These beads (from 384 to 250,000 types) are pooled and self-assembled onto an array. On average, 30 copies of each bead type are present on an array.[247] For genome-wide (SNP) genotyping and association studies, two types of assays can be employed. The Infinium whole genome genotyping (WGG) assay employs a one-

colour allele-specific primer extension (ASPE) assay and two bead types per SNP to genotype over 100,000 SNPs. For genetic analysis in large populations, the GoldenGate assay was developed that uses solid-phase allele specific extension and ligation and multiplexed amplification to genotype up-to 1,536 SNPs in each sample.

Instead of fixing beads on a solid support as with the bead arrays, there is also the option to maintain the beads in suspension (e.g. suspension arrays). Luminex (www.luminexcorp.com) use this technology incombination with a fluorescent cell sorter to decode the beads and measure the fluoresence of the target hybridised to the beads. The system is suitable for applications with moderate multiplex levels, because the scanner can distinguish up to 100 bead types and each SNP assay requires two bead types. Applications for this type of platform focus on human, mouse and rat genomes and are currently not available for agricultural crops.

6.3.4 Electronic arrays

Further technology employs semi-conductor-based (electrodes) in situ oligonucleotide synthesis, hybridization and detection methods. Examples of such systems are the NanoChips from NanoGen (www.nanogen.com) and the ElectraSense platform from Combimatrix (www.combimatrix.com). The combimatrix system synthesis oligonuclotides by activating micro-electrodes which selectively generate acid by means of an electrochemical reaction that will deprotect the growing oligonucleotide chain, activating it for binding of the nucleotide. One of the advantages of the CombiMatrix and NanoGen arrays is that they can be re-used, resulting in substantial reduction of assay costs.

6.3.5 Real-time quantitative PCR

There are various protocols available for SNP and polymorphism/mutation analysis using real-time quantitative PCR (qPCR). Real-time qPCR can be used to successfully identify small variations between sequences, with one common method being the analysis of melting curves. A shift in the melting curve of several degrees can be seen when analylising the metling curves of short fluorescent probes bound to wild-type and mutant genes.

Another method is to use dual-labelled, hydroloysis or TaqMan (Roche Molecular Systems) probes, where the 5'-3' endonuclease reaction is greatly impared when a well-designed probe mismatches its target sequence, even by a single base. Where there is a mismatch, the fluorfor is not excised from the probe and not removed from its close proximity to the quencher, hence no fluorescence is observed. This is compared to the wild-type where the probe binds perfectly, the fluorofor is excised and removed from the quencher and fluoresces.

The immediate disadvantages to these methods are that they are low through-put, often single or small multiplex reactions requiring sequence specific primers and probes. Costs are high due to custom primers and probes and they require real-time qPCR instrumentation.

6.3.6 Next-Generation sequencing

High-throughput sequencing or 'next generation sequencing' allows for millions of sequences to be analysed at once.[248] There are three main technologies available. Roche 454 (www.454.com) sequencing uses water droplets in an oil emulsion where the DNA template attached to a bead forms a clonal colony. One bead is sorted into one well and sequenced using luciferase to generate light. Intimediate read lengths are then combined. Illumiuna (Solexa) (www.illumina.com) sequencing uses reversible dye-terminators where DNA is attached to primers on a glass slide and amplified so that local colonal colonies are formed. Each base ddNTP is added and extended one nucleotide at a time. Fluorescent signals from labeled nucleotides are detected and then removed ready for the next cycle. The Applied Biosystems SOLiD (www.appliedbiosystems.com) system employs sequencing by ligation, where DNA is amplified by emulsion PCR then the resulting bead is deposited on a glass slide and the fluorescence read then compiled.

This technology is expensive but very fast and high-throughput. Individual's genomes could be sequenced and compared for individuals/populations identifying any mutations present. As the prices per sample and the price of instrumentation reduces, other applications are being developed such as SNP detection, where methods for parallel enrichment are used for known SNPs and mutations, such as NimbleGen's targeted resequencing arrays or comparative genome sequencing (CGS) protocols (www.nimblegen.com). In addition, Applied Biosystems use a

DiBayes algorithym for SNP detection during next generation sequencing. The algorithm can distinguish between the sequenced sequence and the reference sequence. Again, this technology is useful for small genomes but not cost effective for large genomes such as wheat where little sequence data is available.

6.4 DArT vs array-based genotyping technologies

Major drawbacks from the use of the above mentioned array-based SNP genotyping methods are that prior sequence information is required and the high set-up costs involved in development of such chips. In addition, large and polyploid genomes such as wheat may not be amenable to the whole-genome hybridisation approach, although attempts to reduce the complexity for such crops have been made.[249] [250] It remains to be seen whether the development of such sequenced-based arrays could become affordable for a broad range of agricultural species. DArT therefore is specifically a good alternative for crops with no or limited genetic resources available.[251] [252] DArT is independent of prior sequence information and can be fine tuned to detect polymorphism in genomes of virtually any size as has been proven successfully in the 16,000 Mbp genome of hexaploid wheat,[253] the complex and polyploid genome of sugarcane[254], in the triploid banana 873 Mbp genome[255] and in the 90,000 Mbp lily genome (unpublished data), currently one of the largest known plant genomes.

6.5 A new model for technology delivery

The biotechnology industry is protecting an ever-increasing amount of DNA sequences, technologies and methods by means of intellectual property (IP). This is done because of both the long timelines associated with product development and the high costs of commercialising these products. The protection of technologies however limits the use and can slow down the development of new innovative products.[256] [257] The Center of Application of Molecular Biology to International Agriculture (CAMBIA) has enacted an initiative that aims to make the biotechnology patent landscape more transparent and provide opportunities for open access to technology.[258] [259] [260] This initiative is called Biological Innovation for Open Society (BiOS, www.bios.net). The BiOS initiative has developed a set if internet-based information tools to promote collaborative work among researchers and to ctitically

analyse existing patent landscapes. The BiOS initiative will stimulate decentralised, cooperative innovation in the application of biological technologies that are available under a cost-free BiOS licence, which ensures that improvements to the technology are shared within the research community.

6.6 Future work

Further experimentation is needed to validate each significant candidate polymorphic marker, and to sequence it so that probes can be designed to easyily detect the presence or absence of this marker in different individuals or populations. Genotyping for trait-linked markers can then be performed in the field using basic PCR or qPCR techniques.

Further experiments can be designed to test additional wheat cultivars, both domesticated and wild varieties to increase the available gene pool. Salt, light and temperature experiments can be expanded to include additional conditions to further discriminate any potential polymorphisms.

These results can then be included into salt and environmental tolerance plant programs that are ongoing through CSIRO Plant Industry, The University of Adelaide, Australian Centre for Plant Functional Development, the GRDC and University of Sydney. CSIRO studies have shown that planting new salt tolerant durum wheats in different levels of salinity and comparing their yield with other durum wheats, an impressive 25 per cent yield advantage under saline soil conditions is observed[261]. This research is being performed by Dr Rana Munns and Dr Richard James in Canberra where the salt tolerance experiments were performed.

Appendices

Chapter 7

7.0 Appendices

7.1 Reagents

7.1.1 Fresh working solution used for DNA extractions

Fresh buffer working solution (60 ml):

Add - 0.3g Sodiumdisulfite (Sodium metasulfide)

- 1.2g (2% w/v) PVP ()
- 25 ml Extraction buffer stock
- 25 ml Lysis buffer stock
- 10 ml sarcosyl (5% w/v)

Stock Solutions

Extraction Buffer:	**Lysis Buffer:**
0.35 M Sorbitol	0.2 M Tris pH 7.5
0.1 M Tris pH 7.5	0.05 M EDTA
5 nM EDTA	2 M NaCl
	2% CTAB

7.1.2 Freezing Media

To 4 litres of Milli-Q water add

80.0 g LB	0.4 g $MgSO_4.7H_2O$
32.8 g $K_2HPO_4.3H_2O$	3.6 G $(NH4)^2SO_4$
7.2 g KH_2PO_4	176 mls glycerol (4.4 % v/v)
2.0 g Na-citrate.$2H_2O$	

Mix thoroughly, autoclave and add 100 mg per litre ampicillin before use.

7.1.3 DArTSpotter2

50% DMSO, 1.5 M sorbitol, 0.1 M TEA.HCl, 0.5 % dextran, 0.02 % CHAPS

7.1.4 S.O.C Medium

2% Tryptone, 0.5% Yeast Extract, 10mM Nacl, 2.5nM KCl, 10nM $MgCl_2$, 10mM $MgSO_4$, 20mM glucose

7.1.5 Insert Amplification Mix

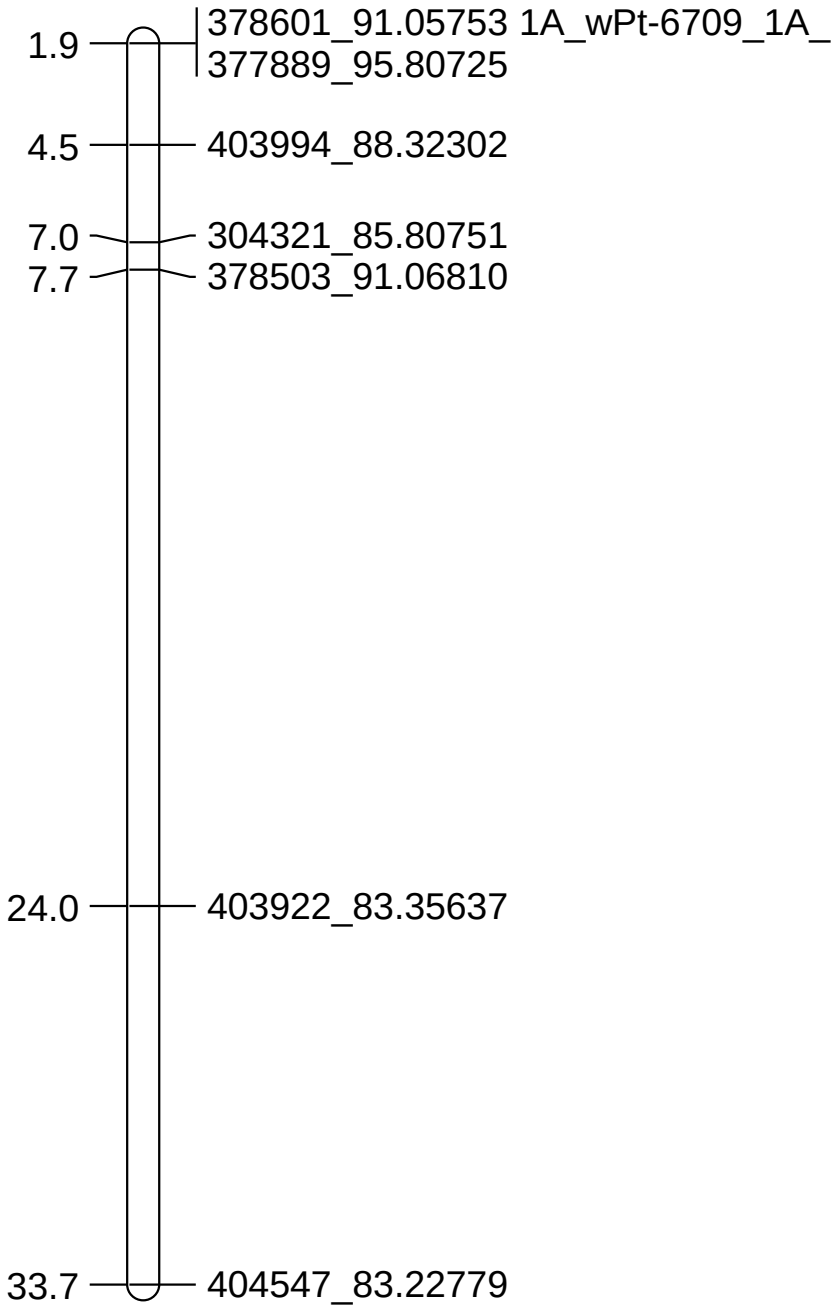
1x Possum Taq buffer (500mM Tris, 60mM HCl, 160mM $(NH_4)_2SO_4$, 15mM $MgCl_2$) (home-made), 200µM dNTPs, 0.2µM M13 forward and 0.2µM M13 reverse primer and 1.0µl Possum Taq (home-made)

7.2 Reagent Suppliers

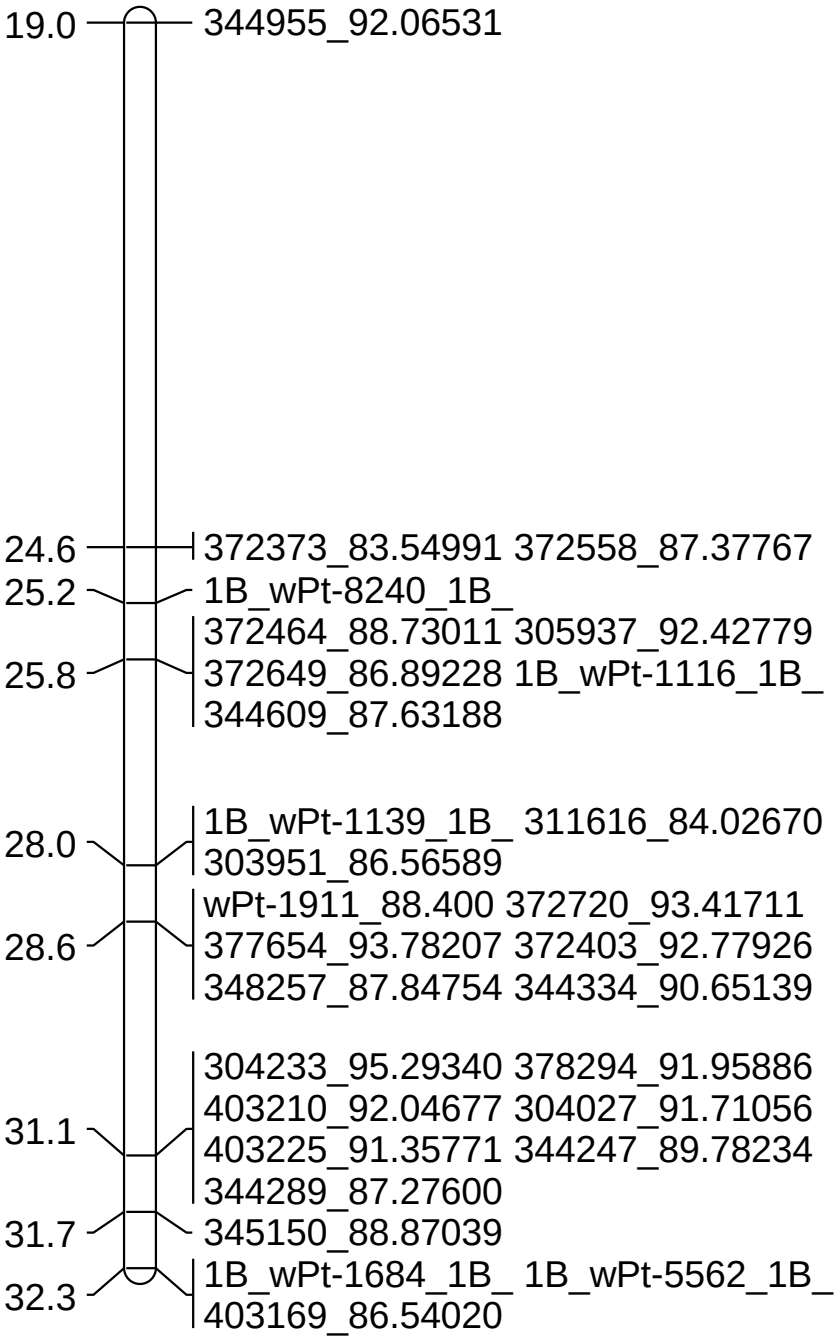
Company	Contact / Internet web address
Sigma-Aldrich	www.sigmaaldrich.com
HOBO	Onset Computer Corporation, www.onsetcomp.com
Invitrogen	www.invitrogen.com
Eppendorf	www.eppendorf.com.au
Erie Scientific	www.eriesci.com
BioRobotics	Genomic Solutions, www.genomicsolutions.com
Tecan	www.tecan.com
GE Health	www.gehealthcare.com
Fermentas	www.fermentas.com
New England Biolabs	GeneSearch, www.genesearch.com.au, www.neb.com
Clonetec	www.clontech.com
Promega	www.promega.com
Sigma-Aldrich	www.sigma-aldrich.com

7.3 Linkage maps (high resolution)

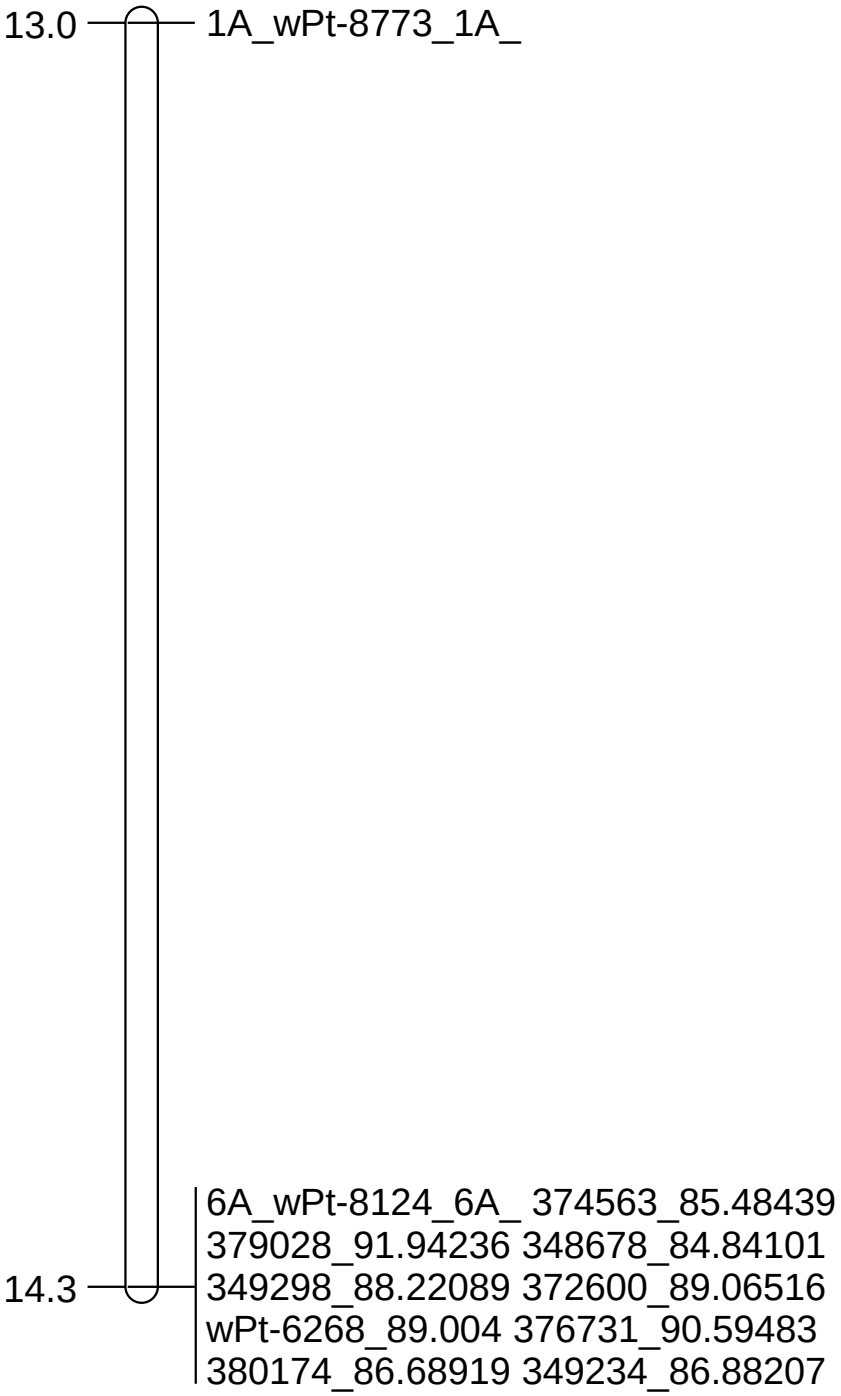
1A



1B


19.0
344955_92.06531
24.6
372373_83.54991 372558_87.37767
25.2
1B_wPt-8240_1B_
372464_88.73011 305937_92.42779
25.8
372649_86.89228 1B_wPt-1116_1B_
344609_87.63188
28.0
1B_wPt-1139_1B_ 311616_84.02670
303951_86.56589
28.6
wPt-1911_88.400 372720_93.41711
377654_93.78207 372403_92.77926
348257_87.84754 344334_90.65139
31.1
304233_95.29340 378294_91.95886
403210_92.04677 304027_91.71056
403225_91.35771 344247_89.78234
344289_87.27600
31.7
345150_88.87039
32.3
1B_wPt-1684_1B_ 1B_wPt-5562_1B_
403169_86.54020

