Characterization of phenotypic and genotypic selection for simple and complex traits of barley (*Hordeum vulgare* L.)

This thesis is presented for the degree of

Doctor of Philosophy of Murdoch University

July 2009

REETINDER GILL

BSc. Agric. (Hons.)
MSc. Vegetable Crops

The School of Biological Sciences,

Division of Science and Engineering,

Murdoch University

Declaration

I declare that the research presented in this thesis is original and was undertaken and written by myself, except where specifically indicated in the text.

The thesis has been completed during the course of enrolment in a PhD degree at Murdoch University and has not been used previously for a degree or diploma at any other institution.

Reetinder Gill

Abstract

The challenge to modern day plant breeding is to constantly strive to improve the efficiency and efficacy of utilising the available resources. This requires the integration of conventional breeding technologies together with molecular genetic markers to significantly improve the breeding programs.

Male sterile facilitated recurrent selection (MSFRS) can improve a breeding program by reducing the variety release time to almost half, maintaining continued out-crossing and recombination, thus broadening the germplasm base. A molecular genetic marker system for the male sterile-orange lemma-shrunken endosperm (*msg6-rob1-sex1)* linkage block was developed during this study, to aid in the differentiation of fully fertile homozygous from fully fertile heterozygous plump individuals in F_2 populations. These individuals may be desired for the production of doubled haploids or for use as pollen donor parents in further cycles of MSFRS. A DF578/6*Gairdner BC_5 population was chosen for the study and SSR markers were identified in the region of the *msg6-rob1-sex1* linkage group on the short arm of chromosome 6H. A linkage map was constructed and it was found that microsatellite markers HVM65, HVM74 and Bmgtttttt1 are tightly linked to *msg6-rob1-sex1* linkage block.

Within the MSFRS process, it was found that genotypic selection with SSR markers is an advantage when the phenotype of interest is controlled by a single major gene and the marker is either “perfect” or closely linked. Genotypic selection was found to be very effective in selecting for the aluminium tolerant genotypes with the SSR marker HVM68,

where aluminium tolerance is known to be controlled by *Alt* or *Alp* gene on chromosome 4H.

For the complex trait such as barley scald resistance, genotypic selection was found to be of limited value. Simulation studies demonstrated that multiple genetic factors need to be taken into account while selecting for a complex phenotype. As expected, phenotypic selection method was found to be efficient in selecting for scald resistance as it selected minor genes along with two known major genes for scald resistance on chromosome 4H ($Rhs_{Vlamingh}$) and 6H ($Rhs_{WABAR2147}$).

Simulation studies based on the experimental results provided a guide for the frequency and timing of the use of molecular markers in the breeding program. It was found that markers that are loosely linked to the gene of interest should only be used once early in the breeding program. As in the case of scald resistance, where the markers-gene distances are 5cM or 30 cM, the genes can only be partially fixed. Further cycles of genotypic selection for scald resistance will lead to the selection of the susceptible genotypes instead of resistant genotypes as the phenotype comes in repulsion rather than the coupling phase with the scald resistant genes.

GGT (an acronym of Graphical GenoTypes) software package was extensively used to study genotypic changes in response to selection and to select for the genotypes carrying resistant genes of interest. Based on the allele frequencies and the marker-scald associations carried out using GGT, the SSR marker Bmac213 (1H) was found to be associated with powdery mildew ($Mla_{WABAR2147}$) resistance and SSR markers GBM1221

(4H) and Bmac316 (6H) were found to be associated with scald ($Rhs_{Vlamingh}$ and $Rhs_{WABAR2147}$) resistance.

Results presented in this thesis have enabled the identification of mechanisms behind the success of phenotypic selection and its use while selecting for quantitative traits as it incorporates the minor gene effects while selecting for the major genes. Genotypic selection method was found to be efficient in selecting for the desired genotypes but may not give the desired result in terms of phenotype when a complex trait is involved. Genotypic selection will be at par with the phenotypic selection for the complex quantitative traits if the associated markers with the minor genes are included in the selection. Both phenotypic and genotypic selection methods together can be used effectively in the breeding program to increase the rate of genetic gain.

