Molecular Breeding Strategies for Improvement of Complex Traits in Soybean

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Christopher Joseph Smallwood

December 2015

Copyright © 2015 by Christopher Joseph Smallwood
All rights reserved.

ACKNOWLEDGEMENTS

I would like to extend my gratitude to all those who helped me throughout the course of this Dissertation. Foremost, I would like to thank Dr. Vincent Pantalone, who has served as my major advisor. Thank you for the excellent educational opportunities; as well as the guidance and support you have provided me with in conducting this research.

I would also like to thank my committee members, Dr. Hem Bhandari, Dr. Jason Gillman, Dr. Arnold Saxton, and Dr. Phillip Wadl. Your direct contributions to ideas and research have proven invaluable. With your help, my abilities and understanding in breeding, statistics, and molecular genetics have improved greatly.

Much assistance was provided by many former and current members of the Soybean Breeding Team and the UTIA farm crews in the completion of this research. To all of you who helped, I thank you: Dr. Ben Fallen, Deborah Landau-Ellis, Elizabeth Meyer, Rachel Fulton, Jeff Boehm, Lauren Richardson, Alison Willette, Laura Betz, Nick Betz, Victoria Benelli, Mia Cunicelli, Greg Allen, Nicole Tacey, Hailee Korotkin, Jimmy McClure, Jason Williams, Chris Bridges, Brad Fisher, BJ DeLozier, Derrick Hopkins, Charles Summey, Vasilj Bobrek, Brad Reagan, Lee Ellis, and too many others to name.

Thank you to Dr. Perry Cregan, Dr. David Hyten and the USDA-ARS Beltsville Agricultural Research Center, Soybean Genomics and Improvement Laboratory for genotyping the material used in this population. Thank you to Dr. Jim Orf and Art Killam of the University of Minnesota for performing and assisting with Near Infrared Reflectance Spectroscopy analyses. Thank you to Dr. Tiffany Langewisch and Dr. Kristin Bilyeu of the USDA-ARS, Columbia, MO, for identifying linked markers to the E1, E3, and Dt1 loci. Thank you to Dr. Jason Gillman and the University of Missouri DNA Core for sequencing the genomes of

the parent lines used in this research. Thank you to Dr. Arnold Saxton of the University of Tennessee for providing statistical expertise, as well as spending countless hours answering questions and running data analyses. Thank you to Dr. Hem Bhandari of the University of Tennessee for giving me the opportunity to give lectures in your plant breeding class; this was a very helpful experience in helping grow my understanding of plant breeding. Thank you to Dr. Phillip Wadl for providing overall guidance and support, along with insight into horticultural breeding.

I would like to extend a special thanks to Dr. Fred Allen for introducing me to plant breeding, as well as for support and guidance in these efforts.

I am very grateful to the United Soybean Board, the Tennessee Soybean Promotion Board, and the University of Tennessee Institute of Agriculture for financial support. This research would not have been possible without your generous contributions.

Finally, I would like to thank my family for their support and encouragement. Especially, I would like to thank my wife Jennifer, who has supported my efforts all the way and inspired me to a higher level of performance.

ABSTRACT

Soybean [Glycine max (L.) Merrill] is the leading oilseed crop grown in the world. Yield, fatty acids, protein, and oil are commercially important soybean traits; thus evaluation of breeding strategies for improvement of these traits is merited. To accomplish this, a comparison of molecular and phenotypic breeding strategies from progeny row selections was performed. From this it was determined that molecular strategies consistently outperformed phenotypic selections (PS) in the progeny row stage for soybean yield, fatty acids, protein, and oil. For yield, Epistacy was the preferred selection method. For fatty acids, protein, and oil, the genomic selection (GS) strategies were preferred. Additionally, a second comparison of molecular and phenotypic strategies was performed with selections from replicated field trials. These comparisons displayed mixed results except for yield, for which PS was the dominant method. With selection from replicated field trials, PS and GS methods were comparable for fatty acids, protein, and oil; indicating that either of these methods could be useful for making improvements. In addition to selection method evaluation, increased knowledge of genomic regions governing soybean yield, fatty acids, protein, and oil would be helpful. Thus, quantitative trait loci (QTL) detection was performed for these traits, with a total of 29 QTLs identified. Of these QTLs, three were candidates for confirmed status and four were candidates for positional confirmations. Additionally, possible candidate genes for soybean yield, fatty acids, protein and oil associated with QTLs in this study were identified; as were pleiotropic effects between protein and oil and between the fatty acids. The results from this research should be beneficial for those seeking to make soybean improvements. Researchers making selections from both progeny rows and replicated field trials can draw from these results when choosing which selection strategy to use. The gained knowledge of influential genomic regions for these traits can have application in improvement efforts. Future research seeking to implement high performing molecular breeding strategies and to identify causative genes for

these and other QTLs impacting targeted traits will be important for the soybean breeding community.