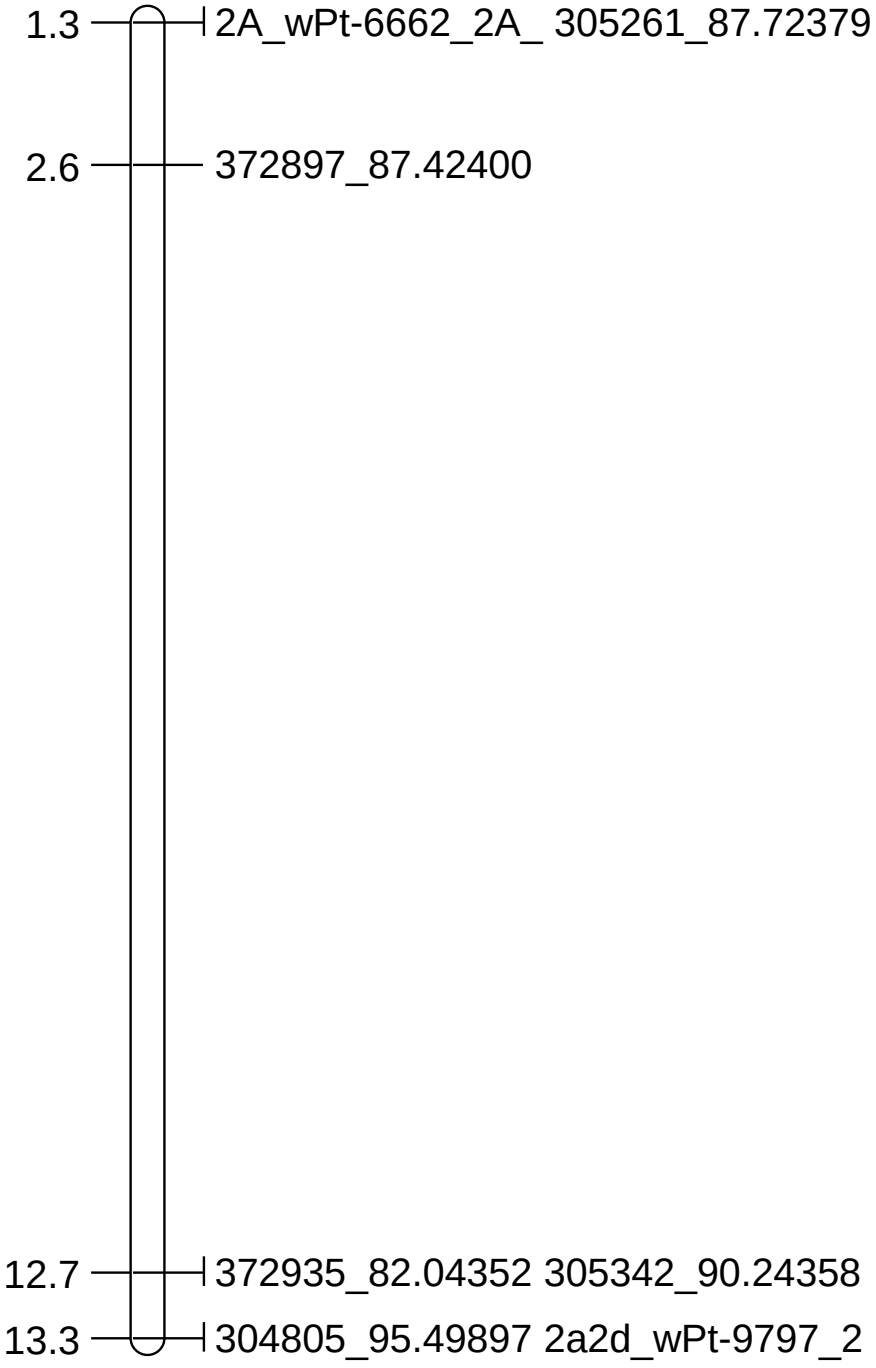
1A


13.0
1A_wPt-8773_1A_
14.3
6A_wPt-8124_6A_ 374563_85.48439
379028_91.94236 348678_84.84101
349298_88.22089 372600_89.06516
wPt-6268_89.004 376731_90.59483
380174_86.68919 349234_86.88207

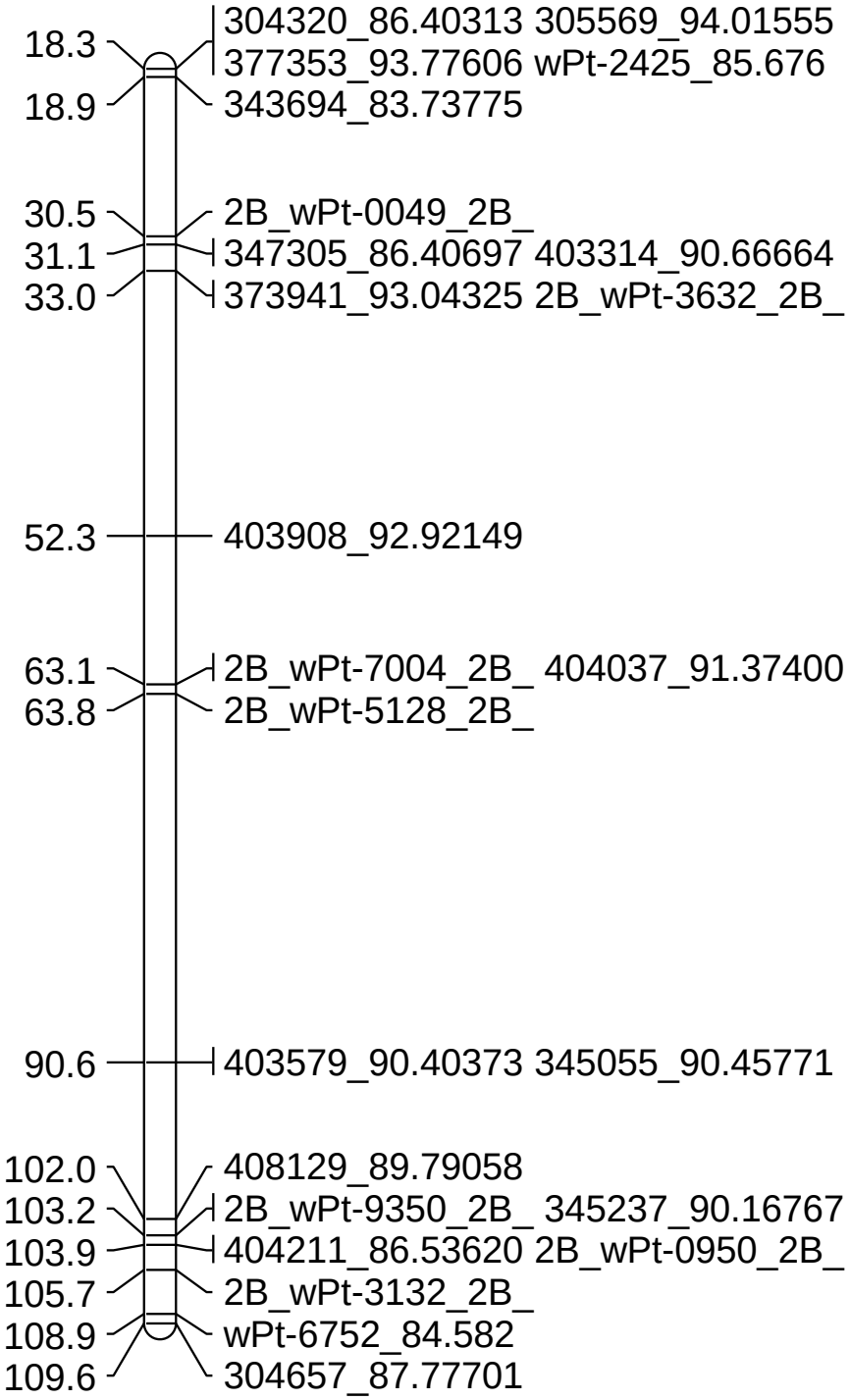
1B


5.6
3BwPt-7212_3B_9
11.5
404584_88.34549
18.8
403903_85.05967 304559_93.06910
22.0
1B_wPt-0705_1B_ 1B_wPt-3227_1B_

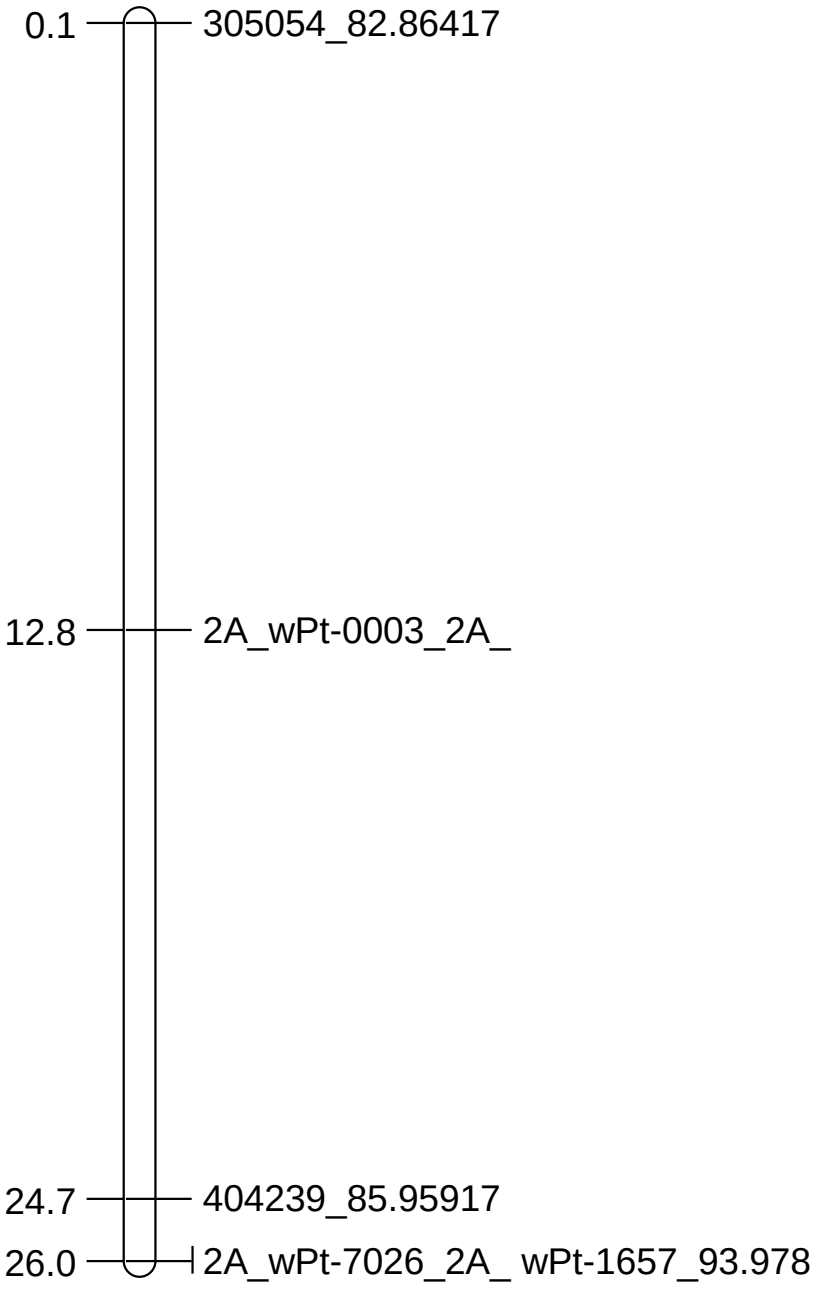
2A


1.3
2A_wPt-6662_2A_ 305261_87.72379
2.6
372897_87.42400
12.7
372935_82.04352 305342_90.24358
13.3
304805_95.49897 2a2d_wPt-9797_2

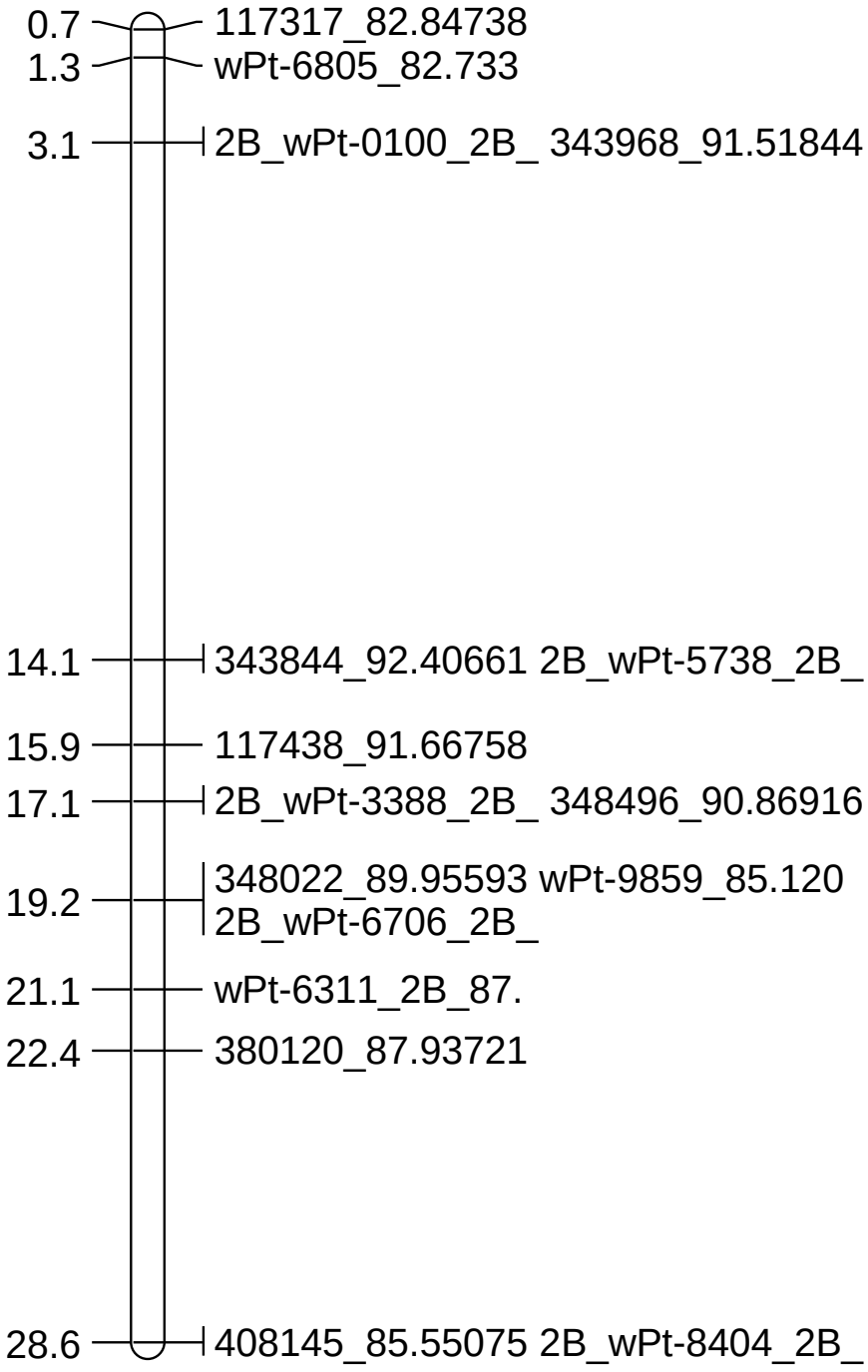
2B


304320_86.40313 305569_94.01555
18.3
377353_93.77606 wPt-2425_85.676
18.9
343694_83.73775
30.5
2B_wPt-0049_2B_
31.1
347305_86.40697 403314_90.66664
33.0
373941_93.04325 2B_wPt-3632_2B_
52.3
403908_92.92149
63.1
2B_wPt-7004_2B_ 404037_91.37400
63.8
2B_wPt-5128_2B_
90.6
403579_90.40373 345055_90.45771
102.0
408129_89.79058
103.2
2B_wPt-9350_2B_ 345237_90.16767
103.9
404211_86.53620 2B_wPt-0950_2B_
105.7
2B_wPt-3132_2B_
108.9
wPt-6752_84.582
109.6
304657_87.77701

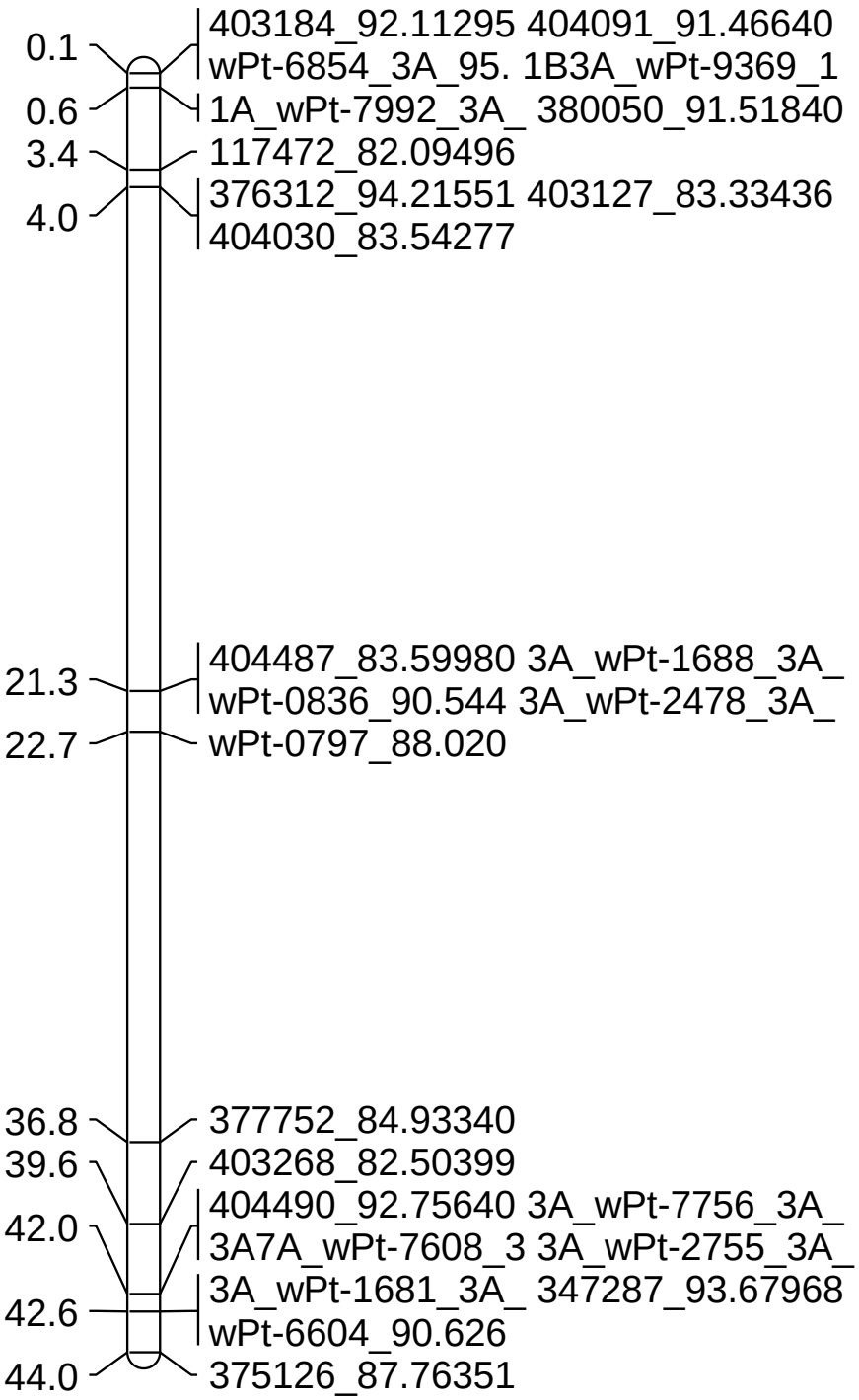
2A


0.1
305054_82.86417
12.8
2A_wPt-0003_2A_
24.7
404239_85.95917
26.0
2A_wPt-7026_2A_ wPt-1657_93.978