Acknowledgements

During the course of this PhD thesis I have come to believe that interdependence is certainly more valuable than independence as a journey is easier when you travel together. First and foremost I would like to express my special thanks to my husband and best friend Mandeep for helping me to concentrate on completing my thesis and supporting me mentally during the course of this work. I would not be sitting in front of my monitor typing these acknowledgement lines without you. I am very grateful for your love, patience, help and encouragement during the PhD period.

This is a great opportunity to thank all the people who accompanied and supported me during my PhD. I extend my most sincere thanks to my supervisors, Dr. Reg Lance, Prof. Rudi Appels and Dr. Chengdao Li for their advice and support over the period of my candidature. Thank you Rudi for giving me the opportunity to take on this project and for being supportive during all the ups and downs of this journey. I am really grateful to you for all those hours that I am sure you have spent outside of work hours to read my work and respond in the punctual time, which you always do. I wish to thank

Reg for his constant encouragement, great understanding and for the lessons he gave me on work-life balance - you have always put in a large effort to help make the fieldwork successful. Internationally, I would like to acknowledge A/Professor Duane Falk for his help and support whenever I needed it and for taking time out during holidays to meet me to discuss my project and general breeding and genetics, when I visited University of Guelph, Canada.

I thank everyone from the Barley group at the Department of Agriculture and Food, Western Australia and Research station, UWA, for assisting me with phenotyping experiments. I also appreciate the contributions of Dr. Katia Stefanova with the statistical analysis for the experiments. My experience at SABC would not have been such a good one without the presence of all the 'girls' working there. Thanks Dora, Vera, Julie, Esther, Sharon, Fiona, JJ and Steve for putting up with my questions and for their advice. I have learned a lot from you all.

Funding for this research was provided by Murdoch University and the Molecular Plant Breeding CRC. Without this financial support my research would not have been possible. Financial support for conference attendance has

also been provided by Murdoch University and MPBCRC with additional funds from DAFWA. Additionally, I acknowledge the facilities provided for this research by The Western Australian State Agricultural Biotechnology Centre and the Department of Agriculture and Food, Western Australia.

Special thanks to my friend Yumi, whose friendship I will cherish forever. Over the years I have been grateful for your support and friendship. You made my time at Murdoch good fun, as well as stimulating. I really enjoyed all our wide-ranging discussions and miss them. Many thanks to my friend Dean Deipeveen for always supporting and believing in me which has led to the next step in my scientific career. Thanks are due to my friends Jyoti, Meenu, Neetu, Bobby and Navi for all their support during my PhD candidature. Writing these acknowledgements, I am realizing that never again will I spend the entire weekends working in the lab or thinking about the 'perfect sentences' for the discussion sections in my thesis.

Finally, I would like to thank my parents and my brothers, Puneet and Sumeet, who have always been there for me as an unwavering support.

Dedication

I wish to dedicate this thesis to my daughter, Raavi - born in March 2009 - who is too young to understand this but I hope this will inspire her to achieve her goals under any circumstances. One of the best experiences that I and Mandeep lived through in this period was waiting for our first child to be born and our little princess has provided an additional and joyful dimension to our life.

Table of Contents

Page #

List of Figures

Chapter 1:

Chapter 2:

Chapter 3:

of the parents, using a SSR marker HVM68 on chromosome 4H. The white coloured numbers are representing the homozygotes, red coloured representing the heterozygotes for aluminium tolerance and the blue colour represents the aluminium intolerant individuals in the population screened. Where BK – Birka (DF168), Sk – Skiff, 96 – WABAR2096, 94 – WABAR2094, Vlm – Vlamingh, 229 – WB229 and Bdn – Baudin .

3.7 Showing the number of individuals with different genotypes, selected using the SSR marker HVM68 (Figure 3.5). Encircled homozygote aluminium tolerant individuals (Al-tol (AA)) were planted in the crossing block after selection to develop F_1 seed. Where Al-tol (Aa) – heterozygotes for aluminium tolerance and Al-sensitive – individuals which were found to be sensitive to aluminium.

3.8 Results based on the predicted means of the phenotyping at F_2 are shown. The Comparisons between phenotypic (Treatment 1) and genotypic (Treatment 2) selections for Aluminium tolerance carried out at F_1 stage have been made based on the data collected on the different traits at segregating F_2 stage. Comparisons were made based on the data on (A) Yield (B) 1000-kernel weight (C) Plumpness (%Plump, >2.5mm) and (D) number of tillers. Values for the parents of the population are indicated in arrows.