TABLE OF CONTENTS

Chapter 1 Introduction and Literature Review 1

References 8

Chapter 2 Molecular Breeding Strategies Outperform Phenotypic Selection For Soybean Quantitative Traits in the Progeny Row Stage 13

Abstract 14
Introduction 14
Materials and Methods 17
Results 23
Discussion 27
Conclusions 30
References 32
Appendix B-Chapter 2 Tables and Figures 39

Chapter 3 Mixed Results Between Phenotypic and Molecular Breeding Methods for Soybean Quantitative Traits Predicted from Replicated Field Trials 67

Abstract 68
Introduction 69
Materials and Methods 71
Results 77
Discussion 81
Conclusions 84
References 86
Appendix C-Chapter 3 Tables and Figures 93

Chapter 4 Identifying and Exploring Significant Genomic Regions For Soybean Yield, Fatty Acids, Protein, and Oil 121

Abstract 122
Introduction 122
Materials and Methods 124
Results and Discussion 129
Conclusions 139
References 141
Appendix D-Chapter 4 Tables and Figures 147

Chapter 5 Conclusions 194

References 197

Vita 199

LIST OF TABLES

Table 2.1 Simple statistics for soybean population E×W-50K consisting of 860 F5 derived RILs planted in single rep plots in 2010 in Knoxville, TN. This dataset was used to make performance predictions for traits of interest in a subset of the population (276 RILs) grown in replicated field trials in 2013 at three locations (Knoxville, TN; Springfield, TN; and Milan, TN).....40

Table 2.2 Simple statistics for soybean population E×W-50K subset consisting of 276 F5 derived RILs planted in replicated field trials at three locations in 2013 (Knoxville, TN; Springfield, TN; and Milan, TN). Information from this dataset was compared with performance predictions for traits of interest in the full population (860 RILs) grown in 2010 in single rep plots planted at Knoxville, TN. 41

Table 2.3 Comparison of cross-validations for G-BLUP and BayesB methods of GS for soybean population E×W-50K consisting of 860 F5 derived RILs grown in 2010 at Knoxville, TN. Cross-validations were replicated 50 times for each trait. In each rep, a randomly chosen 1/5 of the population had phenotypic data removed (test set), while phenotypic and genotypic information were retained for the remaining 4/5 of the population (training set). The values displayed for G-BLUP and BayesB are the mean Pearson correlation coefficients for the predicted and observed values in the test set. For each trait there was no statistical difference (P > 0.05) between G-BLUP and BayesB methods.....42

Table 2.4 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for soybean yield in population E×W-50K subset consisting of 276 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), and E (2010 Epistacy). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.....43

Table 2.5 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for palmitic acid in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. 44

Table 2.6 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for stearic acid in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. 45

Table 2.7 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for oleic acid in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. 46

Table 2.8 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for linoleic acid in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. 47

Table 2.9 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for linolenic acid in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. 48

Table 2.10 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for protein in soybean population E×W-50K subset consisting of 271 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. 49

Table 2.11 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for oil in soybean population E×W-50K subset consisting of 271 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. 50

Table 2.12 Yield contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 276 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, and G-BLUP) selection methods. 51

Table 2.13 Palmitic acid contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K

subset consisting of 275 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. 52

Table 2.14 Stearic acid contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. 53

Table 2.15 Oleic acid contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. 54

Table 2.16 Linoleic acid contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. 55

Table 2.17 Linolenic acid contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. 56

Table 2.18 Protein contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 271 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. 57

Table 2.19 Oil contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 271 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. 58

Table 3.1 Simple statistics for soybean population E×W-50K subset consisting of 276 F5 derived RILs from the combined 2010 and 2013 field seasons. This

dataset was used to make performance predictions for traits of interest in a subset of the population (203 RILs) grown in replicated field trials in 2014 at three locations (Knoxville, TN; Springfield, TN; and Milan, TN). Yield predictions were made using data from only the Knoxville, TN location (displayed in bold text). ... 94

Table 3.2 Simple statistics for soybean population E×W-50K subset consisting of 203 F5 derived RILs, planted in replicated field trials at three locations in 2014 (Knoxville, TN; Springfield, TN; and Milan, TN). Information from this dataset was compared with performance predictions for traits of interest in the larger subset (276 RILs) from the combined 2010 and 2013 field seasons. Yield comparisons were made using data from only the Knoxville, TN location (displayed in bold text). ... 95

Table 3.3 Comparison of cross-validations for G-BLUP and BayesB methods of GS for soybean population E×W-50K subset consisting of 276 F5 derived RILs; the combined 2010 and 2013 datasets were used for analysis. All traits were analyzed with the multi-location dataset except yield, which was only analyzed with data from the Knoxville, TN location. Cross-validations were replicated 50 times for each trait. In each rep, a randomly chosen 1/5 of the population had phenotypic data removed (test set), while phenotypic and genotypic information were retained for the remaining 4/5 of the population (training set). The values displayed for G-BLUP and BayesB are the mean Pearson correlation coefficients for the predicted and observed values in the test set. Listed P values indicate if there is a statistically significant difference between GS methods. Methods with higher cross-validation correlation values are displayed with bold text. ... 96

Table 3.4 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for soybean yield from only the Knoxville, TN location in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. ... 97

Table 3.5 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for palmitic acid in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. ... 98

Table 3.6 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for stearic acid in soybean

population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. 99

Table 3.7 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for oleic acid in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. 100

Table 3.8 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for linoleic acid in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. 101

Table 3.9 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for linolenic acid in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. 102

Table 3.10 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for protein in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. 103

Table 3.11 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for oil in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings. ...104

Table 3.12 Yield contrasts from only the Knoxville, TN location of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. 105

Table 3.13 Palmitic acid contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.....106

Table 3.14 Stearic acid contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.....107

Table 3.15 Oleic acid contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.....108

Table 3.16 Linoleic acid contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.....109

Table 3.17 Linolenic acid contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived

RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.....110

Table 3.18 Protein contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.....111

Table 3.19 Oil contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.....112

Table 4.1 Simple statistics taken from combined analysis using data from 2010, 2013, and 2014 field seasons for soybean population E×W-50K subset consisting of 203 F5 derived RILs.....148

Table 4.2 Quantitative trait loci (QTL) for yield, fatty acids, protein, and oil detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was estimated from the combined analysis using data from 2010, 2013, and 2014 field seasons. Significant QTL at the 1% threshold (based on 10,000 permutations) are displayed in bold text...149

Table 4.3 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 4 based on Glyma 2.0 position.....150