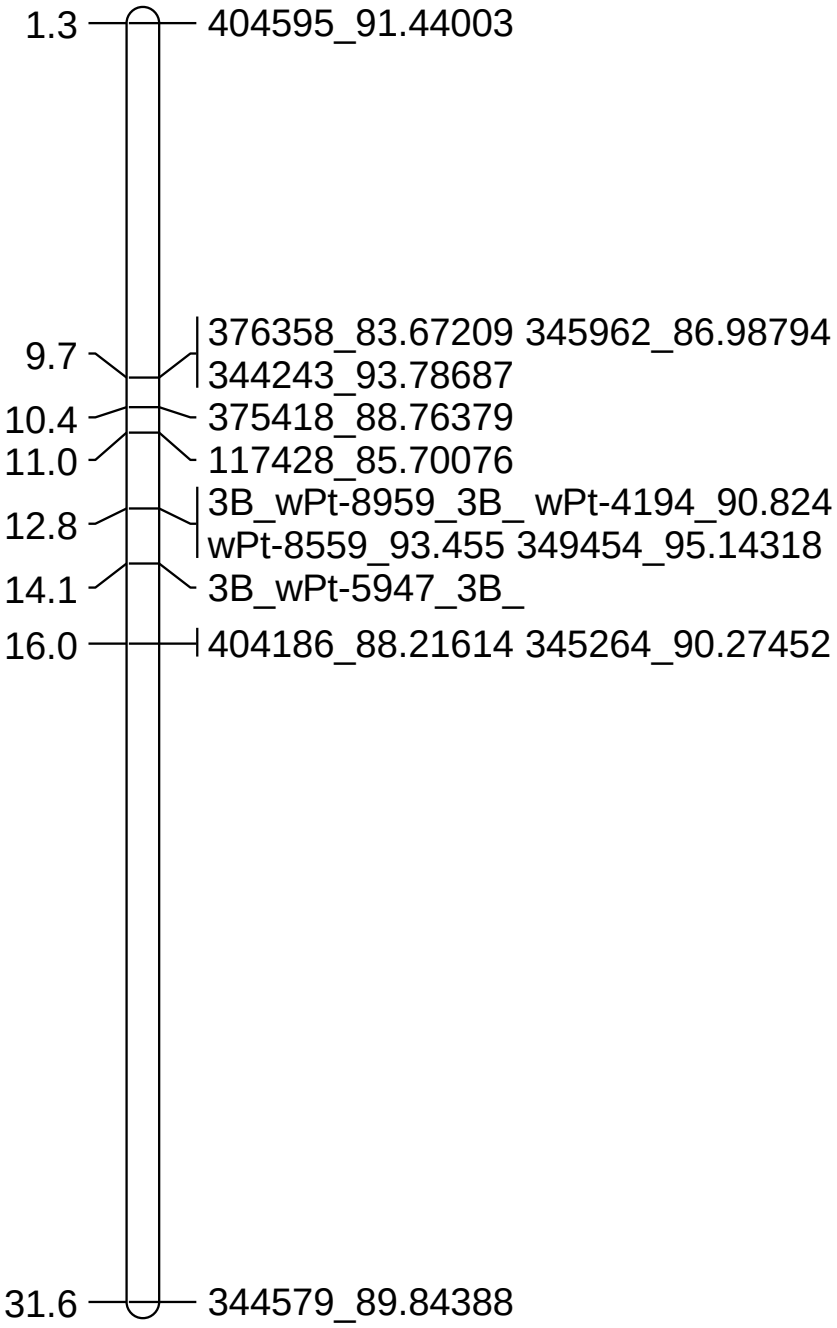
2B


0.7
117317_82.84738
1.3
wPt-6805_82.733
3.1
2B_wPt-0100_2B_ 343968_91.51844
14.1
343844_92.40661 2B_wPt-5738_2B_
15.9
117438_91.66758
17.1
2B_wPt-3388_2B_ 348496_90.86916
19.2
348022_89.95593 wPt-9859_85.120
2B_wPt-6706_2B_
21.1
wPt-6311_2B_87.
22.4
380120_87.93721
28.6
408145_85.55075 2B_wPt-8404_2B_

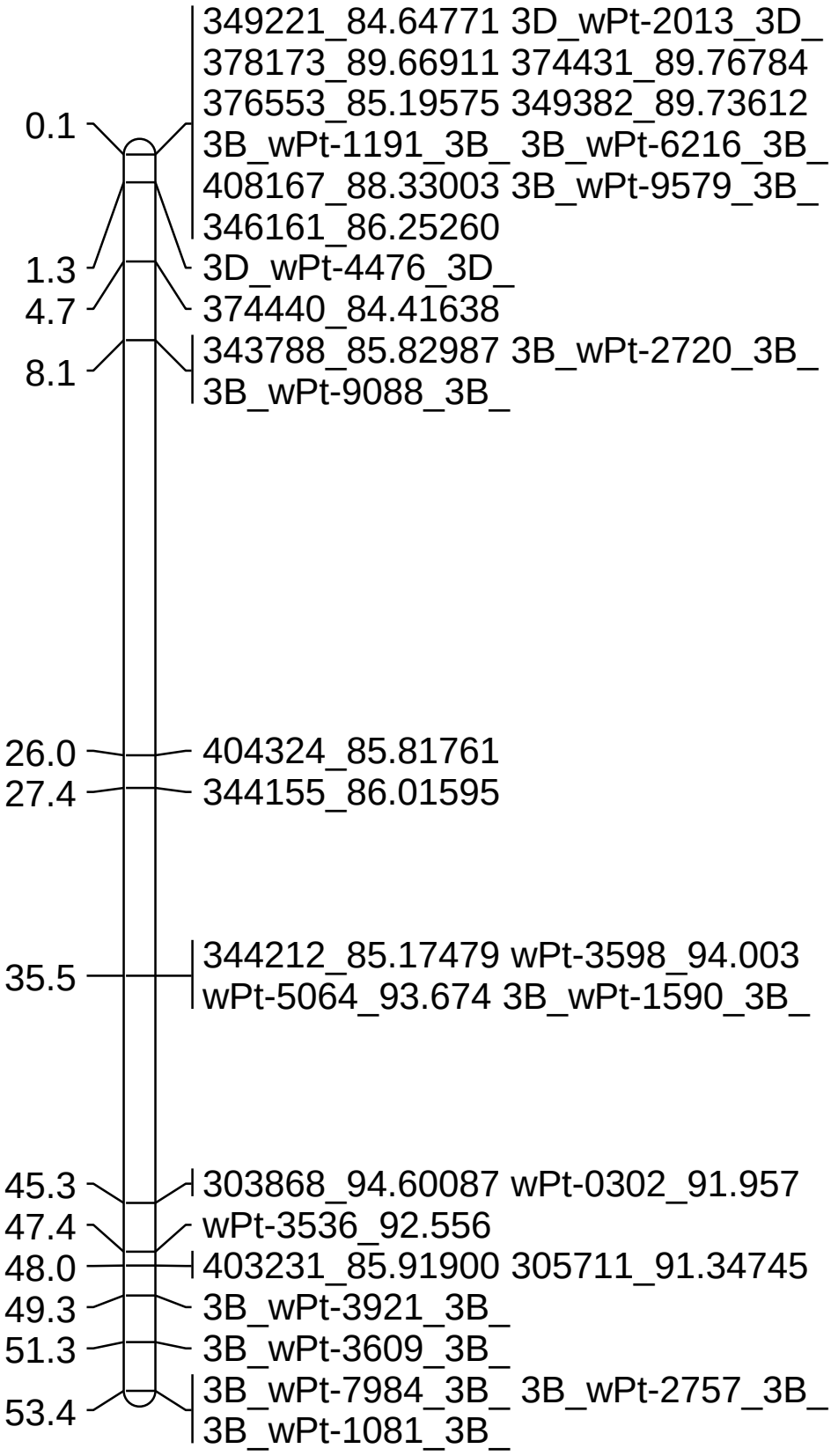
3A


403184_92.11295 404091_91.46640
0.1
wPt-6854_3A_95. 1B3A_wPt-9369_1
0.6
1A_wPt-7992_3A_ 380050_91.51840
3.4
117472_82.09496
376312_94.21551 403127_83.33436
4.0
404030_83.54277
404487_83.59980 3A_wPt-1688_3A_
21.3
wPt-0836_90.544 3A_wPt-2478_3A_
22.7
wPt-0797_88.020
36.8
377752_84.93340
39.6
403268_82.50399
404490_92.75640 3A_wPt-7756_3A_
42.0
3A7A_wPt-7608_3 3A_wPt-2755_3A_
3A_wPt-1681_3A_ 347287_93.67968
42.6
wPt-6604_90.626
44.0
375126_87.76351

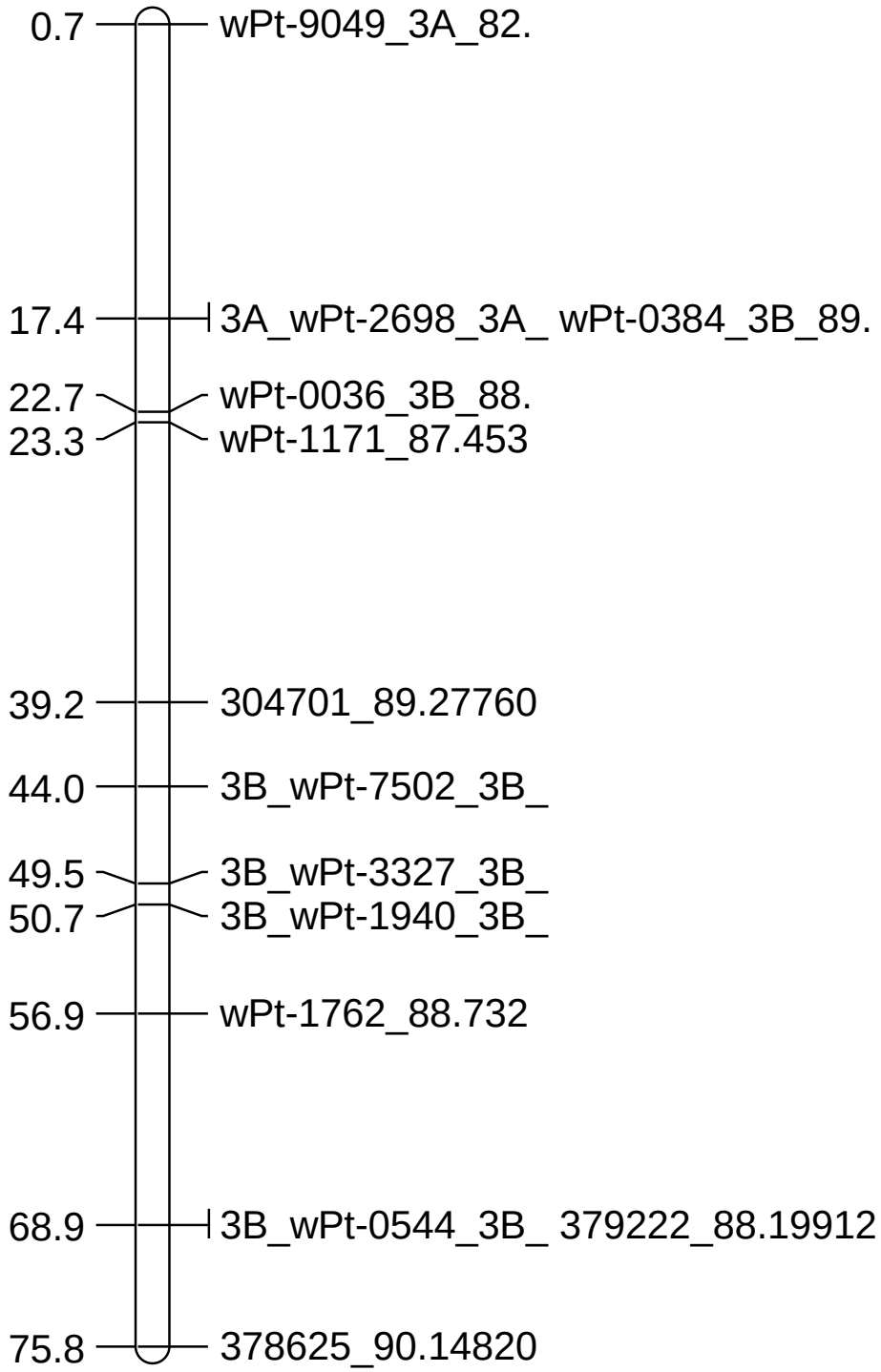
3B


1.3
404595_91.44003
9.7
376358_83.67209 345962_86.98794
344243_93.78687
10.4
375418_88.76379
11.0
117428_85.70076
12.8
3B_wPt-8959_3B_ wPt-4194_90.824
wPt-8559_93.455 349454_95.14318
14.1
3B_wPt-5947_3B_
16.0
404186_88.21614 345264_90.27452
31.6
344579_89.84388

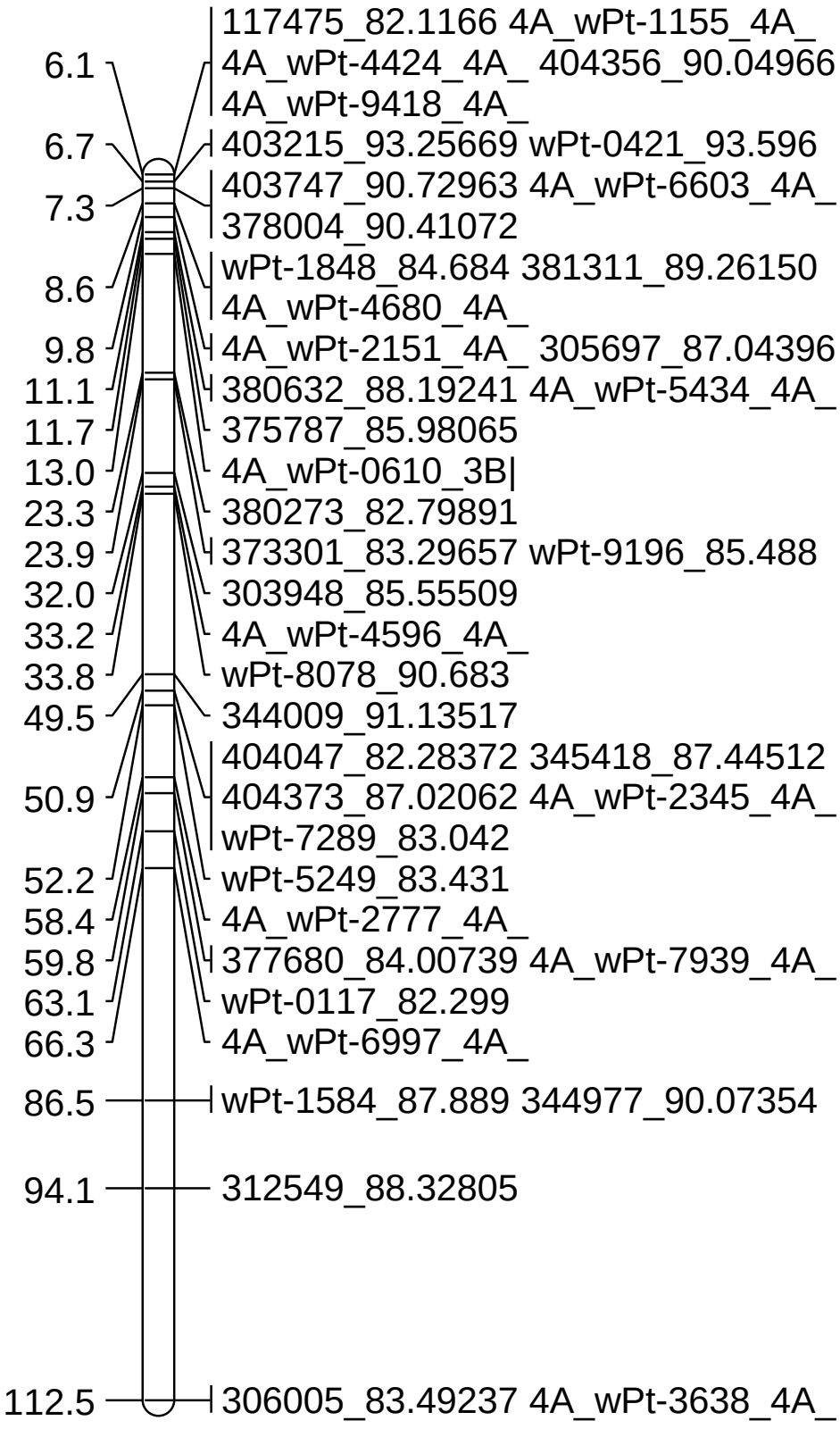
3B


349221_84.64771 3D_wPt-2013_3D_
378173_89.66911 374431_89.76784
376553_85.19575 349382_89.73612
0.1
3B_wPt-1191_3B_ 3B_wPt-6216_3B_
408167_88.33003 3B_wPt-9579_3B_
346161_86.25260
1.3
3D_wPt-4476_3D_
4.7
374440_84.41638
343788_85.82987 3B_wPt-2720_3B_
8.1
3B_wPt-9088_3B_
26.0
404324_85.81761
27.4
344155_86.01595
344212_85.17479 wPt-3598_94.003
35.5
wPt-5064_93.674 3B_wPt-1590_3B_
45.3
303868_94.60087 wPt-0302_91.957
47.4
wPt-3536_92.556
48.0
403231_85.91900 305711_91.34745
49.3
3B_wPt-3921_3B_
51.3
3B_wPt-3609_3B_
3B_wPt-7984_3B_ 3B_wPt-2757_3B_
53.4
3B_wPt-1081_3B_

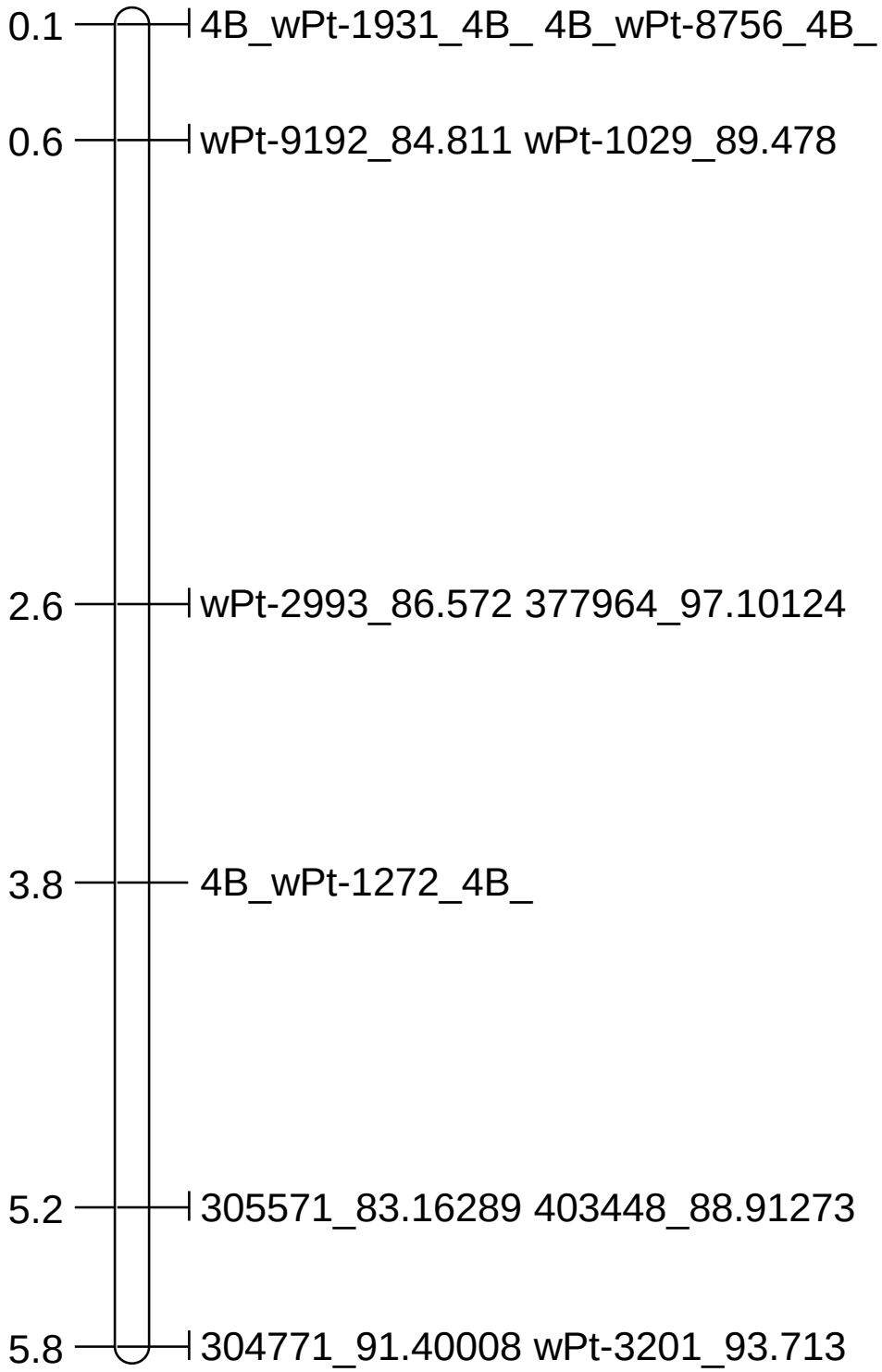
3B


0.7
wPt-9049_3A_82.
17.4
3A_wPt-2698_3A_ wPt-0384_3B_89.
22.7
wPt-0036_3B_88.
23.3
wPt-1171_87.453
39.2
304701_89.27760
44.0
3B_wPt-7502_3B_
49.5
3B_wPt-3327_3B_
50.7
3B_wPt-1940_3B_
56.9
wPt-1762_88.732
68.9
3B_wPt-0544_3B_ 379222_88.19912
75.8
378625_90.14820