Chapter 4:

4.1 Showing the pedigree of Cross1. Each colour represents the parent and the proportion of each colour corresponds to the proportion of that specific parent in the population.

Bdn – Baudin, Hml – Hamelin, Vlm – Vlamingh and 47 – WABAR2147.

4.9 Results based on the predicted means of the phenotyping at $F_{2,}$ carried out in year 2007 are shown. The comparison between phenotypic selection (Treatment 1) and genotypic selection (Treatment 2) methods for scald carried out at F_1 have been made for both the populations. (A) Cross 1 population (see Figure 4.1) and (B) Cross 2 population (see Figure 4.2).

4.10 Box plot for scald phenotype for each simulation strategy, 1 – original set of 235 DH lines, 2 – 53 DH lines selected phenotypically and 3 – 53 DH lines selected genotypically. Along x-axis are the three simulations and along y-axis are the scald scores on the scale of 0-9 where 9 represent the most resistant DH lines for scald.

4.11 Box plots for each DH genotype based on the three simulation strategies, 1 – original set of 235 DH lines, 2 – 53 DH lines selected phenotypically and 3 – 53 DH lines selected genotypically. Along x-axis are the three simulations for each genotype and along y-axis are the scald scores on the scale of 0-9 where 9 represent the most resistant DH lines for scald. Genotype 11- both the alleles from Vlamingh, genotype 12 - one allele from Vlamingh and one from WABAR2147 and genotype 22 - both the alleles from WABAR2147. Black spot represents the scald means for the populations.

4.12 Phenotypic and genotypic changes in response to phenotypic selection (Treatment 1) and genotypic selection (Treatment 2) carried out in year 2006 and 2007 for Cross 1. The data on the genotypic selection for the Year 2007 had been taken from Chapter 5 which was based on the Graphical Genotyping of solely the homozygote scald and powdery mildew resistant genotypes. Where PS –

phenotypic selection, GS (1H) - genotypic selection for powdery mildew on chromosome 1H, GS (4H) - genotypic selection for scald on chromosome 4H and GS (6H) - genotypic selection for scald on chromosome 6H.

Chapter 5:

5.1 The histogram showing the selected genotypes after the phenotypic selections where the leaf area damage (%) ranged from 0-3.5. The genotypes found to be resistant (circled) on the disease scale with 0 leaf area damage (%) were used for the graphical genotyping of the scald and powdery mildew resistances.

5.2 Schematic presentation of gene transfer from the scald and powdery mildew resistant parents to a susceptible population to get the moderately resistant population with all the resistant genes after crossing. In the figure, resistant genes are represented by: A – powdery mildew resistance gene; $Mla_{WABAR2147}$, on chromosome 1H, B and C – scald resistance genes; Rhs_{Vlamingh} and $Rhs_{\text{WABAR2147}}$, on chromosomes 4H and 6H respectively.

5.3 Representation of the GeneMapper results showing the differences in the allele sizes for SSR markers. (A) – for the marker hv0020, two genotypes; 2 and 3 were found to be homozygotes for the alleles a and b and the genotype 1 was found to be heterozygote, as exhibited both the alleles; a and b, (B) – this graph shows the alleles exhibited by three out of five parents for the marker hv0268, all the three parents contained allele a but were homozygotes for allele b, c and d respectively and (C) – for marker hv0614, the first two parents were found to contain allele b

and the third parent had allele a.

5.4 Allele frequencies at the marker loci of 203 F_2 plants under study to monitor the allelic changes in response to genotypic selection (Treatment 2). The allele codes are based on Table 5.3. Along the x-axis are the Graphically Genotyped individuals and along y-axis are the SSR markers used for Graphical Genotyping. The allele frequencies of the major alleles at a given locus are given on the right hand side of the figure.

5.5 Allele frequencies at the marker loci of 213 F_2 plants under study to monitor the allelic changes in response to genotypic selection (Treatment 2). The allele codes are based on Table 5.3. Along the x-axis are the Graphically Genotyped individuals and along y-axis are the SSR markers used for Graphical Genotyping. The allele frequencies of the major alleles at a given locus are given on the right hand side of the figure.