Table 4.4 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 6 based on Glyma 2.0 position.....151

Table 4.5 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 7 based on Glyma 2.0 position.....155

Table 4.6 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 9 based on Glyma 2.0 position.....157

Table 4.7 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 11 based on Glyma 2.0 position.....167

Table 4.8 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 13 based on Glyma 2.0 position.....168

Table 4.9 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 14 based on Glyma 2.0 position. 176

Table 4.10 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 17 based on Glyma 2.0 position. 179

Table 4.11 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 19 based on Glyma 2.0 position. 180

Table 4.12 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 5 mb from stearic acid QTL on soybean chromosome 14 based on Glyma 2.0 position. 183

LIST OF FIGURES

Figure 2.1 Yield (kg ha^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 276 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, and G-BLUP) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes. 59

Figure 2.2 Palmitic acid (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes. 60

Figure 2.3 Stearic acid (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes. 61

Figure 2.4 Oeic acid (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes. 62

Figure 2.5 Linoleic acid (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes. 63

Figure 2.6 Linolenic acid (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes. 64

Figure 2.7 Protein (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 271 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and

G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes. 65

Figure 2.8 Oil (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 271 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes. 66

Figure 3.1 Yield (kg ha^{-1}) performance comparisons from only the Knoxville, TN location between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes. 113

Figure 3.2 Palmitic acid (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes. 114

Figure 3.3 Stearic acid (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes. 115

Figure 3.4 Oeic acid (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes. 116

Figure 3.5 Linoleic acid (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

Predictions with higher R2 were more closely related to 2014 observed phenotypes. 117

Figure 3.6 Linolenic acid (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes. 118

Figure 3.7 Protein (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes. 119

Figure 3.8 Oil (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes. 120

Figure 4.1 Quantitative trait loci (QTL) for yield detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against ‘ghost’ QTL (Martinez and Curnow, 1992).....184

Figure 4.2 Quantitative trait loci (QTL) for palmitic acid detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against ‘ghost’ QTL (Martinez and Curnow, 1992).....185

Figure 4.3 Quantitative trait loci (QTL) for stearic acid detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).....188

Figure 4.4 Quantitative trait loci (QTL) for oleic acid detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).....189

Figure 4.5 Quantitative trait loci (QTL) for linoleic acid detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).....190

Figure 4.6 Quantitative trait loci (QTL) for linolenic acid detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).....191

Figure 4.7 Quantitative trait loci (QTL) for protein detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010,

2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).....192

Figure 4.8 Quantitative trait loci (QTL) for oil detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).....193

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Soybean [Glycine max (L.) Merrill] is the leading oilseed crop grown in the world (Sharma et al., 2012). The oil content of soybean seed, approximately 200 g kg^{-1} of total composition, is of primary importance in many products. Products that use soybean oil are wide ranging and include both industrial (e.g. candles, ink, paint, plastic, and biodiesel) and food (e.g. vegetable oil, margarine, mayonnaise, and salad dressings) applications. While an overall increase in soybean oil would be beneficial for the industries mentioned above, it must be balanced carefully with other targeted goals of soybean production. Soy protein also has a wide range of uses, and it consists of approximately 400 g kg^{-1} of the total composition of soybean seed. There is much historical evidence that oil and protein in soybean seed are negatively correlated (Yaklich et al., 2002). Also, oil and yield share a positive relationship, while protein and yield have a negative relationship (Morrison et al., 2008). Thus, increases in soybean oil and yield must be sought after with the added goal of maintaining adequate protein levels (Cober et al., 2009) and protein quality (Panthee and Pantalone, 2006; Pantalone, 2012). Additionally, soybean oil would be far more useful to various industries with improvements in the relative abundances of particular soybean fatty acids to one another. Given these needs for soybean improvement, researchers must pursue breeding methods that can accomplish all of these goals in a complementary manner.

Five primary fatty acids compose soybean oil: palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1), linoleic acid (18:2) and linolenic acid (18:3). In normal soybean seeds these fatty acids exist in approximate relative concentrations of 100, 40, 220, 540, and 100 g kg^{-1} of total lipids, respectively (Wilson, 2004). Goals of fatty acid improvement are variable, but the current major goals are increasing oleic acid, decreasing linolenic acid, and decreasing the saturates. Fatty acid improvement goals may conflict depending on the end product. An example is the saturated fatty acids: reduction in palmitic acid and stearic acid is often sought due to negative cardiovascular effects in humans associated with

dietary consumption of saturated fatty acids. However, margarine production benefits from increased saturated fatty acid (Cober et al., 2009). When comparing saturated fatty acids, there is evidence that stearic acid is neutral with respect to cholesterol in humans, which differs from the hypercholesteromic effect of palmitic acid (Kris-Etherton and Yu, 1997). Thus, aside from specialty uses, improving soybean saturated fatty acid involves increasing stearic acid while reducing palmitic acid.

Another conflicting goal of fatty acid improvement is the polyunsaturated linolenic acid. The ink industry would benefit from increased linolenic acid, which is in contrast to modern vegetable oil goals requiring decreased linolenic acid (Cober et al., 2009). This decrease in linolenic acid demanded by vegetable oil production is due to the need for its hydrogenation, which is necessary to improve the oxidative stability of soybean oil. It is widely known that hydrogenated soybean oil, resulting in trans-fat formation, has deleterious effects on human cardiovascular health (Hunter, 2006). Strikingly, the U.S. Food and Drug Administration (FDA) has issued its final ruling that partially hydrogenated oils are no longer considered Generally Recognized as Safe (GRAS) for use in any food product (https://www.federalregister.gov/articles/2015/06/17/2015-14883/final-determination-regarding-partially-hydrogenated-oils). Thus decreasing linolenic acid usually takes precedence over its increase outside of specialty markets (Cober et al., 2009), and has become one of the major goals of soybean fatty acid improvement.