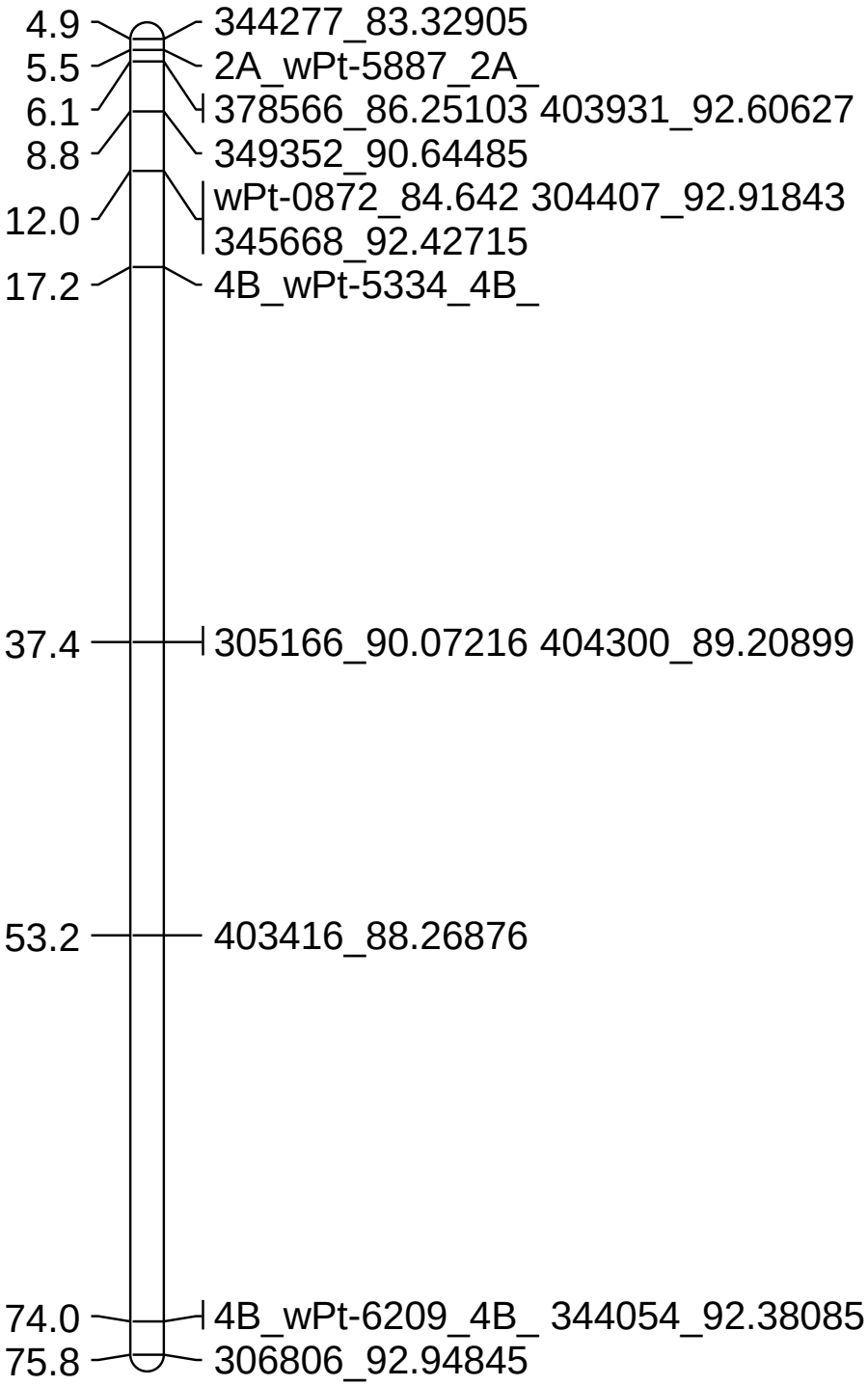
4A


117475_82.1166 4A_wPt-1155_4A_
6.1
4A_wPt-4424_4A_ 404356_90.04966
4A_wPt-9418_4A_
6.7
403215_93.25669 wPt-0421_93.596
7.3
403747_90.72963 4A_wPt-6603_4A_
378004_90.41072
8.6
wPt-1848_84.684 381311_89.26150
4A_wPt-4680_4A_
9.8
4A_wPt-2151_4A_ 305697_87.04396
11.1
380632_88.19241 4A_wPt-5434_4A_
11.7
375787_85.98065
13.0
4A_wPt-0610_3B|
23.3
380273_82.79891
23.9
373301_83.29657 wPt-9196_85.488
32.0
303948_85.55509
33.2
4A_wPt-4596_4A_
33.8
wPt-8078_90.683
49.5
344009_91.13517
404047_82.28372 345418_87.44512
50.9
404373_87.02062 4A_wPt-2345_4A_
wPt-7289_83.042
52.2
wPt-5249_83.431
58.4
4A_wPt-2777_4A_
59.8
377680_84.00739 4A_wPt-7939_4A_
63.1
wPt-0117_82.299
66.3
4A_wPt-6997_4A_
86.5
wPt-1584_87.889 344977_90.07354
94.1
312549_88.32805
112.5
306005_83.49237 4A_wPt-3638_4A_

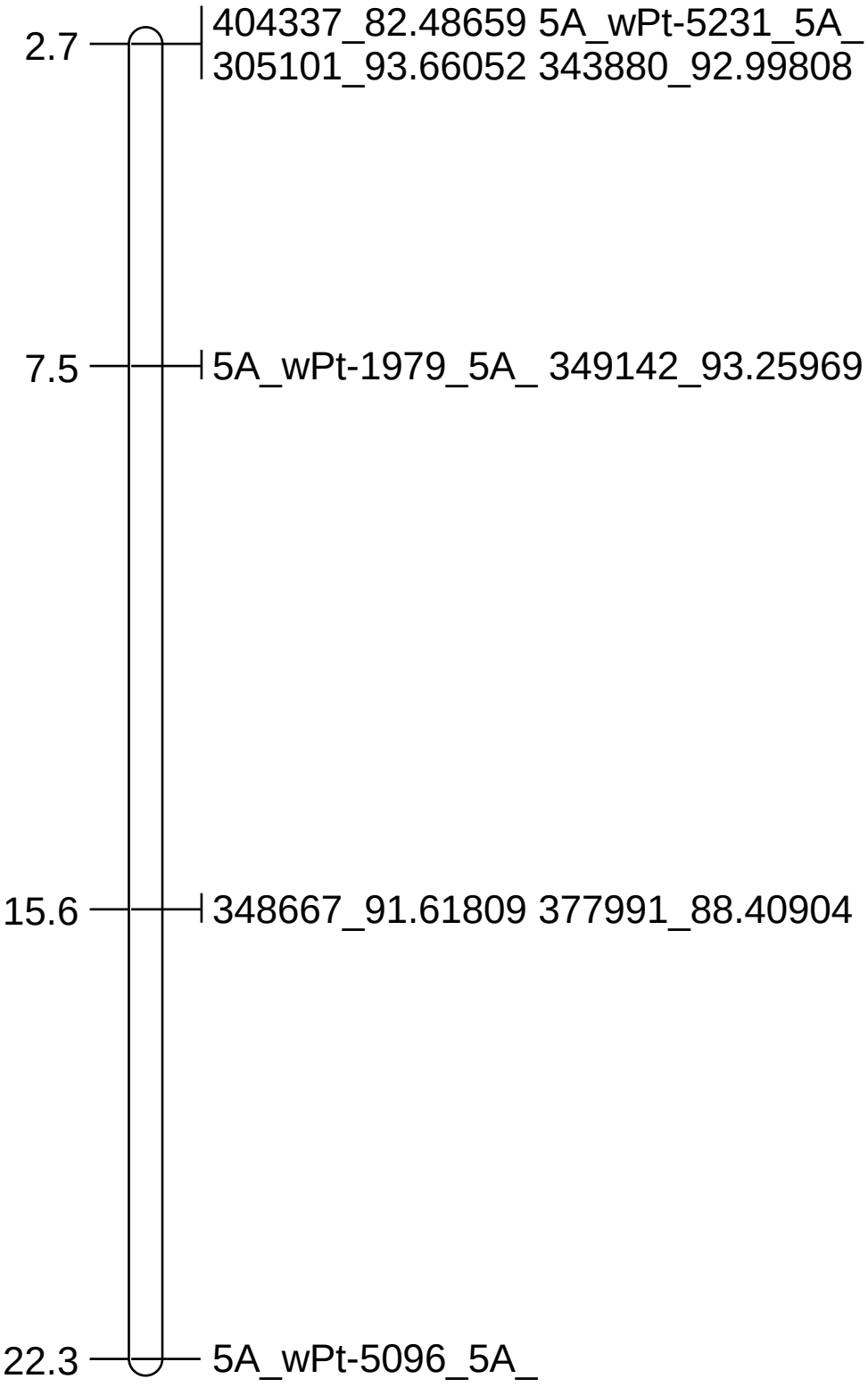
4B


0.1
4B_wPt-1931_4B_ 4B_wPt-8756_4B_
0.6
wPt-9192_84.811 wPt-1029_89.478
2.6
wPt-2993_86.572 377964_97.10124
3.8
4B_wPt-1272_4B_
5.2
305571_83.16289 403448_88.91273
5.8
304771_91.40008 wPt-3201_93.713

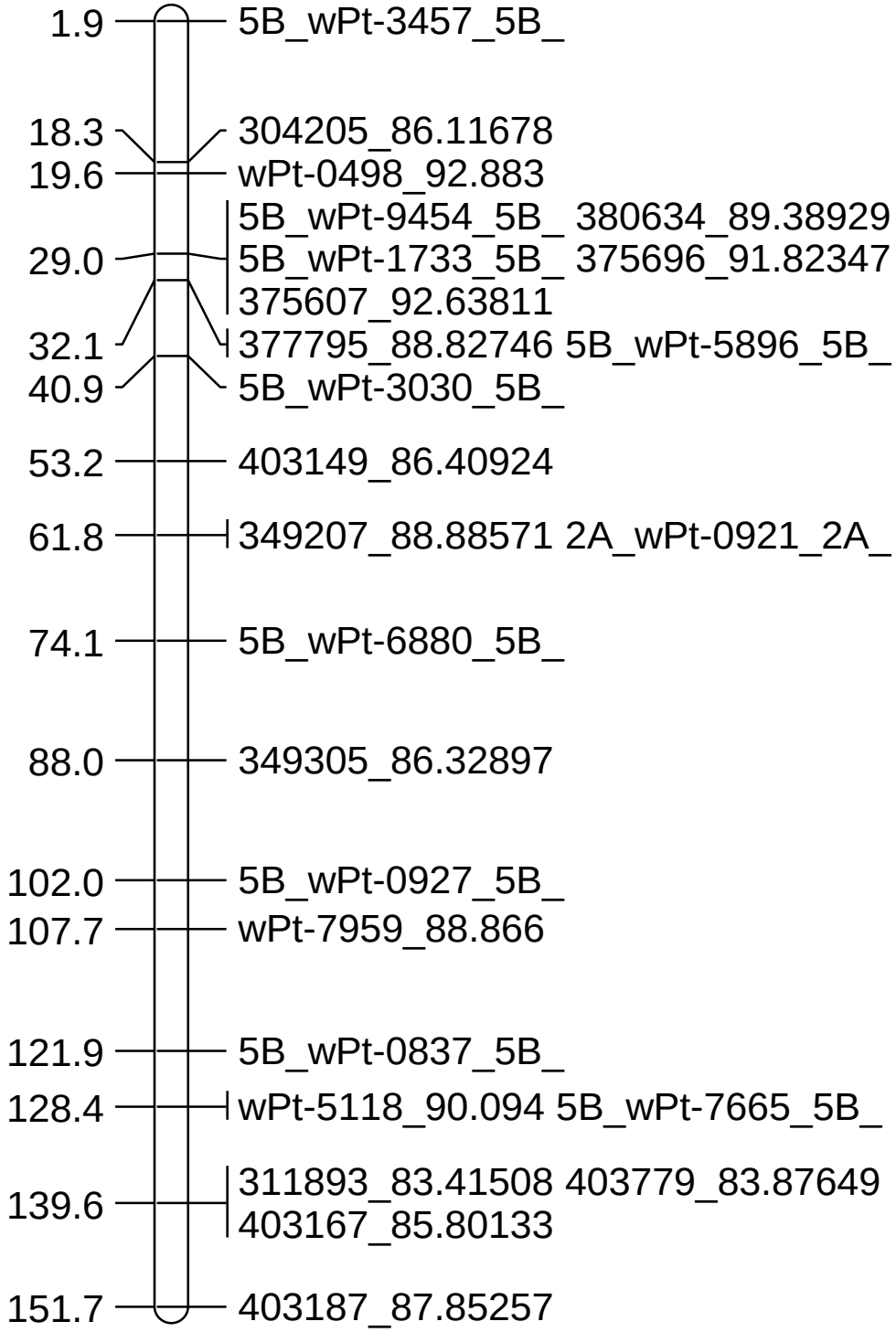
4B


4.9
5.5
6.1
8.8
12.0
17.2
37.4
53.2
74.0
75.8
344277_83.32905
2A_wPt-5887_2A_
378566_86.25103 403931_92.60627
349352_90.64485
wPt-0872_84.642 304407_92.91843
345668_92.42715
4B_wPt-5334_4B_
305166_90.07216 404300_89.20899
403416_88.26876
4B_wPt-6209_4B_ 344054_92.38085
306806_92.94845

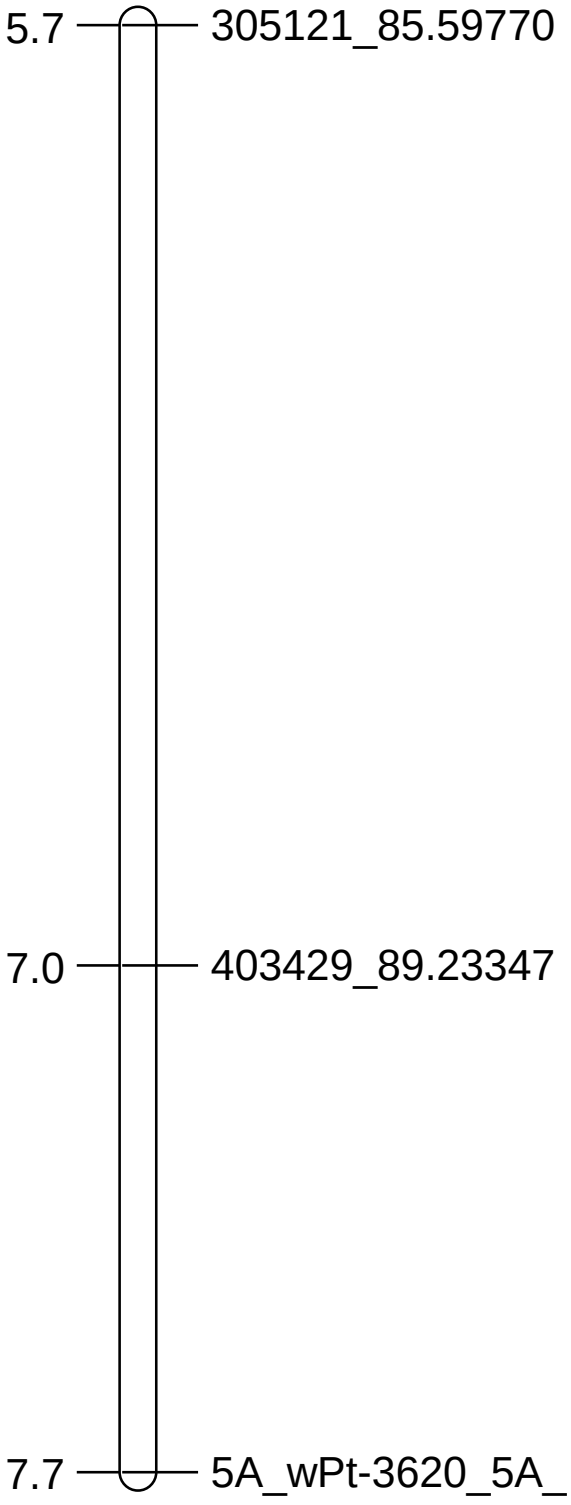
5A


2.7
404337_82.48659 5A_wPt-5231_5A_
305101_93.66052 343880_92.99808
7.5
5A_wPt-1979_5A_ 349142_93.25969
15.6
348667_91.61809 377991_88.40904
22.3
5A_wPt-5096_5A_

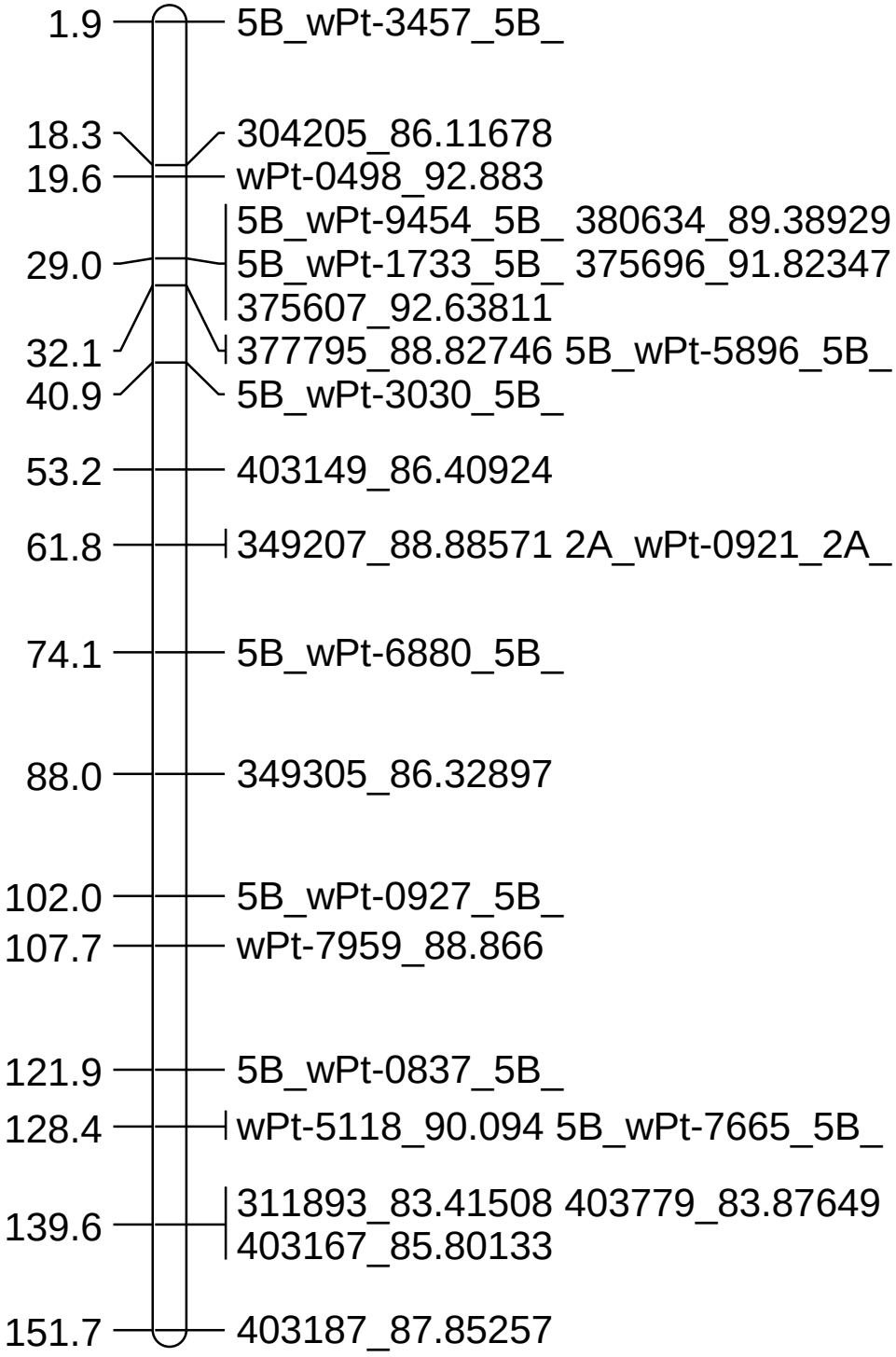
5B


1.9
5B_wPt-3457_5B_
18.3
304205_86.11678
19.6
wPt-0498_92.883
5B_wPt-9454_5B_ 380634_89.38929
29.0
5B_wPt-1733_5B_ 375696_91.82347
375607_92.63811
32.1
377795_88.82746 5B_wPt-5896_5B_
40.9
5B_wPt-3030_5B_
53.2
403149_86.40924
61.8
349207_88.88571 2A_wPt-0921_2A_
74.1
5B_wPt-6880_5B_
88.0
349305_86.32897
102.0
5B_wPt-0927_5B_
107.7
wPt-7959_88.866
121.9
5B_wPt-0837_5B_
128.4
wPt-5118_90.094 5B_wPt-7665_5B_
311893_83.41508 403779_83.87649
139.6
403167_85.80133
151.7
403187_87.85257