5.6 A graphical display of the overview association plots along the chromosome bars developed using GGT. Only the markers with the R^2 values above the threshold are shown. The red bars circled blue show strong association between the marker and the trait and the one circled yellow show moderate associations. (A) Marker-trait association for phenotypically selected (Treatment 1) population. Yellow circled bar is for marker Bmac213 (1H) and blue circled bars are markers GBM1501 (4H) and Bmac316 (6H), similarly, (B) Marker-trait association for genotypically (Treatment 2) selected population. Yellow circled bars are for marker Bmac213 (1H) and Bmac316 (6H) and blue circled bars are markers

Bmac310 (4H) and GMS89 (4H). Only chromosomes 1H, 4H and 6H have been considered.

5.7 Forked-line method showing the formation of gametes for markers (A) Bmac213 (*Mla*$_{WABAR2147}$), (B) GBM1221 (*Rhs*$_{Vlamingh}$) and (C) Bmac316 (*Rhs*$_{WABAR2147}$).

5.8 Punnet-square presentation of all the possible 64 genotypes which can be developed from the 8 gametes (Figure 5.7). Each colour represents a genotype present in (A) phenotypically selected (Treatment 1) and (B) genotypically selected (Treatment 2) populations. The numbers in the boxes along the genotypes represent the number the genotypes in the population.

5.9 The selection strategies simulated to select for scald and powdery mildew, (A) phenotypically, (B) and (C) genotypically, when the markers are 30cM (B) and 5cM (C) away from the gene of interest. Where red lines are the scald scores and green lines are the sum of scald marker scores.

5.10 The frequencies of the genes and the associated markers at (A) 30cM and (B) 5cM distances. Four bars for each gene/marker represent 0, 1, 2, 3 cycles of selection. Left hand set of four bars for each gene/marker represents the frequencies after phenotypic selection and the right hand set of bars represents the frequencies after genotypic selection. Where: scald-VLM – *Rhs*$_{Vlamingh}$ and scald-2147 – *Rhs*$_{WABAR2147}$. It was assumed that no selections were carried out for *Mla*$_{WABAR2147}$ gene on chromosome 1H.

5.11 Histogram of genotype frequencies for three simulations, 1 – original set of 235 DH lines, 2 – 53 DH lines selected phenotypically (Treatment 1) and 3 – 53 DH

lines selected genotypically (Treatment 2). Where genotype 11 - both the alleles from Vlamingh, genotypes 12 and 21 - one allele from Vlamingh and one from WABAR2147 and genotype 22 - both the alleles from WABAR2147.

Chapter 6:

6.1 The distribution of the population without selection (starting population), after phenotypic selection (solid black line), after genotypic selection (dotted black line) for the major gene and after the genotypic selection of the major gene plus 3-4 minor genes controlling scald resistance. Where R represents the desirable genotypes and S represents the undesirable genotypes in the population. The key Chapters involved in establishing the features of the distributions shown are indicated.

List of Tables

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

phenotypic selection (Treatment1), T2 – genotypic selection (Treatment2), SC1 – scald score 1, SC2 – scald score 2 (taken 30 days after SC1) and SC3 – scald score 3 (taken 30 days after SC2).

4.7 Correlation coefficients between the three scald scores for both the Crosses and Treatments. T1 – phenotypic selection (Treatment1), T2 – genotypic selection (Treatment2), SC1 – scald score 1, SC2 – scald score 2 (taken 30 days after SC1) and SC3 – scald score 3 (taken 30 days after SC2).

Chapter 5:

5.1 Expected genotype frequencies based on the pedigree of the population used for graphical genotyping for scald and powdery mildew resistances.

5.2 Details of the SSR markers used for graphical genotyping.

5.3 Allelic codes assigned to the Parents and their different combinations for graphical genotyping display. Different combinations are based on the five parents; A-WABAR2096, B-WABAR2147, C-Birka, D-Skiff and E-Vlamingh.

5.4 Categorizing the alleles into parental alleles according to Identity by Descent. Allele codes are based on Table 5.3.

5.5 Cumulative allele frequencies of all the alleles present in the graphically genotyped population for scald and powdery mildew resistances. For allele codes please see Table 5.3.

5.6 The genotypes with their genotypic frequencies for both phenotypically selected