One goal that is shared by most researchers seeking to improve the fatty acid profile of soybean oil is the increase of monounsaturated oleic acid. In human consumption, oleic acid has been shown to lower cholesterol when compared with saturated fatty acids (Kris-Etherton and Yu, 1997). Additionally, soybean oil with enhanced oleic acid exhibits improved oxidative stability, resulting in increased shelf life of soybean oil food products (Kinney, 1996) and biodiesel

(Kinney and Clemente, 2005; Fallen et al, 2012). Thus, increasing the oleic acid content of soybean oil is a primary goal of soybean oil improvement.

Much research has taken place in the effort to increase soybean oil and improve its fatty acid profile. A recent study by Eskandari et al. (2013) detected quantitative trait loci (QTL) that increase both oil and protein simultaneously, as well as QTL that simultaneously increase oil and yield. Toward fatty acid improvement, several studies have identified mutations in delta-9-stearoyl-acyl carrier protein desaturase (SACPD) enzymes that result in increased levels of stearic acid (Pantalone et al., 2002; Boersma et al., 2012; Ruddle et al., 2013b; Gillman et al., 2014). Decreased levels of linolenic acid have been reported in soybean cultivars specifically bred for that goal, and recent efforts combining three individual omega-3 fatty acid desaturase (FAD3) genes have achieved lows near 10 g kg^{-1} for linolenic acid (Bilyeu et al., 2011). Additionally, greatly increased levels of oleic acid (> 800 g kg^{-1}) are available in soybean through both genetic engineering (Buhr et al., 2002) and mutant detection (Pham et al., 2010) of fatty acid desaturase 2 (FAD2) genes. These findings represent important milestones for soybean oil and fatty acid improvement. Further efforts seeking to characterize significant genetic regions for these traits and implement selection of desirable allele forms into high yielding cultivars will be essential for making desired improvements.

While the advances described above will be pivotal in the pursuit of soybean oil and fatty acid improvement, there are challenges that remain. One challenge is poor agronomic performance of mutant genotypes at key loci. An example is noted by Lee et al. (2012). To solve this problem, high oleic alleles must be bred into agronomically superior soybean cultivars (Lee et al., 2012). This work is in progress, with high oleic breeding lines competing closely with check cultivars for soybean yield (Thang et al. 2014). An illustration of poor agronomic performance for high stearic mutant alleles can be found in Ruddle et al. (2013a), in which

lines containing the SACPD-B mutant allele were outperformed in agronomic traits by lines with either the SACPD-C mutant allele or the wild type allele. Also, lines containing both mutants were agronomically inferior to those with only SACPD-B mutants (Ruddle et al., 2013a). An additional problem is that only with both mutant alleles combined can a stearic acid level of 146 g kg^{-1} be achieved (Ruddle et al., 2013b), which still falls short of the 200 g kg^{-1} stearic acid goal for solid fat uses. A study by Gillman et al. (2014) demonstrated that breeding lines with 100-150 g kg^{-1} stearic acid can be competitive with check cultivars in yield, but may come with the pleiotropic side effect of decreased root nodule development.

The quantitative inheritance of soybean fatty acids, oil, protein, and yield represents another breeding challenge for these traits. Examples of genetic complexity can be illustrated by the 237 and 57 genes listed in the gene ontology (GO) categories for fatty acid biosynthetic process (GO:0006633) and lipid biosynthetic process (GO:0008610), respectively (www.SoyBase.org, "GO Term Enrichment Tool", accessed 7/17/2015). In the case of stearic acid, competitively yielding breeding lines have yet to achieve the targeted goal of 200 g kg^{-1} of total oil (Gillman et al., 2014). Breeding methods incorporating a greater portion of genetic effects may help to increase stearic acid to the desired level. For oleic acid (> 800 g kg^{-1}) and linolenic acid (< 30 g kg^{-1}) targeted goals have been achieved through breeding efforts with relatively few loci. However, for oleic acid, environmental variation may result in levels that are below 800 g kg^{-1} (Lee et al., 2012; Fallen et al., 2012). Breeding methods accounting for a broader range of genetic effects could be useful in a fine-tuning approach to provide more consistent results. In addition to fatty acids, it would be useful to explore such breeding methods for oil, protein, and yield improvement.

Given the need to account for a greater range of genetic effects than is possible using a QTL approach, genomic selection (GS) would be a useful strategy to

investigate. Methodology for GS was first described by Meuwissen et al. (2001) as genetic maps increased in density due to the increased availability of molecular markers. Meuwissen (2007) defines GS as the simultaneous selection of many thousands of markers which densely cover the entire genome; and any gene affecting the targeted trait would be expected to occur in linkage disequilibrium with some markers. A key difference between GS and QTL based marker-assisted selection (MAS) is the amount of genetic information used for selection; GS uses the entire genome in contrast to MAS, which only accounts for genetic information from targeted regions (Nakaya and Isobe, 2012). Prediction accuracy of breeding values in GS simulation studies has been shown to be as high as 0.85 (Meuwissen et al., 2001), however, GS may have limited success for low heritability traits (Nakaya and Isobe, 2012). Given the potential for success in complex traits, numerous studies have continued to explore the potential of GS in animal and plant breeding with some success (Ødegård et al., 2009; Lillehammer et al., 2010; Resende et al., 2012; Poland et al., 2012; Sitzenstock et al., 2013; Crossa et al., 2014). With this potential comes the need for further research, such as testing the accuracy of GS over multiple generations, rather than only reporting cross-validation results from the same generation as has often been done in crop studies (Jonas and de Koning, 2013).