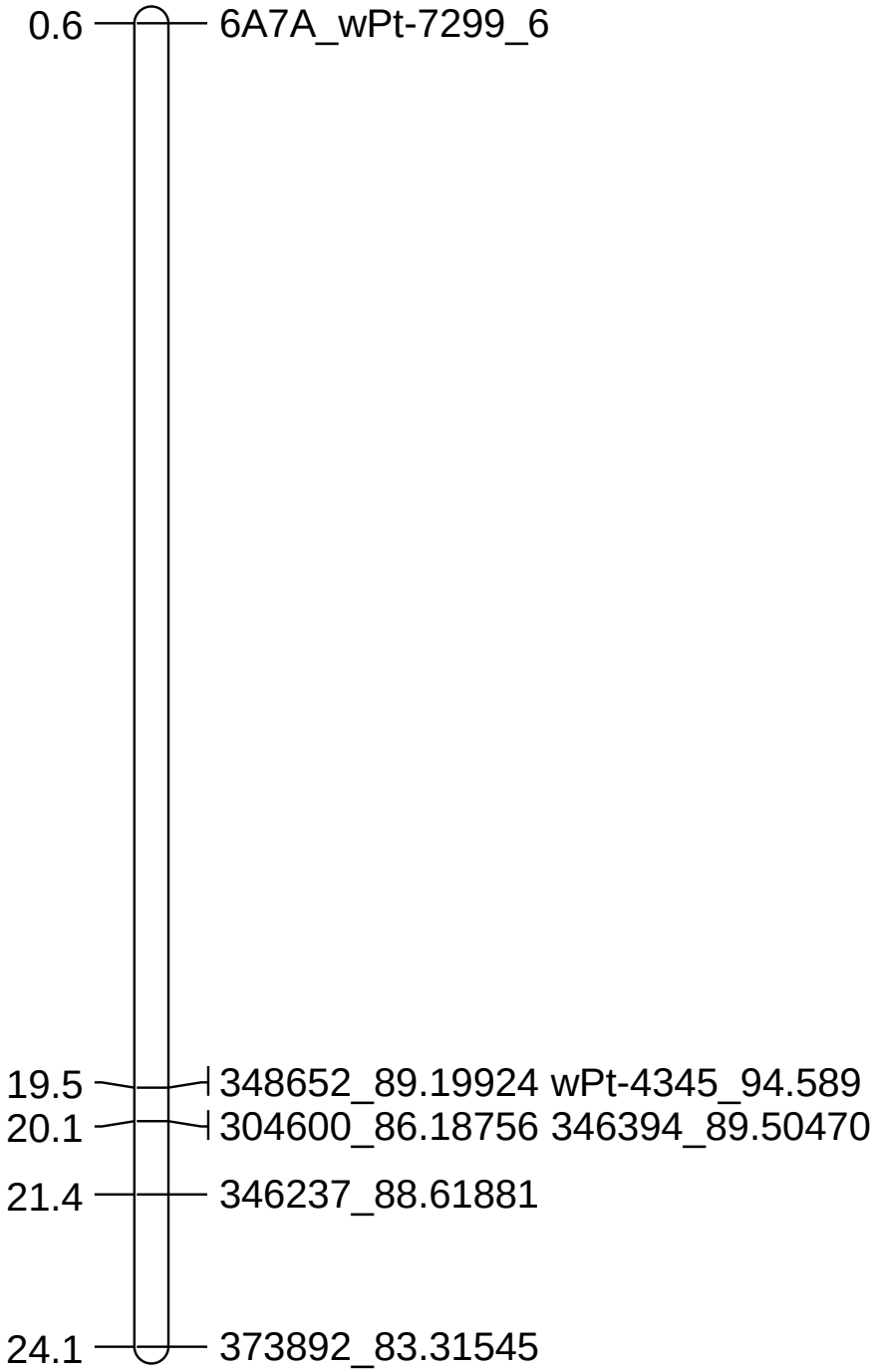
5A



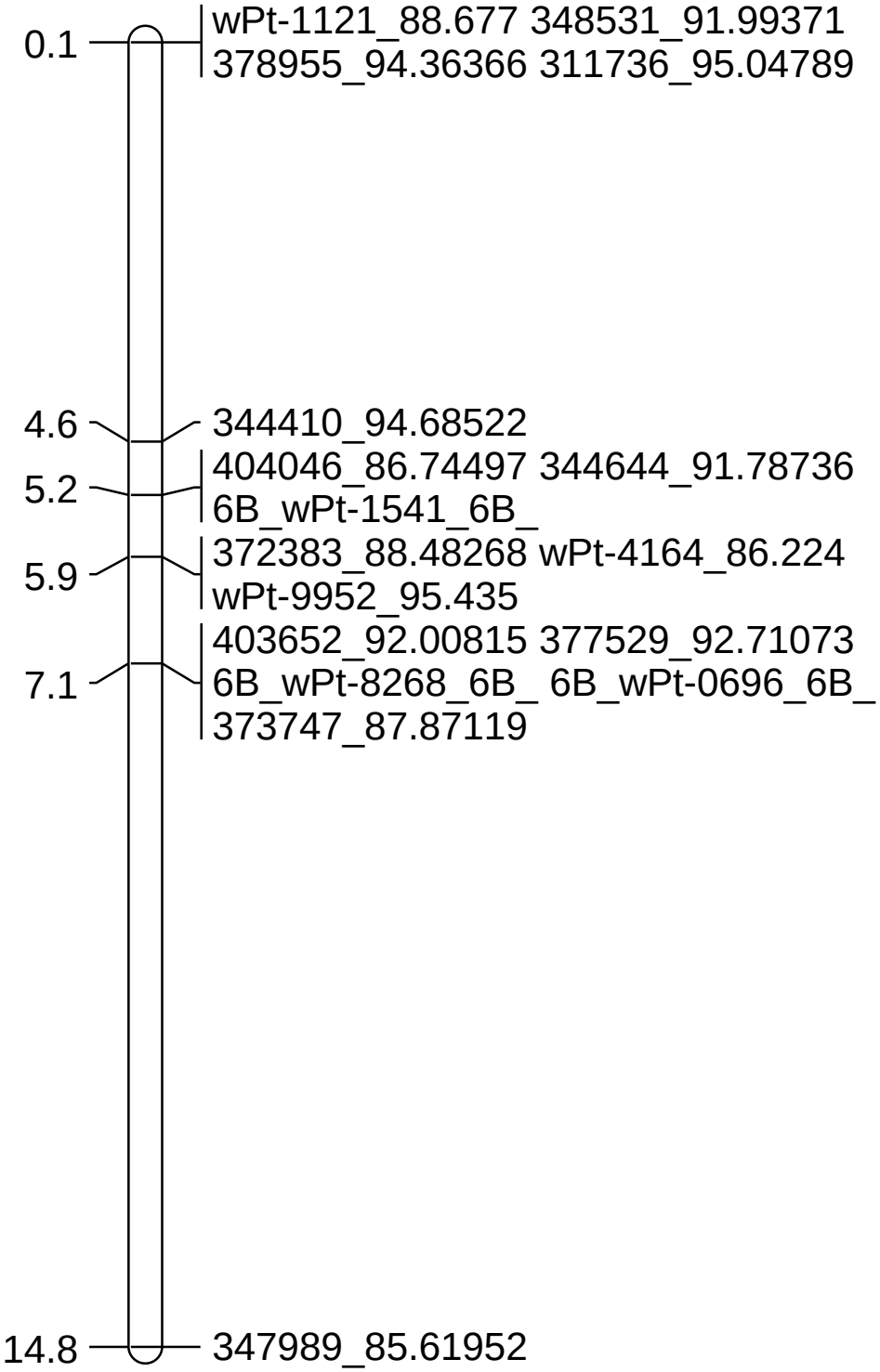
5B


1.9
5B_wPt-3457_5B_
18.3
304205_86.11678
19.6
wPt-0498_92.883
5B_wPt-9454_5B_ 380634_89.38929
29.0
5B_wPt-1733_5B_ 375696_91.82347
375607_92.63811
32.1
377795_88.82746 5B_wPt-5896_5B_
40.9
5B_wPt-3030_5B_
53.2
403149_86.40924
61.8
349207_88.88571 2A_wPt-0921_2A_
74.1
5B_wPt-6880_5B_
88.0
349305_86.32897
102.0
5B_wPt-0927_5B_
107.7
wPt-7959_88.866
121.9
5B_wPt-0837_5B_
128.4
wPt-5118_90.094 5B_wPt-7665_5B_
139.6
311893_83.41508 403779_83.87649
403167_85.80133
151.7
403187_87.85257

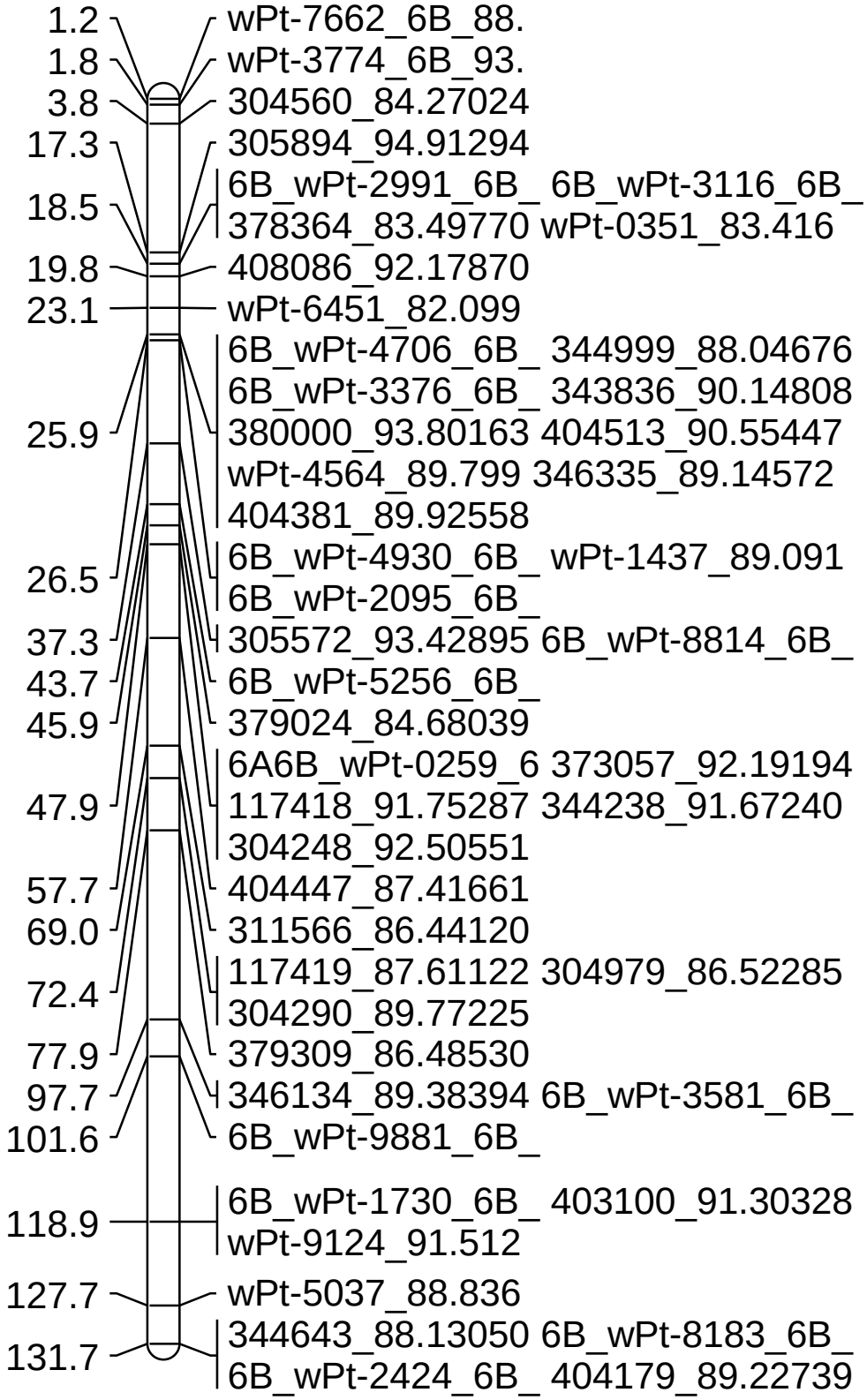
6A/7A


0.6
6A7A_wPt-7299_6
19.5
348652_89.19924 wPt-4345_94.589
20.1
304600_86.18756 346394_89.50470
21.4
346237_88.61881
24.1
373892_83.31545

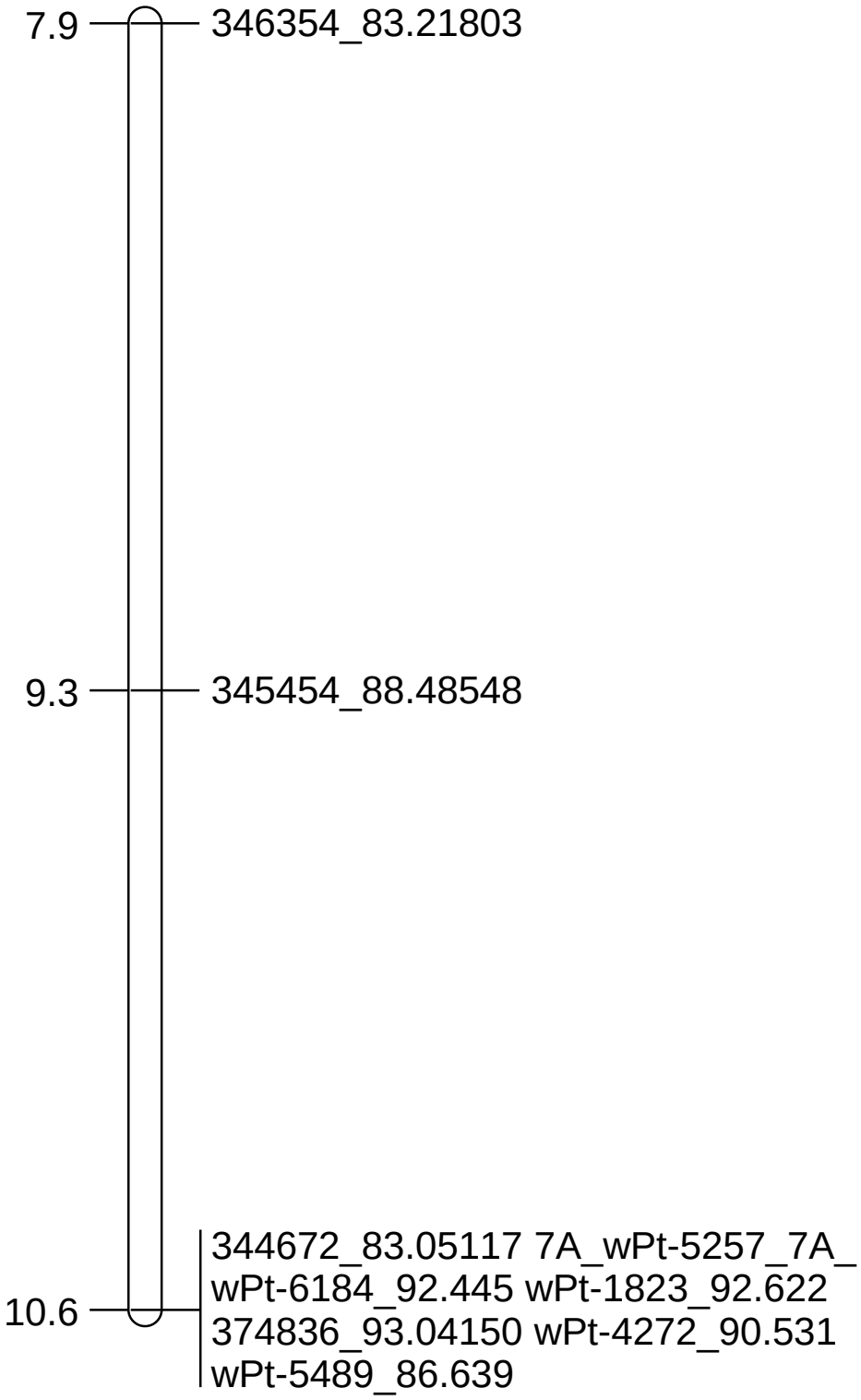
6B


0.1
wPt-1121_88.677 348531_91.99371
378955_94.36366 311736_95.04789
4.6
344410_94.68522
5.2
404046_86.74497 344644_91.78736
6B_wPt-1541_6B_
5.9
372383_88.48268 wPt-4164_86.224
wPt-9952_95.435
7.1
403652_92.00815 377529_92.71073
6B_wPt-8268_6B_ 6B_wPt-0696_6B_
373747_87.87119
14.8
347989_85.61952

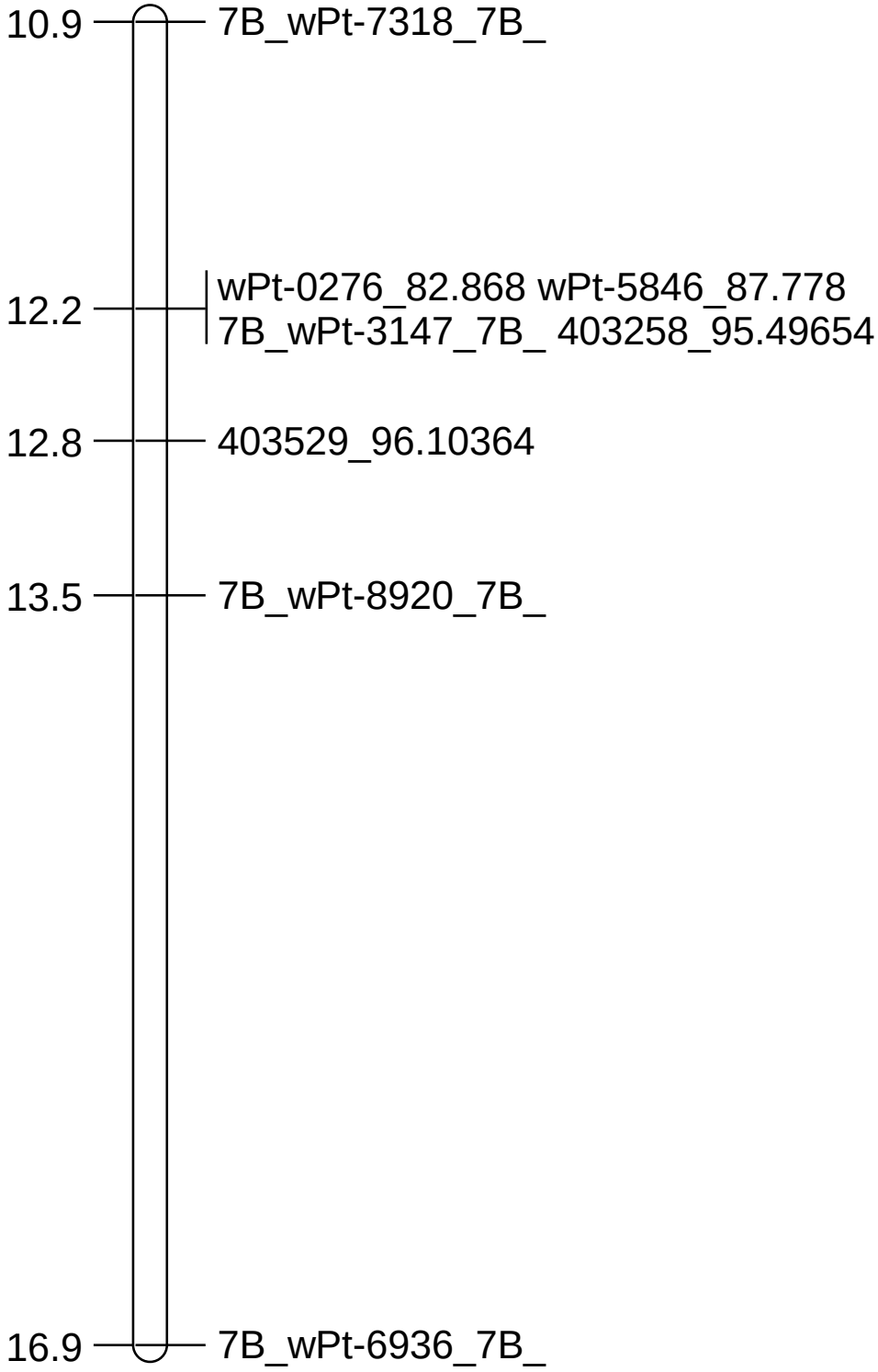
6B


1.2
1.8
3.8
17.3
18.5
19.8
23.1
25.9
26.5
37.3
43.7
45.9
47.9
57.7
69.0
72.4
77.9
97.7
101.6
118.9
127.7
131.7
wPt-7662_6B_88.
wPt-3774_6B_93.
304560_84.27024
305894_94.91294
6B_wPt-2991_6B_ 6B_wPt-3116_6B_
378364_83.49770 wPt-0351_83.416
408086_92.17870
wPt-6451_82.099
6B_wPt-4706_6B_ 344999_88.04676
6B_wPt-3376_6B_ 343836_90.14808
380000_93.80163 404513_90.55447
wPt-4564_89.799 346335_89.14572
404381_89.92558
6B_wPt-4930_6B_ wPt-1437_89.091
6B_wPt-2095_6B_
305572_93.42895 6B_wPt-8814_6B_
6B_wPt-5256_6B_
379024_84.68039
6A6B_wPt-0259_6 373057_92.19194
117418_91.75287 344238_91.67240
304248_92.50551
404447_87.41661
311566_86.44120
117419_87.61122 304979_86.52285
304290_89.77225
379309_86.48530
346134_89.38394 6B_wPt-3581_6B_
6B_wPt-9881_6B_
6B_wPt-1730_6B_ 403100_91.30328
wPt-9124_91.512
wPt-5037_88.836
344643_88.13050 6B_wPt-8183_6B_
6B_wPt-2424_6B_ 404179_89.22739