Increasing the oil content and improving the fatty acid profile of soybean seed remain important breeding goals. Incorporating these traits into high yielding genetic backgrounds while maintaining or increasing seed protein content are critical for any improvements in oil and fatty acids to be adapted on a wide scale. Detection and confirmation from previous studies, as well as genetic characterization of significant regions will be influential to the improvement of these traits. Additionally, comparison of molecular breeding methods with each other and with phenotypic selection will be important in determining how to achieve maximum improvement while selecting for quantitative traits. Toward these ends, this research will seek to detect and confirm significant genetic

regions impacting important soybean traits, and to evaluate breeding strategies for maximum improvement.

References

Bilyeu, K., J.D. Gillman, and A.R. LeRoy. 2011. Novel FAD3 mutant allele combinations produce soybeans containing 1% linolenic acid in the seed oil. Crop Sci. 51:259–264.

Boersma, J.G., J.D. Gillman, K.D. Bilyeu, G.R. Ablett, C. Grainger, and I. Rajcan. 2012. New mutations in a delta-9-stearoyl-acyl carrier protein desaturase gene associated with enhanced stearic acid levels in soybean seed. Crop Sci. 52:1736–1742.

Buhr, T., S. Shirley, F. Ebrahim, A. Xing, Y. Zhou, M. Mathiesen, B. Schweiger, A. Kinney, P. Staswick, and T. Clement. 2002. Ribozyme termination of RNA transcripts down-regulate seed fatty acid genes in transgenic soybean. Plant J. 30:155–163.

Cober, E.R., S.R. Cianzio, V.R. Pantalone, and I. Rajcan. 2009. Soybean. In: J. Vollman and I. Rajcan, editors, Oil crops: Handbook of plant breeding, volume 4. Springer Science + Business Media LLC. p. 57-90.

Crossa, J., P. Pérez, J. Hickey, J. Burgeño, L. Ornella, J. Cerón-Rojas, X. Zhang, S. Dreisigacker, R. Babu, Y. Li, D. Bonnett, and K. Mathews. 2014. Genomic prediction in CYMMIT maize and wheat breeding programs. Heredity 112:48-60.

Eskandari, M., E.R. Cober, and I. Rajcan. 2013. Genetic control of soybean seed oil: II. QTL and genes that increase oil concentration without decreasing protein or with increased yield. Theor. Appl. Genet. 126:1677–1687.

Fallen, B.D., K. Rainey, C.E. Sams, D.A.Kopsell, and V.R. Pantalone. 2012. Evaluation of agronomic and seed characteristics in elevated oleic acid soybean lines in the south-eastern US. J. Am. Oil Chem. Soc. 89:1333-1343.

Federal Register. 2015. Final determination regarding partially hydrogenated oils. https://www.federalregister.gov/articles/2015/06/17/2015-14883/final-determination-regarding-partially-hydrogenated-oils (accessed 24 July 2015).

Gillman, J.D., M.G. Stacy, Y. Cui, H.R. Berg, and G. Stacey. 2014. Deletions of the SACPD-C locus elevate seed stearic acid but also result in fatty acid and morphological alterations in nitrogen fixing nodules. BMC Plant Biol. 14:143.

Grant, D., R.T. Nelson, S.B. Cannon, and R.C. Shoemaker. 2010. SoyBase, the USDA-ARS soybean genetics and genomics database. Nucl. Acids Res. 38:D843-D846.

Hunter, J.E. 2006. Dietary trans fatty acids: Review of recent human studies and food industry responses. Lipids 41:967–992.

Jonas, E., and D.J. de Koning. 2013. Does genomic selection have a future in plant breeding? Trends Biotechnol. 31:497-504.

Kinney, A.J. 1996. Development of genetically engineered soybean oils for food application. J. Food Lipids. 3:273-292.

Kinney, A.J., and T.E. Clemente. 2005. Modifying soybean oil for enhanced performance in biodiesel blends. Fuel Pro. Technol. 86:1137–1147.

Kris-Etherton, P.M., and S. Yu. 1997. Individual fatty acid effects on plasma lipids and lipoproteins: Human studies. Am. J. Clin. Nutr. 65:S1628–S1644.

Lee, J.D., K.D. Bilyeu, V.R. Pantalone, A.M. Gillen, Y.S. So, and J.G. Shannon. 2012. Environmental stability of oleic acid concentration in seed oil for soybean lines with FAD2-1A and FAD2-1B mutant genes. Crop Sci. 52:1290–1297.

Lillehammer, M., T.H.E. Meuwissen, and A.K. Sonesson. 2011. A comparison of dairy cattle breeding designs that use genomic selection. J. Dairy Sci. 94:493-500.

Meuwissen, T. 2007. Genomic selection: Marker assisted selectionon a genome wide scale. J. Anim. Breed. Genet. 124:321–322.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.

Morrison, M.J., E.K. Cober, M.F. Saleem, N.B. McLaughlin, J. Fregeau-Reid, B.L.

Ma, W. Yan, and L. Woodrow. 2008. Changes in isoflavone concentration with 58 years of genetic improvement of short-season soybean cultivars in Canada. Crop Sci. 48:2201–2208.

Nakaya, A., and S.N. Isobe. 2012. Will genomic selection be a practical method for plant breeding? Ann. Bot. 110:1303-1316.

Ødegård, J., A.K. Sonesson, M.H. yazdi, and T.H.E. Meuwissen. 2009. Introgression of a major QTL from an inferior into a superior population using genomic selection. Genet. Sel. Evol. 41:38.

Pantalone, V.R. 2012. Modern breeding approaches for enhancing soybean protein quality. In: R.F. Wilson, editor, Designing soybean for 21st century markets. AOCS Press, Urbana, IL. p. 197-226.

Pantalone, V.R., R.F. Wilson, W.P. Novitzky, and J.W. Burton. 2002. Genetic regulation of elevated stearic acid concentration in soybean oil. J. Am. Oil Chem. Soc. 79:543–553.

Panthee, D.R., and Pantalone, V.R. 2006. Registration of soybean germplasm lines TN03-350 and TN04-5321 with improved protein concentration and quality. Crop Sci. 46:2328-2329.

Pham, A.T., J.D. Lee, J.G. Shannon, and K.D. Bilyeu. 2010. Mutant alleles of FAD2-1A and FAD2-1B combine to produce soybeans with the high oleic acid seed oil trait. BMC Plant Biol. 10:195.

Poland, J., J. Endelman, J. Dawson, J. Rutkoski, S. Wu, Y. Manes, S. Dreisigacker, J. Crossa, H. Sanchez-Villeda, M. Sorrells, and J.L. Jannink. 2012. Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Gen. 5:103–113.

Resende Jr., M.F.R., P. Muñoz, J.J. Acousta, G.F. Peter, J.M. Davis, D. Grattapaglia, M.D.V Resende, and M. Kirst. 2012. Accelerating the domestication of trees using genomic selection: Accuracy of prediction models across ages and environments. New Phytol. 193:617-624.

Ruddle II, P., A. Cardinal, R.G. Upchurch, C. Arellano, and L. Miranda. 2013a. Agronomic effects of mutations in two soybean Δ9-stearoyl-acyl carrier

protein-desaturases. Crop Sci. 53:1887-1893.

Ruddle II, P., R. Whetten, A. Cardinal, R.G. Upchurch, and L. Miranda. 2013b. Effect of a novel mutation in a Δ9-stearoyl-ACP-desaturase on soybean seed oil composition. Theor. Appl. Genet. 126:241–249.

Sharma, M., S.K. Gupta, and A.K. Mondal. 2012. Production and trade of major world oil crops. In: S. K. Gupta, editor, Technological innovations in major world oil crops, volume 1. Springer Science + Business Media LLC. p. 1-15.

Sitzenstock, F., F. Ytournel, A.R. Sharifi, D. Cavero, H. Täubert, R. Preisinger, and H. Simianer. 2013. Effeciency of genomic selection in an established commercial layer breeding program. Gent. Sel. Evol. 45:29.

Thang, C.L., S.M. Pathan, T. Vuong, J. Lee, A.M. Scaboo, J.R. Smith, A.M. Gillen, J. Gillman, M.R. Ellersieck, H.T. Nguyen, and J.G. Shannon. 2014. Effect of high-oleic acid soybean on seed oil, protein concentration, and yield. Crop. Sci. 54:2054-2062.

Wilson, R.F. 2004. Seed composition. In: H.R. Boerma and J.E. Specht, editors, Soybeans: Improvement, production, and uses. 3rd ed. ASA, CSSA, and SSSA, Madison, WI. p. 621–678.

Yaklich, R.W., B. Vinyard, M. Camp, and S. Douglass. 2002. Analysis of seed protein and oil from soybean northern and southern region uniform tests. Crop Sci. 42:1504–1515.

CHAPTER 2

MOLECULAR BREEDING STRATEGIES OUTPERFORM PHENOTYPIC SELECTION FOR SOYBEAN QUANTITATIVE TRAITS IN THE PROGENY ROW STAGE

Abstract

Evaluating different selection methods for relative utility is necessary in order to choose those which maximize breeding results. Soybean [Glycine max (L.) Merrill] yield, fatty acids, protein, and oil are all commercially important traits that display quantitative inheritance. Thus, it is of interest to evaluate breeding methods for these traits that can account for the entire genome. In addition to phenotypic selection (PS), the molecular breeding methods chosen for this study were BayesB, G-BLUP, Epistacy, BayesB(Epi), and G-BLUP(Epi). These methods were evaluated in a soybean population consisting of 860 F5 derived recombinant inbred lines (RILs), which was genotyped with 17,236 polymorphic SNPs using the Illumina Infinium beadchip SoySNP50K. In order to simulate progeny rows, each RIL was grown in a single plot in 2010 in Knoxville, TN and phenotyped. The combined phenotypic and genotypic datasets were used to make predictions with the methods mentioned above. A subset of 276 RILs from this population was then grown in multi-location, replicated field trials in 2013 in order to evaluate the relative utility of each selection method; Spearman correlations and 15% tail selection contrasts were used for comparison. For yield, Epistacy was the preferred method; however for all other traits Epistacy was the least influential method. For the fatty acids, BayesB and G-BLUP were the best methods, with the slight overall edge going to BayesB. For protein, the preferred method was G-BLUP and for oil the preferred method was BayesB(Epi). Notably, for each trait the preferred method was a molecular selection strategy. This provides important implications for how soybean breeders could maximize selections from the progeny row stage for yield, fatty acids, protein, and oil.

Introduction

Soybean [Glycine max (L.) Merrill] is a major crop produced globally for a wide range of purposes. Protein (~400 g kg^{-1}) and oil (~200 g kg^{-1}) are major components of soybean seed that contribute to its high value. Historically, oil

and protein in soybean seed are negatively correlated (Yaklich et al., 2002). Oil and yield share a positive relationship, and protein and yield have a negative relationship (Morrison et al., 2008). Because of this, increases in soybean oil and yield must be sought after while simultaneously seeking to maintain adequate protein levels (Cober et al., 2009).