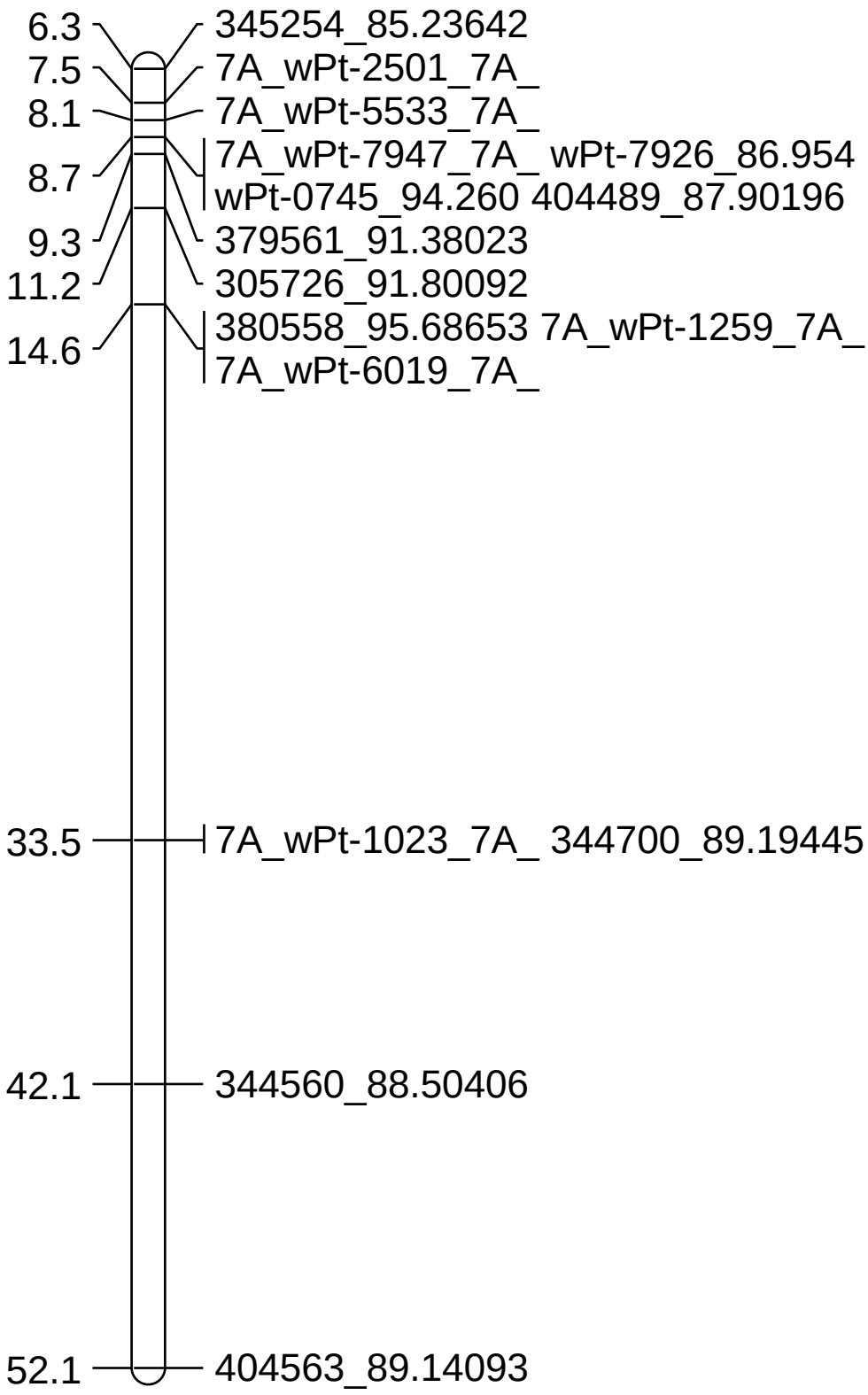
7A


7.9
346354_83.21803
9.3
345454_88.48548
10.6
344672_83.05117 7A_wPt-5257_7A_
wPt-6184_92.445 wPt-1823_92.622
374836_93.04150 wPt-4272_90.531
wPt-5489_86.639

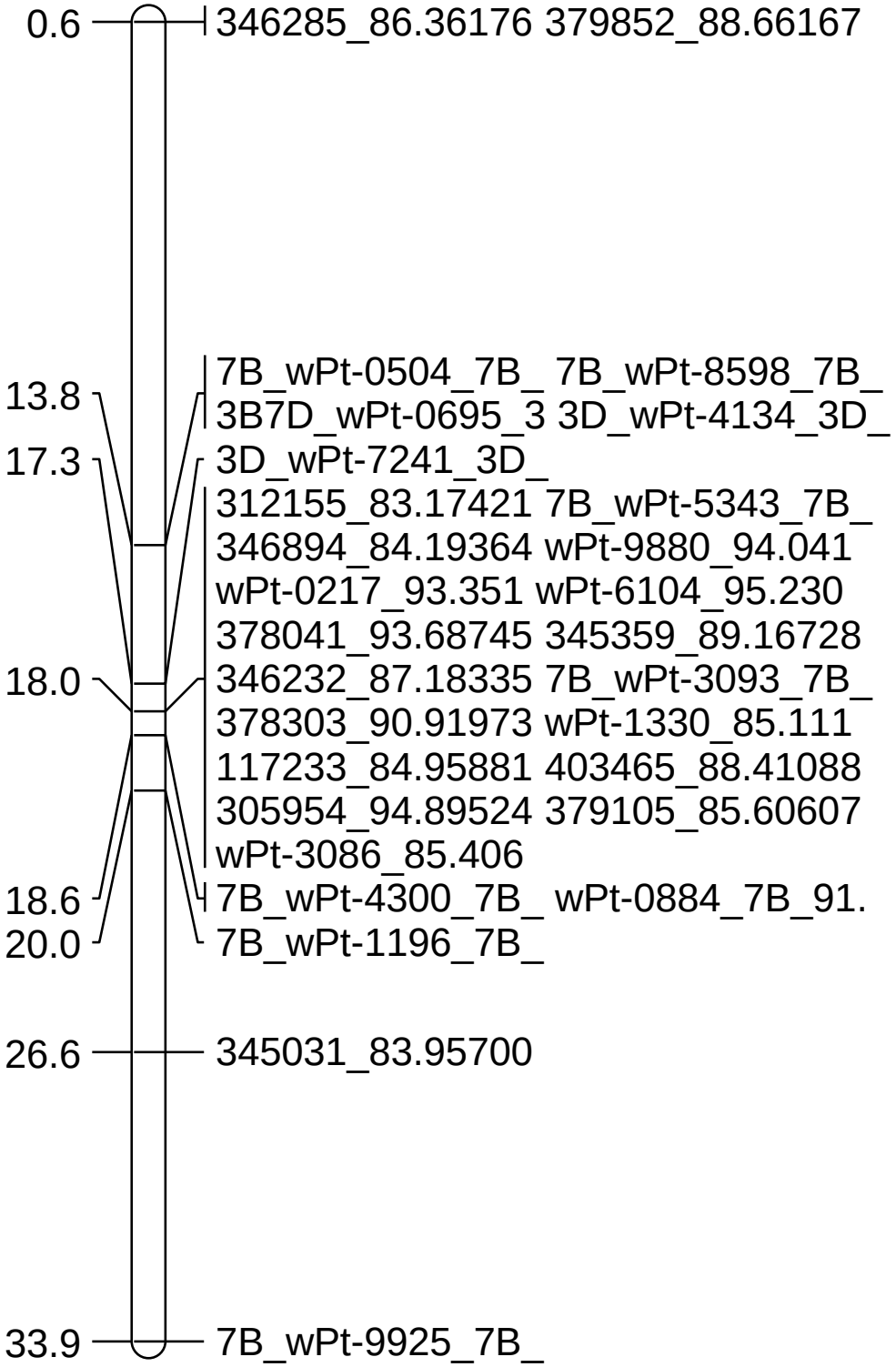
7B


10.9
7B_wPt-7318_7B_
12.2
wPt-0276_82.868 wPt-5846_87.778
7B_wPt-3147_7B_ 403258_95.49654
12.8
403529_96.10364
13.5
7B_wPt-8920_7B_
16.9
7B_wPt-6936_7B_

7A


6.3
7.5
8.1
8.7
9.3
11.2
14.6
33.5
42.1
52.1
345254_85.23642
7A_wPt-2501_7A_
7A_wPt-5533_7A_
7A_wPt-7947_7A_ wPt-7926_86.954
wPt-0745_94.260 404489_87.90196
379561_91.38023
305726_91.80092
380558_95.68653 7A_wPt-1259_7A_
7A_wPt-6019_7A_
7A_wPt-1023_7A_ 344700_89.19445
344560_88.50406
404563_89.14093

7B


0.6
346285_86.36176 379852_88.66167
13.8
7B_wPt-0504_7B_ 7B_wPt-8598_7B_
3B7D_wPt-0695_3 3D_wPt-4134_3D_
17.3
3D_wPt-7241_3D_
312155_83.17421 7B_wPt-5343_7B_
346894_84.19364 wPt-9880_94.041
wPt-0217_93.351 wPt-6104_95.230
378041_93.68745 345359_89.16728
18.0
346232_87.18335 7B_wPt-3093_7B_
378303_90.91973 wPt-1330_85.111
117233_84.95881 403465_88.41088
305954_94.89524 379105_85.60607
wPt-3086_85.406
18.6
7B_wPt-4300_7B_ wPt-0884_7B_91.
20.0
7B_wPt-1196_7B_
26.6
345031_83.95700
33.9
7B_wPt-9925_7B_

Cited Literature

Chapter Eight

8.0 Cited Literature

[1] Botstein, D., White, R. L., Skalnick, M. H., and Davies, R. W. (1980), Construction of a genetic linkage map in man using restriction fragment length polymorphism, Am. J. Hum. Genet. 32, 314_331.

[2] Abrol, IP, Yadav, JSP and Massoud, FI, (1988) "Salt-affected Soils and their management", FAO Soilds Bulletin 39, Food and Agriculture organisation of The United Nations.

[3] Briggle, L.W., Curtis, B.C. (1987). Wheat Worldwide. In "Wheat and Wheat Improvement", EG Heyne, ed Ed. 2nd. American Society of AgronomyInc. Publishers (2 more societies), Madison, Wisconsin USA. pp 4-31.

[4] T.G. Reeves et al., (2002), "History and past achievements of plant breeding", Aust. J. Agric. Res., 53, 851-863.

[5] T.G. Reeves et al., (2002), "History and past achievements of plant breeding", Aust. J. Agric. Res., 53, 851-863.

[6] Simmonds, D.H. (1989). Wheat and wheat quality in Australia. CSIRO Australia, Queensland.

[7] Academy of Technological Sciences and Engineering, (2002), "*Technology in Australia 1788-1988*", Online edn, Australian Science and Technology Heritage Centre, Melbourne, www.austehc.unimelb.edu.au/tia/scripts/tiadynindex.php3?EID=P000401.

[8] Academy of Technological Sciences and Engineering, (2002), "*Technology in Australia 1788-1988*", Online edn, Australian Science and Technology Heritage Centre, Melbourne, www.austehc.unimelb.edu.au/tia/scripts/tiadynindex.php3?EID=P000401.

[9] United States Department of Agriculture, Foreign Agricultural Service, www.fas.usda.gov, accessed 22/03/2004.

[10] Australian Bureau of Agricultural and Resource Economics, Agricultural Commodities, Australia, (2010), "7121.0 – Agricultural Commodities, Australia, 2009-10", www.abs.gov.au.

[11] Australian Bureau of Agricultural and Resource Economics, Agricultural Commodities, Australia, (2010), "7121.0 – Agricultural Commodities, Australia, 2009-10", www.abs.gov.au.

[12] Australian Bureau of Agricultural and Resource Economics, Agricultural Commodities, Australia, (2010), "7121.0 – Agricultural Commodities, Australia, 2009-10", www.abs.gov.au.

[13] E.S. Lagudah *et al*. (2001), Plant Physiol. Biochem. 39, 335-344.

[14] Hare, R, (2006). Agronomy of the durum wheats Kamilarui, Yallaroi, Wollaroi and EGA Bellaroi. In “Primefacts 140”, New South Wales Department of Primary Industries, pp 1-8.

[15] Hare, R, (2006). Agronomy of the durum wheats Kamilarui, Yallaroi, Wollaroi and EGA Bellaroi. In “Primefacts 140”, New South Wales Department of Primary Industries, pp 1-8.

[16] Evans, L.T., Bingham, J., Fischer, R.A. (1975). Wheat. In "Crop Physiology" Ed. Evans, L. T. Cambridge Univ Press, Cambridge, pp 101-149.

[17] Simmons, S.R. (1987). Growth, development, and Physiology. In "Wheat and Wheat Improvement", EG Heyne, ed Ed. 2. ASA Inc, CSSA, Inc and SSS of America Inc., Madison Wisconsin, USA. pp 77-104.

[18] Klepper, B., Belford, R.K., Rickman, R.W. (1984), “Root and shoot development in winter wheat”, Crop Science 76: 117-122.

[19] New South Wales Department of Primary Industries (2011), http://www.dpi.nsw.gov.au/agriculture.

[20] Australian Bureau of Agricultural and Resource Economics (ABARE), (2009), “Australian Crop Report 151”, www.abareconomics.com/interactive/09acr_sept.

[21] Australian Wheat Board, (2004), “Wheat Statistics”, www.health.gov.au/internet/ogtr/publishing.nsf/Content/wheat3/$FILE/biologywheat.pdf

[22] The Australian Bureau of Agricultural and Resource Economics (ABARE), (2010), “Australians Crop Report Number 152”, www.abare.gov.au/interactive/09acr_dec/excel/h_aus.xls

[23] Kirby, E.J.M. (1983), Developoment of the cereal plant. In "The yield of cereals", WW D, ed. Royal Agriculture society of England, London. pp 1-3.

[24] Kirby, E.J.M. (1983), Developoment of the cereal plant. In "The yield of cereals", WW D, ed. Royal Agriculture society of England, London. pp 1-3.

[25] Patrick, J.W. (1972), “Distribution of assimilate during stem elongation in wheat”, Australian Journal of Biological Sciences 25: 455-467.

[26] Austin, R.B., Jones, H.G. (1975), “The Physiology of wheat: In Plant Breeding”, Institute Annual Report-1974. Cambridge UK, Cambridge. pp 20-73.

[27] Austin, R.B., Jones, H.G. (1975), “The Physiology of wheat: In Plant Breeding”, Institute Annual Report-1974. Cambridge UK, Cambridge. pp 20-73.

[28] Kirby, E.J.M. (1983), Developoment of the cereal plant. In "The yield of cereals", WW D, ed. Royal Agriculture society of England, London. pp 1-3.

[29] Slafer, G.A. & Rawson, H.M. **(**1994), "Sensitivity of wheat phasic development to major environmental factors: a reexamination of some assumptions made by physiologists and modelers", Austr. J. Plant Physiol., 21: 393-426.

[30] Acevedo, E., Silva, P., Silva, H., (2002), Wheat growth and physiology, in "Bread Wheat. Improvement and Production", BC Curtis, S Rajaram, H Gomez MacPherson, eds. Food and Agriculture Orginisation, Rome, Italy.

[31] R.A. McIntosh et al., (2000), "Catalogue of gene symbols for wheat", http://wheat.pw.usda.gov/ggpages/wgc/2000upd.html.

[32] E Martinez-Perez, P Shaw, S Reader, L Aragon-Alcaide, T Miller and G Moore, (1999), "Homologous chromosome pairing in wheat", Journal of Cell Science, Vol 112, Issue 11 1761-1769.

[33] E Martinez-Perez, P Shaw, S Reader, L Aragon-Alcaide, T Miller and G Moore, (1999), "Homologous chromosome pairing in wheat", Journal of Cell Science, Vol 112, Issue 11 1761-1769.

[34] E Martinez-Perez, P Shaw, S Reader, L Aragon-Alcaide, T Miller and G Moore, (1999), "Homologous chromosome pairing in wheat", Journal of Cell Science, Vol 112, Issue 11 1761-1769.

[35] E Martinez-Perez, P Shaw, S Reader, L Aragon-Alcaide, T Miller and G Moore, (1999), "Homologous chromosome pairing in wheat", Journal of Cell Science, Vol 112, Issue 11 1761-1769.

[36] Human Genome Project, "Functional and Comparative Genomics Facts Sheet", www.ornl.gov, accessed November 2006.

[37] Kimber G and Sears ER. (1987), "Evolution in the genus *Triticum* and the origin of cultivated wheat". In: **"**Wheat and wheat improvement", 2nd Ed (Heyne, E.G. ed.). Ameican Society of Agronomy, Madison, WI. pp. 154-164.

[38] Kimber, G. and K. Tsunewaki, (1988), "Genome symbols and plasma types in the wheat group", Pp. 1209-1211 in T.E. Miller and R.M. D. Koebner (eds.), Proc. Seventh International Wheat Symposium, Cambridge, 1987.

[39] McFadden, E. S., and E. R. Sears, (1946), "The origin of *Triticum spelta* and its free-threshing hexaploid relatives", J. Hered. 37: 81–89.

[40] Kimber G and Sears ER. (1987), "Evolution in the genus *Triticum* and the origin of cultivated wheat". In: **"**Wheat and wheat improvement", 2nd Ed (Heyne, E.G. ed.). Ameican Society of Agronomy, Madison, WI. pp. 154-164.

[41] Friebe B et al. (1990) Transfer of hessian fly resistance from 'Chaupon' rye to hexaploid wheat via a 2BS-2RL wheat rye chromosome translocation. Theoretical and Applied Genetics 79:385-389.

[42] P. Sarkar, G. L. Stebbins, (1956), "Morphological Evidence Concerning the Origin of the B Genome in Wheat", American Journal of Botany, Vol. 43, No. 4, pp. 297-304.

[43] Gill, SB., Friebe, B., and Endo, TR., (1991), "Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*)" Genome, 34: 830-839.

[44] Department of Health and Ageing, Office of the Gene Technology Regulator, Australian Government. (2005), "The biology and ecology of breat wheat (Triticum aestivum L. Em Thell) in Australia".

[45] Blumenthal CS., Bekes F., Bately IL., Wrigley CW., Moss HJ., Mares DJ., Barlow EWR., (1991) "Interpretation of grain quality results from wheat variety trails with reference to high temperature stress", Australian Journal of Agricultural Research 42(3) 325-334.

[46] Gibson LR., Paulsen GM., (1999), "Yield components of Wheat grown under high temperature stress during reproductive growth", Crop Science, 39:1841-1846.

[47] Ford GW, Martin JJ, Rengasamy P, Boucher SC, Ellington A, (1993), 'Soil sodicity in Victoria', Australian Journal of Soil Research, 31, 869-909.

[48] Francois LE, Maas EV, Donovan TJ, Youngs VL, (1986), 'Effect of salinity on grain yield and quality, vegetative growth, and germination of semi-dwarf and durum wheats', Agronomy Journal, 78, 1053-1058.

[49] Greenway H and Munns R, (1980), 'Mechanisms of salt tolerance in non-halophytes', Annual Reviews of Plant Physiology, 31, 149-190.

[50] Shah SH, Gorham J, Forester BP, Wyn Jones RG, (1987), 'Salt tolerance in the Triticeae: the contribution of the D genome to cation selectivity in hexaploid wheat', Journal of Experimental Botany, 38, 254-269.

[51] Gorham J, Hardy C, Wyn Jones RG, Joppa LR, Law CN, (1987), 'Chromosomal location of a K/Na discrimination character in the D genome of wheat', Theoretical and Applied Genetics, 74, 584-588.

[52] Dubcovsky J, Santa Maris G, Epstein E, Luo MC, Dvorak J, (1996), 'Mapping of the K=/Na+ discrimination locus *Kna*1 in wheat', Theoretical Applied Genetics, 2, 448-454.

[53] Munns, R., and James, R.A., (2003), 'Screening methods for salinity tolerance: a case study with tetraploid wheat', Plant and Soil, 253:201-218.

[54] Srivastava, J.P., and Jana, S., (1984), 'Screening wheat and barley germplasm for salt tolerance. *In* Salinity tolerance in plants: Stratergies for crop improvement', Ed. Staples, R.C., and Toenniessen, G.H., pp. 273-283.

[55] Francois, L.E., Mass, E.V., Donovan, T.J. and Youngs, V.L., (1986), 'Effects of salinity on grain yield and quality, vegetative growth and germination of semi-dwarf and durum wheat' Agron. J. 78,1053-1058.

[56] Gorham, J., Hardy, C., Wyn Jones, R.G., Joppa, L.R. and Law, C.N., (1987) 'Chromosomal loation of a K/Na discrimination character in the D genome of wheat', Theor. Appl. Genet., 74, 584-588.

[57] Anderson GR, Papa D, Peng J, Tahir M, Lapitan NL., (2003), "Genetic mapping of Dn7, a rye gene conferring resistance to the Russian wheat aphid in wheat", Theor Appl Genet. Jul 19.

[58] Brown GN (1993) A seedling marker for gene Sr2 in wheat. In 'Proceedings of the Tenth Australian Plant Breeding Conference, Volume two. (Eds BC Imrie and JB Hacker) pp. 139-140 (Conference Organising Committee: Cold Coast, QLD)

[59] H.A. Eagles et al. (2001) 'Implementation of marker in Australian wheat breeding' Aust. J. Agric. Res. 52, 1349-1356.