Within soybean oil there are five primary fatty acids; palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2) and linolenic (18:3). These typically occur in relative concentrations of 100, 40, 220, 540, and 100 g kg^{-1} of total lipids, respectively (Wilson, 2004). Improving the fatty acid profile in soybean is gaining importance, particularly with the Food and Drug Administration (FDA) recently banning partially hydrogenated oils (PHOs) as no longer Generally Recognized as Safe (GRAS) (https://www.federalregister.gov/articles/2015/06/17/2015-14883/final-determination-regarding-partially-hydrogenated-oils). Due to this ban in all food products, a primary goal of fatty acid improvement is to reduce linolenic acid (< 30 g kg^{-1}), thus reducing the need to partially hydrogenate soybean oil. Coinciding with this is the goal of increasing monounsaturated oleic acid (> 800 g kg^{-1}). Oleic acid has been shown to lower cholesterol when compared with saturated fatty acids in human consumption (Kris-Etherton and Yu, 1997). Additionally, soybean oil with increased oleic acid leads to improved oxidative stability, resulting in increased shelf life of soybean oil food products (Kinney, 1996) and biodiesel (Kinney and Clemente, 2005; Fallen et al, 2012). While much recent work has occurred in the improvement of soybean fatty acids (Pantalone et al., 2002; Pham et al., 2010; Bilyeu et al., 2011; Boersma et al., 2012; Gillman et al., 2014), there is still a need for continued advancement.

Given the need to improve the soybean yield, fatty acids, protein, and oil, it is essential to evaluate the relative utility of various selection strategies. Since each of these traits displays quantitative inheritance, it is important to evaluate selection methods that can account for a broad range of genetic effects.

Targeted goals have been achieved for oleic acid and linolenic acid using relatively few loci (Pham et al., 2010; Bilyeu et al. 2011). However, for oleic acid there is still concern that environmental variation may result in levels that are below 800 g kg^{-1} (Lee et al., 2012; Fallen et al., 2012). In such cases, it would be useful to evaluate breeding methods that account for a broader range of genetic effects, acting as a fine-tuning approach to provide more consistent results. In addition to fatty acids, such methods would also be worth exploring for oil, protein, and yield improvement.

Since quantitative trait loci (QTL) based selection strategies only account for a limited amount of genetic information, a more robust method such as genomic selection (GS), which accounts for the entire genome (Nakaya and Isobe, 2012), would be worth investigating. First described by Meuwissen et al. (2001), GS is the simultaneous selection of many thousands of markers which densely cover the entire genome; with any gene affecting the targeted trait expected to occur in linkage disequilibrium with some markers (Meuwissen, 2007). Numerous studies have explored the potential of GS in animal and plant breeding with some success (Ødegård et al., 2009; Lillehammer et al., 2010; Resende et al., 2012; Poland et al., 2012; Sitzenstock et al., 2013; Crossa et al., 2014). Given this potential, there is a need to evaluate the accuracy of GS over multiple generations rather than only reporting cross-validation results from the same generation, as has been common in crop studies (Jonas and de Koning, 2013). Thus, the purpose of this research will be to evaluate the relative utility for soybean yield, fatty acids, protein, and oil with various GS strategies in comparison with other molecular breeding strategies and with phenotypic selection (PS).

Materials and Methods

Plant Materials

A population of 860 recombinant inbred lines (RILs) with both genotypic and phenotypic data was developed from the cross between 'Essex' and 'Williams 82' (hereafter known as E×W-50K). Essex is a maturity group (MG) V soybean cultivar with a determinate growth habit, purple flower, and gray pubescence (Smith and Camper, 1973), while Williams 82 is an MG III soybean cultivar with indeterminate growth habit, white flower, and tawny pubescence (Bernard and Cremeens, 1988). The seed of Essex and Williams 82 were obtained from the USDA soybean germplasm collection (www.ars-grin.gov), and a random single plant of each parental line was intentionally selfed for two generations to provide highly homozygous parental lines to be crossed for RIL development. The initial cross was made in the summer of 2005 at the East Tennessee Research and Education Center (ETREC) in Knoxville, TN. The hybrid seed resulting from the cross were harvested in the fall of 2005 and grown as F1 single plants in Puerto Rico at the Tropical Agricultural Research Station (TARS) in Isabela, Puerto Rico, in the winter of 2005-06. Following the single seed descent method (Brim, 1966), the population was advanced from the F2 to the F5 generation. In the summer of 2009, F5 plants were grown in Beltsville, MD, and leaf tissue was collected individually from each plant. Seed harvested from each plant was used to grow F5:6 plant rows in Homestead, FL in the fall of 2009. The F5:6 plant rows were harvested individually and planted as F5:7 RIL in the summer of 2010 in Knoxville, TN.

Each entry in the 2010 field test was planted in a single plot consisting of two adjacent rows 6.1 m in length, with the rows spaced 0.8 m apart. Along with the RIL and the parents, four checks with relevant maturities were included in the 2010 field test. The checks were 'LD00-3309' (MG IV-early) (Diers et al., 2006), 'IA4004' (MG IV-early), '5002T' (MG V-early) (Pantalone et al., 2004), and

‘5601T’ (MG V-mid) (Pantalone et al., 2003). Flower color was determined at the R2 growth stage; pubescence color, plant height, and maturity were determined at the R8 growth stage (Fehr and Caviness, 1977).