[60] Dholakia BB, Ammiraju JS, Santra DK, Singh H, Katti MV, Lagu MD, Tamhankar SA, Rao VS, Gupta VS, Dhaliwal HS, Ranjekar PK, (2001), "Molecular marker analysis of protein content using PCR-based markers in wheat", Biochem Genet. Oct;39 (9-10):325-38.

[61] Bhattramakki D, Dolan M, Hanafey M, Wineland R, Vaske D, Register JC 3rd, Tingey SV, Rafalski A., (2002), "Insertion-deletion polymorphisms in 3' regions of maize genes occur frequently and can be used as highly informative genetic markers", Plant Mol Biol. Mar-Apr;48(5-6):539-47.

[62] Spielmeyer W, Lagudah S., (2003), "Homoeologous set of NBS-LRR genes located at leaf and stripe rust resistance loci on short arms of chromosome 1 of wheat", Funct Integr Genomics. Mar;3(1-2):86-90.

[63] Boyett, V., Agrama, H., Gibbons, J., Bryant, R.J., Eizenga, G.C., Moldenhauer, K.K. (2007), "Cooking quality and blast disease resistance linked markers: Genotyping a working rice germplasm collection for future marker-assisted breeding applications", Plant and Animal Genome XV. p. 150, www.intl-pag.org/15/abstracts/PAG15_P03f_197.html.

[64] Boyett, V., Agrama, H., Gibbons, J., Bryant, R.J., Eizenga, G.C., Moldenhauer, K.K. (2007), "Cooking quality and blast disease resistance linked markers: Genotyping a working rice germplasm collection for future marker-assisted breeding applications", Plant and Animal Genome XV. p. 150, www.intl-pag.org/15/abstracts/PAG15_P03f_197.html.

[65] Botstein D, White RL, Skolnick M, Davis RW., (1980), "Construction of a genetic linkage map in man using restriction fragment length polymorphisms," Am J Hum Genet. May;32(3):314-31.

[66] CSIRO, (2011), "Wheat Markers", www.csiro.au/science/wheatmarkers.

[67] Botstein D, White RL, Skolnick M, Davis RW., (1980), “Construction of a genetic linkage map in man using restriction fragment length polymorphisms,” Am J Hum Genet. May;32(3):314-31.

[68] Weber JL, May PE., (1989) “Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction,” Am J Hum Genet. 1989 Mar;44(3):388-96.

[69] Weber JL, May PE., (1989) “Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction,” Am J Hum Genet. 1989 Mar;44(3):388-96.

[70] Weber JL, May PE., (1989) “Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction,” Am J Hum Genet. 1989 Mar;44(3):388-96.

[71] Weber JL, May PE., (1989) “Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction,” Am J Hum Genet. 1989 Mar;44(3):388-96.

[72] Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto E., Hoff, M., Kumlin, E., and White, R. (1987). “Variable number tandem repeat (VNTR) markers for human gene mapping”, Science, 235,1616-1622.

[73] Jeffreys, A. J., Wilson, V., and Thein, S. L. (1985). “Hypervariable "minisatellite" regions in human DNA”, Nature, 314,67-73.

[74] Litt, M. and Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am. J. Hum. Genet. 44,397-401.

[75] Edwards, A., Civitello, A., Hammond, H. A., and Caskey, C. T. (1991). DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am. J. Hum. Genet. 49,746-756.

[76] JG Williams, AR Kubelik, KJ Livak, JA Rafalski and SV Tingey, (1980) “A polymorphisms amplified by arbitrary primers are useful as genetic markers,” Central Research and Development Department, E.I. du Pont de Nemours & Co., Inc, Wilmington, DE.

[77] Hahn, R.C., Macedo, A.M., Fernandes Fontes, C.J., Batista, R.D., Santos, N.L., Hamdan, J.S., (2003) “Randomly Amplified Polymorphic DNA as a Valuable Tool for Epidemiological Studies of Paracoccidioides brasiliensis,” J Clin Microbiol. Jul;41(7):2849-54.

[78] Hahn, R.C., Macedo, A.M., Fernandes Fontes, C.J., Batista, R.D., Santos, N.L., Hamdan, J.S., (2003) “Randomly Amplified Polymorphic DNA as a Valuable Tool for Epidemiological Studies of Paracoccidioides brasiliensis,” J Clin Microbiol. Jul;41(7):2849-54.

[79] Hahn, R.C., Macedo, A.M., Fernandes Fontes, C.J., Batista, R.D., Santos, N.L., Hamdan, J.S., (2003) “Randomly Amplified Polymorphic DNA as a Valuable Tool for Epidemiological Studies of Paracoccidioides brasiliensis,” J Clin Microbiol. Jul;41(7):2849-54.

[80] Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, M., Pot, J., Peleman, J., Kuiper, M., Zabeau, M., (1995) “AFLP: a new technique for DNA fingerprinting,” Nucleic Acids Res. Nov 11;23(21):4407-14.

[81] Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, M., Pot, J., Peleman, J., Kuiper, M., Zabeau, M., (1995) “AFLP: a new technique for DNA fingerprinting,” Nucleic Acids Res. Nov 11;23(21):4407-14.

[82] Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, M., Pot, J., Peleman, J., Kuiper, M., Zabeau, M., (1995) “AFLP: a new technique for DNA fingerprinting,” Nucleic Acids Res. Nov 11;23(21):4407-14.

[83] Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, M., Pot, J., Peleman, J., Kuiper, M., Zabeau, M., (1995) “AFLP: a new technique for DNA fingerprinting,” Nucleic Acids Res. Nov 11;23(21):4407-14.

[84] Unrau P, Deugau KV., “Non-cloning amplification of specific DNA fragments from whole genomic DNA digests using DNA 'indexers',” Gene. 1994 Aug 5;145(2):163-9.

[85] Unrau P, Deugau KV., (1994), “Non-cloning amplification of specific DNA fragments from whole genomic DNA digests using DNA 'indexers',” Gene. Aug 5;145(2):163-9.

[86] Unrau P, Deugau KV., (1994), “Non-cloning amplification of specific DNA fragments from whole genomic DNA digests using DNA 'indexers',” Gene. Aug 5;145(2):163-9.

[87] U Unrau P, Deugau KV., (1994), “Non-cloning amplification of specific DNA fragments from whole genomic DNA digests using DNA 'indexers',” Gene. Aug 5;145(2):163-9.

[88] Unrau P, Deugau KV., (1994), “Non-cloning amplification of specific DNA fragments from whole genomic DNA digests using DNA 'indexers',” Gene. Aug 5;145(2):163-9.

[89] Unrau P, Deugau KV., (1994), “Non-cloning amplification of specific DNA fragments from whole genomic DNA digests using DNA 'indexers',” Gene. Aug 5;145(2):163-9.

[90] Prashar, Y., Weissman, S.M., (1996) “Analysis of differential gene expression by display of 3' end restriction fragments of cDNAs,” Biochemistry, Vol. 93, Issue 2, 659-663, January 23, 1996.

[91] Prashar, Y., Weissman, S.M., (1996) “Analysis of differential gene expression by display of 3' end restriction fragments of cDNAs,” Biochemistry, Vol. 93, Issue 2, 659-663, January 23, 1996.

[92] Prashar, Y., Weissman, S.M., (1996) “Analysis of differential gene expression by display of 3' end restriction fragments of cDNAs,” Biochemistry, Vol. 93, Issue 2, 659-663, January 23, 1996.

[93] Prashar, Y., Weissman, S.M., (1996) “Analysis of differential gene expression by display of 3' end restriction fragments of cDNAs,” Biochemistry, Vol. 93, Issue 2, 659-663, January 23, 1996.

[94] Prashar, Y., Weissman, S.M., (1996) “Analysis of differential gene expression by display of 3' end restriction fragments of cDNAs,” Biochemistry, Vol. 93, Issue 2, 659-663, January 23, 1996.

[95] Langridge, P., Lagudah, E.S., Holton, T.A., Appeles, R., Sharp, P.J., Chalmers, K.J., (2001), “Trends in genetic and genome analyses in wheat: a review”, Aust. J. Agric. Res., 52, 1040-1077.

[96] Quarrie, S. A., Lazic-Jancic, V., Kovacevic, D., Steed, A. and Pekic, S., (1999) “Bulk segregant analysis with molecular markers and its use for improving drought resistance in maize”, Journal of Experimental Botany, Vol. 50, No. 337, pp. 1299-1306, August.

[97] Michelmore, RW, Paran, I, Kesseli, RV, (1991) ‘Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations’, Proc. Natl. Acad. Sci. USA, Vol 88, pp 9828-9832.

[98] Michelmore, RW, Paran, I, Kesseli, RV, (1991) ‘Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations’, Proc. Natl. Acad. Sci. USA, Vol 88, pp 9828-9832.

[99] Michelmore, RW, Paran, I, Kesseli, RV, (1991) 'Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations', Proc. Natl. Acad. Sci. USA, Vol 88, pp 9828-9832.

[100] Michelmore, RW, Paran, I, Kesseli, RV, (1991) 'Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations', Proc. Natl. Acad. Sci. USA, Vol 88, pp 9828-9832.

[101] Michelmore, RW, Paran, I, Kesseli, RV, (1991) 'Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations', Proc. Natl. Acad. Sci. USA, Vol 88, pp 9828-9832.

[102] Twardowska, M., Masojc, P., Milczarski, P., (2005), "Pyramiding genes affects sprouting resistance in Rye by means of marker assisted selection", Euphytica, Vol. 143(3), September, pp. 257-260.

[103] Browman, KW, (2004) "The Genomes of Recombinant Inbred Lines", Genetics 169:1133-1146.

[104] Tiwari VK, Rawat N, Chhuneja P, Neelam K, Aggarwal R, Randhawa GS, Dhaliwal HS, Keller B, Singh K., 2009, "Mapping of quantitative trait Loci for grain iron and zinc concentration in diploid A genome wheat", J Hered. Nov-Dec;100(6):771-6.

[105] Tiwari VK, Rawat N, Chhuneja P, Neelam K, Aggarwal R, Randhawa GS, Dhaliwal HS, Keller B, Singh K., (2009), "Mapping of quantitative trait Loci for grain iron and zinc concentration in diploid A genome wheat", J Hered. Nov-Dec;100(6):771-6.

[106] Tiwari VK, Rawat N, Chhuneja P, Neelam K, Aggarwal R, Randhawa GS, Dhaliwal HS, Keller B, Singh K., (2009), "Mapping of quantitative trait Loci for grain iron and zinc concentration in diploid A genome wheat", J Hered. Nov-Dec;100(6):771-6.

[107] Jaccoud, D., Peng, K., Feinstein, D., Kilian, A., (2001), "Diversity Arrays: a solid state technology for sequence information independent genotyping," Nucleic Acids Res. Feb 15;29(4).

[108] Jaccoud, D., Peng, K., Feinstein, D., Kilian, A., (2001), "Diversity Arrays: a solid state technology for sequence information independent genotyping," Nucleic Acids Res. Feb 15;29(4).

[109] Adams, R. L. P. and R. H. Burdon, (1985),. "Molecular biology of DNA methylation", Springer-Verlag, New York.

[110] Jablonka, E. and M. J. Lamb. (1995), "Epigenetic inheritance and evolution. The Lamarkian dimension", Oxford University Press, New York.

[111] Gruenbaum, Y., T. Naveh-Many, H. Cedar and A. Razin., (1981), “Sequence specificity of methylation in higher plant DNA”, Nature 292: 860-862.

[112] Vanyushin, B. F., S. G. Tkacheva and A. N. Belozerski., (1970), “Rare bases in animal DNA”, Nature 225: 948-949.

[113] Gruenbaum, Y., T. Naveh-Many, H. Cedar and A. Razin., (1981), “Sequence specificity of methylation in higher plant DNA”, Nature 292: 860-862.

[114] Coulondre, C., J. H. Miller, P. J. Farabaugh and W. Gilbert., (1978), “Molecular basis of base substitution hotspots in *Escherichia coli*”, Nature 274: 775-780.

[115] Josse, J., A. D. Kaiser and A. Kornberg., (1961), “Enzymatic synthesis of deoxyribonucleic acid VIII. Frequencies of nearest neighbor base sequences in deoxyribonucleic acid”, J. Biol. Chem. 236: 864-875.

[116] Russell, G. J., P. M. B. Walker, R. A. Elton and J. H. Subak-Sharpe., (1976), “Doublet frequency analysis of fractionated vertebrate nuclear DNA”. J. Mol. Biol. 108: 1-23.

[117] Gardiner-Garden, M., J. A. Sved and M. Frommer., (1992), “Methylation sites in angiosperm genes”, J. Mol. Evol. 34: 219-230.

[118] Leutwiler, L. S., B. R. Haigh-Evans and E. M. Meyerowitz., (1984), “The DNA of *Arabidopsis thaliana*”, Mol. Gen. Genet. 194: 15-23.

[119] Messeguer, R., M. W. Ganal, J. C. Steffens and S. D. Tanskley., (1991), “Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA”, Plant Mol. Biol. 16: 753-770.

[120] Bernacchia, G., A. Primo, L. Giorgetti, L. Pitto and R. Cella., (1998), “Carrot DNA-methyltransferase is encoded by two classes of genes with differing patterns of expression”, Plant J. 13: 317-329.

[121] Messeguer, R., M. W. Ganal, J. C. Steffens and S. D. Tanskley., (1991), “Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA”, Plant Mol. Biol. 16: 753-770.

[122] Razin, A. and T. Kafri., (1994), “DNA methylation from embryo to adult”, Prog. Nucl. Acid Res. Mol. Biol. 48: 53-81.

[123] Belanger, F. C. and A. G. Hepburn., (1990), “The evolution of CpNpG methylation in plants”, J. Mol. Evol. 30: 26-35.

[124] Marinus, M.G. and Morris, N.R. (1973) J. Bacteriol. 114, 1143-1150.

[125] Geier, G.E.and Modrich, P. (1979) J. Biol. Chem. 254, 1408-1413.

[126] Marinus, M.G. and Morris, N.R. (1973) J. Bacteriol. 114, 1143-1150.

[127] May, M.S. and Hattman, S.(1975) J. Bacteriol. 123, 768-770.

[128] Steele-Scott, N., T. A. Kavanagh and J. N. Timmis., (1984), "Methylation of rRNA genes in some higher plants", Plant Sci. Letters 35: 213-217.

[129] New England Biolabs Inc (2007) Restriction Endonuclease technical reference, www.neb.com/nebecomm/products/productR0106.asp, accessed 04/12/07.

[130] New England Biolabs Inc (2007) Restriction Endonuclease technical reference, www.neb.com/nebecomm/products/productR0106.asp, accessed 04/12/07.

[131] New England Biolabs Inc (2007) Restriction Endonuclease technical reference, www.neb.com/nebecomm/products/productR0106.asp, accessed 04/12/07.

[132] Cai, Q., C. L. Guy and G. A. Moore., (1996), "Detection of cytosine methylation and mapping of a gene influencing cytosine methylation in the genome of Citrus", Genome 39: 235-242.

[133] Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV., (1990), "DNA polymorphisms amplified by arbitrary primers are useful as genetic markers." Nucleic Acids Res. Nov 25;18(22):6531-5.

[134] Bedford, M. T. and P. P. Van Helden., (1990), "Analysis of allele-specific methylation", Trends Genet. 6: 140-141.

[135] Tsaftaris and Polidoros, Plant Breeding Reviews: 18:87-176.

[136] Church, G. M. and W. Gilbert., (1984), "Genomic sequencing", Proc. Nat. Acad. Sci. (USA) 81: 1991-1995.

[137] Hayashizaki, Y. and S. Watanabe (eds.), (1997), "Restriction landmark genomic scanning (RLGS)", Springer-Verlag, New York.

[138] Wagner, I. and I. Capesius., (1981), "Determination of 5-methylcytosine from plant DNA by high performance liquid chromatography", Biochim. Biophys. Acta 654: 52-56.

[139] Citti, L., P. G. Gervasi, G. Turchi and L. Mariani., (1983), "Improved method for determination of 5-methylcytosine by ion-pair reversed-phase high-performance liquid chromatography", J. Chromatog. 261: 315-319.

[140] Vilpo, J. A., S. Rasi, E. Suvato and L. M. Vilpo., (1986), "An improved radioimmunoassay for quantitation of DNA methylation", J. Immun. Methods. 87: 179-183.

[141] Jost, J.-P. and H.-P. Saluz., (1993), “DNA methylation: molecular biology and biological significance”, Brinkhauser Verlag, Boston.

[142] Spena, A., A. Viotti and V. Pirrotta., (1983), “Two adjacent genomic zein sequences: Structure, organization and tissue specific restriction pattern”, J. Mol. Biol. 169: 799-811.

[143] Bianchi, M. W. and A. Viotti., (1988), “DNA methylation and tissue-specific transcription of the storage proteins of maize”, Plant Mol. Biol. 11:

[144] Ngernprasirtsiri, J., R. Chollet, H. Kobayashi, T. Sugiyama and T. Akazawa., (1989), “DNA methylation and the differential expression of C4”, J Biol Chem. May 15;264(14):8241-8.

[145] Nick, H., B. Bowen, R. J. Ferl and W. Gilbert., (1986), “Detection of cytosine methylation in the maize alcohol dehydrogenase gene by genomic sequencing”, Nature 319: 243-246.

[146] Langdale, J. A., W. C. Taylor and T. Nelson., (1991), “Cell-specific accumulation of maize phosphoenolpyruvate carboxylase is correlated with demethylation at a specific site >3 kb upstream of the gene”, Mol. Gen. Genet. 225: 49-55.

[147] Adams, R. L. P. and R. H. Burdon., (1985), “Molecular biology of DNA methylation”, Springer-Verlag, New York.

[148] Weber, H., C. Ziechmann and A. Graessmann., (1990), “ In vitro DNA methylation inhibits gene expression in transgenic tobacco”, EMBO J. 9: 4409-4415.

[149] McClintock, B., (1957), “Genetic and cytological studies of maize”, Carnegie Inst. Wash. Yearb. 56: 393-401.

[150] McClintock, B., (1958), “The suppressor-mutator system of control of gene action in maize”, Carnegie Inst. Wash. Yearb. 57: 415-452.

[151] Peterson, P. A., (1965), “A relationship between the Spm and En control systems in maize”, Am. Nat. 44: 391-398.

[152] Peterson, P. A., (1986), “Mobile elements in maize: a force in evolutionary and plant breeding processes”, Stadler Genet. Symp. 17: 47-78.

[153] McClintock, B., (1959), “Genetic and cytological studies of maize”, Carnegie Inst. Wash. Yearb. 58: 415-429.

[154] McClintock, B., (1964), “Aspects of gene regulation in maize”, Carnegie Inst. Wash. Yearb. 63: 592-602.

[155] McClintock, B., (1965), “Components of action of the regulators Spm and Ac”, Carnegie Inst. Wash. Yearb. 64: 527-534.

[156] Flavell, R. B., (1994), “Inactivation of gene expression in plants as a consequence of novel sequence duplications”, Proc. Nat. Acad. Sci. (USA) 91: 3490-3496.

[157] Jorgensen, R. A., (1995), “Cosuppression, flower color patterns, and metastable gene expression states”, Science 268: 686-691.

[158] Matzke, M. A. and A. J. M. Matzke., (1995), ”How and why do plants inactivate homologous (trans)genes? Plant Physiol. 107: 679-685.

[159] Coulondre, C., J. H. Miller, P. J. Farabaugh and W. Gilbert., (1978), “Molecular basis of base substitution hotspots in *Escherichia coli*”, Nature 274: 775-780.

[160] Jones, P. A., W. M. Rideout, J.-C. Shen, C. H. Spruck and Y. C. Tsai., (1992), “Methylation, mutation and cancer”, BioEssays 14: 33-36.

[161] Coulondre, C., J. H. Miller, P. J. Farabaugh and W. Gilbert., (1978), “Molecular basis of base substitution hotspots in *Escherichia coli*”, Nature 274: 775-780.

[162] Coulondre, C., J. H. Miller, P. J. Farabaugh and W. Gilbert., (1978), “Molecular basis of base substitution hotspots in *Escherichia coli*”, Nature 274: 775-780.

[163] Duncan, B. K. and J. H. Miller, (1980), “Mutagenic deamination of cytosine residues in DNA”, Nature 287: 560-561.

[164] Shenoy, S., K. C. Ehrlich and M. Ehrlich, (1987), “Repair of thymine cntdot guanine and uracil cntdot guanine mismatched base-pairs in bacteriophage M13mp18 DNA heteroduplexes”, J. Mol. Biol. 197: 617-626.

[165] Bird, A. P. (1986), “CpG-rich islands and the function of DNA methylation”, Nature 321: 209-13.

[166] Shen, J. C., W. M. Rideout, 3rd and P. A. Jones., (1994), “The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA”, Nucleic Acids Res; 22(6): 972–976.

[167] Cooper, D. N. and M. Krawczak., (1990), “The mutational spectrum of single base-pair substitutions causing human genetic disease: patterns and predictions”, Human Genet. 85: 55-74.

[168] Jones, P. A., W. M. Rideout, J.-C. Shen, C. H. Spruck and Y. C. Tsai., (1992), Methylation, mutation and cancer. BioEssays 14: 33-36.

[169] Munns, R and James, R.A., (2003) “Screening methods for salinity tolerance: a case study with tetraploid wheat”, Plant and Soil, 253: 201-218.

[170] Munns, R and James, R.A., (2003) “Screening methods for salinity tolerance: a case study with tetraploid wheat”, Plant and Soil, 253: 201-218.

[171] Mares, D.J.; Mrva, K. (2001), “Mapping quantitative trait loci associated with variation in grain dormancy in Australian wheat.”, in Australian Journal of Agricultural Research, 2001, 52(11-12):1257-1265.

[172] Doyle, JJ., and Doyle, J.L., (1987), “A rapid DNA isolation procedure for small quantities of fresh leaf tissue”, Phytochemistry Bulletin, 19:11-15.

[173] Carling, J., (2003), Unpublished in house DNA extraction protocol, Diversity Arrays Technology Pty. Ltd., Canberra, Australia.

[174] New England Biolabs, USA, www.neb.com, accessed November 2006.

[175] New England Biolabs, ‘McrBC Endonuclease’, www.neb.com/nebecomm/products/productM0272.asp, accessed Feburary 2007.

[176] New England Biolabs, ‘McrBC Endonuclease’, www.neb.com/nebecomm/products/productM0272.asp, accessed Feburary 2007.

[177] Diversity Arrays Technology, PO Box 7141, Yarralumla, ACT, Australia.

[178] Invitrogen, USA, (2003), “TOPO TA Cloning, five minute cloning of Taq polymerase-amplified PCR products”, Instruction Manual, July 21, p 15.

[179] Ullmann, A., Jacob, F., Monod, J., (1967), “Characterization by in vitro complementation of a peptide corresponding to an operator-proximal segment of the β-galactosidase structural gene of Escherichia coli.”, J. Mol. Biol. 24, 339.

[180] Botstein D., White R. L., Skolnick M. and Davis R. W., (1980), “Construction of a genetic map in man using restriction fragment length polymorphism”, Am. J. Hum. Genet. 32, 314-331.

[181] Anderson, J.A., G.A. Churchill, J.E. Autrique, S.D. Tanksley, and M.E. Sorrells., (1993), “Optimizing parental selection for genetic linkage maps”, Genome 36:181-186.

[182] Chmielewicz, K.M., and Manly, K.F., ‘XTX’ (Map Manager), Roswell Park Cancer Instutite, www.mapmanager.org.

[183] Voorrips, R.E., ‘MarChart’, Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands, www.biometris.nl.

[184] Smith, H., (1977), “The Molecular Biology of Plant Cells”, Berkeley: University of California Press, Chapter 9, Section 9.8.3.

[185] New England Biolabs, www.neb.com, accessed 2009.

[186] New England Biolabs, www.neb.com, accessed 2009.

[187] Zhu, Jian-Kang, (2001), “Plant salt tolerance”, Trends in Plant Science, Vol 6, No. 2, Feburary, pp. 66-71.

[188] Munns, R., (2002), “Comparative physiology of salt and water stress,” Plant Cell Environ. 25, 239–250.

[189] Munns, R., (2002), “Comparative physiology of salt and water stress,” Plant Cell Environ. 25, 239–250..

[190] Munns, R., (2002), “Comparative physiology of salt and water stress,” Plant Cell Environ. 25, 239–250.

[191] CSIRO Plant Industries, GPO Box 1600, Canberra, ACT, Australia.

[192] Munns, R., (2002), “Comparative physiology of salt and water stress,” Plant Cell Environ. 25, 239–250.

[193] Francois, L.E., Maas, E.V., Donovan, T.J. and Youngs, V.L., (1986) “Effect of salinity on grain yield and quality, vegetative growth and germination of semi-dwarf and durum wheat”, Agron. J. 78, 1053–1058.

[194] Gorham, J., Hardy, C., Wyn Jones, R.G., Joppa, L.R. and Law, C.N., (1987) “Chromosomal location of a K/Na discrimination character in the D genome of wheat”, Theor. Appl. Genet. 74, 584–588.

[195] Munns, R and James, R.A., (2003) “Screening methods for salinity tolerance: a case study with tetraploid wheat”, Plant and Soil, 253: 201-218.

[196] Quarrie, S.A., Lazic-Jancic, V., Kovacevic, D., Steed, A., Pekic, S., (1998) “Bluk segregant analysis with molecular markers and its use for improving drought resistance in maize”, Journal of Experimnetal Botany, Vol. 50, No. 337, pp 1299-1306.

[197] Griffiths, A.J.F., Miller, J.H., Suzuki, D.T., Lewontin, R.C., Gelbart, W.M. (1993). "Chapter 5". *An Introduction to Genetic Analysis* (5th ed.). New York: W.H. Freeman and Company.

[198] Voorrips, R.E., (2002). “MapChart: Software for the graphical presentation of linkage maps and QTLs”, The Journal of Heredity 93 (1): 77-78.

[199] Xu, Mingliang, Li, Xiangqian, Korban, Schuyler, (2000), “AFLP-Detection of DNA Methylation”, Plant Molecular Biology reporter, 18:361-368.

[200] Xu, Mingliang, Li, Xiangqian, Korban, Schuyler, (2000), “AFLP-Detection of DNA Methylation”, Plant Molecular Biology reporter, 18:361-368.

[201] Xu, Mingliang, Li, Xiangqian, Korban, Schuyler, (2000), “AFLP-Detection of DNA Methylation”, Plant Molecular Biology reporter, 18:361-368.

[202] Miflin, B, (2000), “Cropimprovment in the 21st Century”, J Exp Bot 51:1-8.

[203] Peleman, JD, van der Hoot, JR, (2003), “Breeding by Design”, Trends Plant Science, 8:330-334.

[204] Joccoud, D, Penf, K, Feinstein, D and Kilian, A, (2001), “Diversity Arrays: a solid state technology for sequence information independent genotyping”, Nucleic Acid Research, 29:E25.

[205] Kiliaan, A, Hutner, E, Wenzel, P, Jaccoud, D, Carling, J, Caig, V, Evers, M, Heller-Uszynska, K, Cayla, C, Patarapuwadol, S, Xia, L, Yang, S, Thomson, B, (2005), “The fast and cheap: SNP and DArT-based whole genome profiling forcrop improvement”, Proceedings of the International Congress ‘in the wake of the double helix: from green revolution to the gene revolution’, Bologna, Italy, Avenue Media, pp 443-461.

[206] Heller-Uszynska, K, Caig, V, Carling, J, Cayla, C, Evers, M, Uszynski, G, Piperidis, G, Gilmore, R, Aitken, K, Jackson, P, Hutner, E, Kilian, A, (20076), “Diversity Arrays Technology (DArT) for high throughput, whole-genome molecular analysis of sugarcane”, Tropical Crops Biotechnology Conference, Cairns, Australia, 16-19 August.

[207] Li, T-X, Wang, Jin-Ke, Bai, Yun-Fei, Lu, Zu-Hong, (2006), “Diversity suppression-subtractive hybridisation array for profiling genomic DNA polymorphisms”, Journal of Integrative Plant Biology, 48:460-467.

[208] Kane, M. D., Jatkoe, T. A., Stumpf, C. R., Lu, J., Thomas, J. D. and Madore, S. J. (2000). “Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays”, Nucleic Acids Res 28: 4552-4557.

[209] Wren, J. D., Kulkarni, A., Joslin, J., Butow, R. A. and Garner, H. R. (2002), ”Cross-hybridization on PCRspotted microarrays” IEEE Eng Med Bioi Mag 21: 71-75.

[210] Reilly, C, Raghavan, A. and Bohjanen, P. (2006). “Global assessment of cross-hybridization for oligonucleotide arrays,” J Biomol Tech 17: 163-172.

[211] Wenzl, P., Li, H., Carling, J., Zhou, M., Raman, H., Edie, E., Hearnden, P., Maier, c., Xia, L., Caig, V., Ovesna, J., Cakir, M., Poulsen, D., Wang, J., Raman, R., Smith, K. P., Muehlbauer, G. J., Chalmers, K. J., KJeinhofs, A., Huttner, E. and Kilian, A. (2006), “A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits,” BMC Genomics 7:206.

[212] Tanksley, S. D., Ganal, M. W., Martin, G. B. (1995), ”Chromosome landing: A new paradigm for map-based cloning in species with large genomes,” Trends Genet 11: 63-68.

[213] Peters, J. L., Cnudde, F. and Gerats, T. (2003), “Forward genetics and map-based cloning approaches” Trends Plant Sci 8: 484-491.

[214] Gupta, P. K, Roy, J. K and Prasad, M. (2001), “Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants,” Curr Science 80: 524-535.

[215] Van Eijk, M. J. T. (2006), ”Complexity Reduction Of Polymorphic Sequences (CRoPS): A Novel Approach For High Througput Polymorphism Discovery,” Plant & Animal Genome XIV Conference, San Diego, USA.

[216] Hardiman, G. (2004). Microarray platforms=comparisons and contrasts. Pharmacogenomics 5: 487-502. Hatta, R., !to, K, Hosaki, Y., Tanaka, T., Tanaka, A., Yamamoto, M., Akimitsu, K and Tsuge, T. (2002), “A conditionally dispensable chromosome controls host-specific pathogenicity in the fungal plant pathogen *Alternaria alternate*,” Genetics 161: 59-70.

[217] Syvanen, A. C. (2001), “Accessing genetic variation: genotyping single nucleotide polymorphisms,” Nat Rev Genet 2: 930-942.

[218] Luikart, G., England, P. R., Tallmon, D., Jordan, S. and Taberlet, P. (2003), ”The power and promise of population genomics: from genotyping to genome typing,” Nat Rev Genet 4: 981-994.

[219] Syvanen, A. C. (2005), “Toward genome-wide SNP genotyping” at Genet 37: S5-1O.

[220] Kilian, A., Huttner, E., Wenzl, P., Jaccoud, D., Carling, J., Caig, V., Evers, M., Heller-Uszynska, Cayla C, Patarapuwadol S., Xia, L., Yang, S., Thomson, B. (2005), “The fast and the cheap: SNP and DArTbased whole genome profilingfor crop improvement,” Proceedings of the International Congress 'in the Wake of the Double Helix: from the Green Revolution to the Gene Revolution'. Bologna, Italy, Avenue Media, pp. 443-461.

[221] Gerhold, D., Rushmore, T. and Caskey, C. T. (1999), “DNA chips: promising toys have become powerful tools,” Trends Biochem Sci 24: l68-173.

[222] Hazen, S. P. and Kay, S. A. (2003), “Gene arrays are not just for measuring gene expression,” 8: 413-416.

[223] Gershon, D. (2005), “DNA microarrays: more than gene expression,” Nature 437: 1195-1198.

[224] Lipshutz, R. J., Fodol', S. P., Cingeras, T. R. and Lockhart, D. J. (1999), “High density synthetic oligonucleotide arrays,” Nat Genet 21: 20-24.

[225] Pease, A. C., Solas, D., Sullivan, E. J., Cronin, M. T., Holmes, C. P. and Fodor, S. P. (1994), “Light generated oligonucleotide arrays for rapid DNA sequence analysis,” Proc Natl Acad Sci USA 91: 5022- 5026.

[226] Cutler, D. J., Zwick, M. E., Carrasquillo, M. M., Yohn, C. T., Tobin, K. P., Kashuk, C., Mathews, D. J., Shah, N. A., Eichler, E. E., Warrington, J. A. and Chakravarti, A. (2001), “High-throughput variation detection and genotyping using microarrays,” Genome Res **11:** 1913-1925.

[227] Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X. c., Stern, D., Winkler, J., Lockhart, D. J., Morris, M. S. and Fodor, S. P. (1996), “Accessing genetic information with high-density DNA arrays,” Science 274: 610-614.

[228] Hacia, J. G., Sun, B., Hunt, N., Edgemon, K., Mosbrook, D., Robbins, c., Fodor, S. P., Tagle, D. A. and Collins, F. S. (1998), "Strategies for mutational analysis of the large multiexon ATM gene using highdensity oligonucleotide arrays," Genome Res 8: 1245-1258.

[229] Wang, D. G., Fan, J. B., Siao, C. J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J., Kruglyak, L., Stein, L., Hsie, L., Topaloglou, T., Hubbell, E., Robinson, E., Mittmann, M., Morris, M. S., Shen, N., Kilburn, D., Rioux, J., Nusbaum, C., Rozen, S., Hudson, T. J., Lipshutz, R., Chee, M. and Lander, E. S. (1998), "Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome," Science 280: 1077-1082.

[230] Dong, S., Wang, E., Hsie, L., Cao, Y., Chen, X., and Gingeras, TR., (2001), "Flexible use of high-density oligonucleotide arrays for single-nucleotide polymorphism discovery and validation," Genome Res 11:1418-1424.

[231] Kennedy, G. C., Matsuzaki, H., Dong, S., Liu, W. M., Huang, J., Liu, G., Su, X., Cao, M., Chen, W., Zhang, J., Liu, W., Yang, G., Di, X., Ryder, T., He, Z., Surti, U., Phillips, M. S., Boyce-Jacino, M. T., Fodor, S. P. and Jones, K W. (2003), "Large-scale genotyping of complex DNA,". Nat Biotechnol 21: 1233-1237.

[232] Winzeler, E. A., Richards, D. R., Conway, A. R., Goldstein, A. L., Kalman, S., McCullough, M. J., McCusker, J. H., Stevens, D. A., Wodicka, L., Lockhart, D. J. and Davis, R. W. (1998), "Direct allelic variation scanning of the yeast genome," Science 281: 1194-1197.

[233] Winzeler, E. A., Richards, D. R., Conway, A. R., Goldstein, A. L., Kalman, S., McCullough, M. J., McCusker, J. H., Stevens, D. A., Wodicka, L., Lockhart, D. J. and Davis, R. W. (1998), "Direct allelic variation scanning of the yeast genome," Science 281: 1194-1197.

[234] Winzeler, E. A., Castillo-Davis, C. I., Oshiro, G., Liang, D., Richards, D. R., Zhou, Y. and Hartl, D. L. (2003), "Genetic diversity in yeast assessed with whole-genome oligonucleotide arrays," Genetics 163:79-89.

[235] Stickney, H. L., Schmutz, J., Woods, I. G., Holtzer, C. c., Dickson, M. c., Kelly, P. D., Myers, R. M. and Talbot, W. S. (2002), "Rapid mapping of zebrafish mutations with SNPs and oligonucleotide microarrays," Genome Res 12: 1929-1934.

[236] Borevitz, J. 0., Liang, D., Plouffe, D., Chang, H. S., Zhu, T., Weigel, D., Berry, C. C., Winzeler, E. and Chory, J. (2003), "Large-scale identification of single-feature polyrnorphisms in complex genomes," Genome Res 13: 513-523.

[237] West, M. A., van Leeuwen, H., Kozik, A., Kliebenstein, D. J., Doerge, R. W., St Clair, D. A. and Michelmore, R. W. (2006), "High-density haplotyping with microarray-based expression and single feature polymorphism markers in *Arabidapsis*,". Genome Res. 16: 787-795.

[238] Kim, S., Zhao, K, Jiang, R., Molitor, J., Borevitz, J. 0., Nordborg, M. and Marjoram, P. (2006), "Association mapping with single-feature polymorphisms," Genetics 173: 1125-1133.

[239] Kirov, G., Nikolov, I., Georgieva, L., Moskvina, V., Owen, M. J. and O'Donovan M, C (2006), "Pooled DNA genotyping on Affymetrix SNP genotyping arrays," BMC Genomics 7: 27.

[240] Brem, R 8., Yvert, G., Clinton, Rand Kruglyak, L. (2002), "Genetic dissection of transcriptional regulation in budding yeast," Science 296: 752-755.

[241] Ronald, J., Akey, J. M., Whittle, J., Smith, E. N., Yvert, G. and Kruglyak, L. (2005), "Simultaneous genotyping, gene-expression measurement, and detection of allele-specific expression with oligonucleotide arrays," Genome Res 15: 284-291.

[242] West, M. A., van Leeuwen, H., Kozik, A., Kliebenstein, D. J., Doerge, R. W., St Clair, D. A. and Michelmore, R. W. (2006), "High-density haplotyping with microarray-based expression and single feature polymorphism markers in *Arabidapsis*" Genome Res. 16: 787-795.

[243] Singh-Gasson, S., Green, R. D., Yue, Y., Nelson, C, Blattner, F., Sussman, M. R. and Cerrina, F. (1999), "Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array," Nat Biotechnol 17: 974-978.

[244] Stengele, K. P., Buhler, J., Buhler, S., Kvassiouk, E., Green, R., Prykota, T. and Pfleiderer, W. (2005), "Recent highlights on photolithic oligonucleotide array in situ synthesis," Nucleosides Nucleotides Nucleic Acids 24: 891-896.

[245] Gunderson, K L., Steemers, F. J., Lee, G., Mendoza, L. G. and Chee, M. S. (2005), "A genome-wide scalable SNP genotyping assay using microarray technology," Nat Genet 37: 549-554.

[246] Shen, R., Fan, J. 8., Campbell, D., Chang, W., Chen, J., Doucet, D., Yeakley, J., Bibikova, M., Wickham Garcia, E., McBride, C, Steemers, F., Garcia, F., Kermani, B. G., Gunderson, K and Oliphant, A. (2005), "High-throughput SNP genotyping on universal bead arrays," Mutat Res 573: 70-82.

[247] Fan, J. B., Hu, S. X., Craumer, W. C. and Barker, D. L. (2005), " Bead Array-based solutions for enabling the promise of pharmacogenomics," Biotechniques 39: 583-588.

[248] Hall, N., (2007), "Advanced sequencing technologies and their wider impact in microbiology". *J. Exp. Biol.* 210 (Pt 9): 1518–25.

[249] Rostoks, N., Borevitz, J. 0., Hedley, P. E., Russell, J., Mudie, S., Morris, J., Cardle, L., Marshall, D. F. and Waugh, R. (2005), "Single-feature polymorphism discovery in the barley Transcriptome," Genome Biol 6: R54.

[250] Van Eijk, M., Bradbury, P., Gore, M., Hogers, R., *Kirst,* M., Verstege, E., Van Oeveren, J., Buckler, E. (2006), "Optimized Target Preparation Method For Single-Feature Polymorphism Detection In Maize," Plant & Animal Genomes XIV Conference, San Diego, CA.

[251] Huttner, E., Wenzl, P., Akbari, M., Caig, Y., Carling, J., CayJa, c., Evers, M., Jaccoud, D., Peng, K, PatarapuwadoJ, S., Uszynski, G., Xia, L., Yang, S., and Kilian, A. (2004), "Diversity Arrays Technology: A novel tool for harnessing the genetic potential of orphan crops," Proceedings of the 2004 Conference of The

World Biological Forum Discovery to Delivery: BioYisionAlexandria 2004, CABI Publishing.

[252] Kilian, A., Huttner, E., Wenzl, P., Jaccoud, D., Carling, J., Caig, V., Evers, M., Heller-Uszynska, Cayla C, Patarapuwadol S., Xia, L., Yang, S., Thomson, B. (2005), The fast and the cheap: SNP and DArT based whole genome profilingfor crop improvement," Proceedings of the International Congress 'in the Wake of the Double Helix: from the Green Revolution to the Gene Revolution'. Bologna, Italy, Avenue Media, pp. 443-461.

[253] Akbari, M., Wenzl, P., Caig, v., Carling, J., Xia, L., Yang, S., Uszynski, G., Mohler, V., Lehmensiek, A., Kuchel, H., Hayden, M. J., Howes, N., Sharp, P., Vaughan, P., Rathmell, B., Huttner, E. and Kilian, A. (2006), "Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome," Theor Appl Genet 113: 1409-1420.

[254] Heller-Uszynska, K., Caig, Y., Carling, J., Elvers, M., Uszynski, G., Piperidis, G., Gilmour, R., Aitken, K., Jackson, P., Huttner, E., Kilian, A. (2006), "Diversity Arrays Technology (DArT) for high throughput, whole-genome molecular analysis in sugarcane," Tropical Crops Biotechnology Conference, Cairns, Australia, 16-19August.

[255] Wongtiem, P., Perera, C., Risterucci, A. M., Xia, L., Patarapuwadol, S., Carling, J., Caig, V., Evers, M., Uszynski, G., Wenzl, P., Huttner, E., Fregene, M., de Vicente, C., Glaszmann, J. C. and Kilian, A. (2006), "Diversity Arrays Technology (DArT): an Efficient Platform for Whole-Genome Profiling of Crops," Presentation delivered at the Generation Challenge Program Workshop of the Plant and Animal Genome XIV Conference, San Diego,14-18 January.

[256] Esquinas-Alcazar, J. (2005), "Science and society: protecting crop genetic diversity for food security: political, ethical and technical challenges," Nat Rev Genet 6: 946-953.

[257] Louwaars, N. (2006), "Controls over plant genetic resources - a double-edged sword," Nat Rev Genet 7: 241.

[258] Dennis, C. (2004), "Biologists launch 'open-source movement'," Nature 431: 494.

[259] Cohn, D. (2005), "Open-Source Biology Evolves," Wired magazine, www.wired.com, 17/01/2005.

[260] Schenker, J. (2006), "Open-Source Biotech," Red Herring magazine, www.redherring.com, accessed 24/10/2008.

[261] CSIRO (2011), "CSIRO develops highest yielding salt tolerant wheat",

www.csiro.au/news/CSIRO-develops-highest-yielding-salt-tolerant-wheat.