The 2010 RIL maturity recorded in Julian calendar date ranged from 251-288 (Table 2.1) (note: all tables and figures in this chapter are located in Appendix B). In order to narrow the maturity range for replicated field testing, 276 RIL with maturities ranging from 266-273 (approximately MG IV-mid to IV-late) were chosen for advancement. The MG IV-mid to IV-late range is of primary importance to Tennessee soybean producers, as evidenced by the number of lines tested in this maturity range relative to others in the Soybean Variety Performance Tests in Tennessee (Allen et al., 2011; 2012, and 2013). Seed harvested from these 276 RIL plots in 2010 was increased as F5:8 RIL in Homestead, FL in the fall-winter of 2012-2013. The resulting F5:9 RIL were planted in 2013 in a randomized complete block design (RCBD) with three replications per environment at three environments (Knoxville, TN; Springfield, TN; and Milan, TN), representative of the eco-geographic regions of East, Middle, and West Tennessee, respectively. In addition to the RILs and parents, three maturity checks were included; ‘LD00-3309’ (MG IV-early), ‘LD00-2817P’ (MG IV-mid) (Diers et al., 2010), and ‘Ellis’ (MG IV-late). As in the 2010 field test, flower color was determined at the R2 growth stage; and pubescence color, plant height, and maturity were determined at the R8 growth stage (Fehr and Caviness, 1977). For both field seasons, plots were harvested at maturity. Yield was measured in kg ha^{-1} after adjusting the plot weight to 13% moisture.

Seed Quality Trait Detection

Fatty acid measurements for each plot from the 2010 and 2013 field tests for palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid were done using gas chromatography with a procedure described by Spencer et al. (2004). This analysis was performed using a Hewlett Packard HP 6890 series gas

chromatograph (Agilent Technologies, Santa Clara, CA) system equipped with a 7683 auto sampler, a 7673 flame ionization detector, an immobilized 30 m x 0.53 mm inner diameter Agilent DB-23 capillary column with 0.5 µm fused stationery phase. Fatty acid estimates were obtained as percentage of seed oil, and converted to g kg^{-1} seed oil.

Following harvest from the 2010 growing season, approximately 25 g of seed from each plot were ground for 20 sec in a Knifetec 1095 Sample Mill (FOSS Tecator, Hoganas, Sweden) to produce ground whole soybean with a uniform consistency and particle size. Samples were analyzed for protein and oil content using the near infrared reflectance spectroscopy (NIRS) instrument (NIR 6500, FOSS North America) as described by Panthee et al. (2006), except that for this study the ground samples were scanned using updated ISIscan software v. 2.85. Plots from the 2013 season were scanned as whole bean samples using a Perten DA 7200 Diode Array (Perten, Hägersten, Sweden) NIRS instrument in collaboration with the University of Minnesota. The calibration equations used for analysis were developed through a cooperative effort between Perten and University of Minnesota (Bolon et al., 2011). For each NIRS analysis, values for protein and oil concentration were adjusted to g kg^{-1} seed on a dry weight basis.

SNP Genotyping and Marker Cleanup

In 2009, samples of DNA were collected from crushed leaves of each F5 greenhouse plant from this population at the Soybean Genomics Laboratory at the USDA Beltsville Agricultural Research Center (USDA-ARS) in Beltsville, MD. The DNA samples were analyzed using the Illumina Infinium beadchip SoySNP50K (Song et al., 2013), with marker positions from the Williams 82 whole genome sequence (WGS) (Glyma 1.01) (Schmutz et al., 2010). In this population there were a total of 17,236 polymorphic SNP markers used for genotyping. These SNPs were screened for missing data using the ‘synbreed’ package (Wimmer et al., 2012) in the R language and environment for statistical

computing (R Core Team, 2015). Markers that were missing in ≥ 5% (43 RILs) of the population were dropped, resulting in 16,718 SNPs used for analysis. Missing data from the remaining SNPs was then imputed using default settings in the Beagle Genetic Analysis Software Package v. 3.3.1 (Browning and Browning 2007; 2009) via the 'synbreed' package (Wimmer et al., 2012). Finally, potential genotyping errors were screened using default settings in the new version of the 'calc.errorlod' function within the 'qtl' package (Broman et al., 2003) in the R language and environment for statistical computing (R Core Team, 2015). No genotyping errors were detected.

Selection Methods and Statistical Analysis

Genomic selections (GS) were performed using the 'BGLR' package (Pérez and de los Campos, 2014) in the R language and environment for statistical computing (R Core Team, 2015). The 860 RILs planted in single rep plots simulating progeny rows in Knoxville, TN, were used to generate predictions for yield, fatty acids (palmitic, stearic, oleic, linoleic, and linolenic), protein, and oil. Because there were no replicates, plots with missing phenotypic data were dropped from the analysis; thus yield, fatty acids, and protein and oil were tested with 860, 855, and 824 RILs, respectively. Since this population segregates at the E1 (maturity), E3 (maturity), and Dt1 (growth habit) loci, SNPs located within or adjacent to (< 5 kb) each locus (ss715593840, ss715635705, and ss715635423 or if missing ss715635422, respectively) were used to predict the expected parental allele. Four exceptions were made at the Dt1 locus, with growth habit call updated based on phenotype. The three loci (E1, E3, and Dt1) were included as covariates in the GS models in order to minimize any associated noise in variability.

The GS models chosen for analysis were genomic best linear unbiased predictor (G-BLUP) and BayesB (Meuwissen et al., 2001). The reason for choosing G-BLUP is that it is a common method in GS studies due to the ease of

implementation; while BayesB was chosen due to its common use, ability for departure from the infinitesimal model, and preference in comparison with Bayesian penalized regression models (de los Campos et al., 2013). Both GS and BayesB included the three covariates mentioned above as fixed effects as well as 40,000 iterations and a burn in of 10,000; otherwise the default settings were used. Cross-validations were replicated 50 times for each trait. In each rep, a randomly chosen 1/5 of the population had phenotypic data removed (test set), while phenotypic and genotypic information were retained for the remaining 4/5 of the population (training set). Since both prediction methods shared the training and test set partitioning for each of the 50 cross-validations, the Pearson correlation coefficients were compared using a paired-t test (Pérez and de los Campos, 2014).

An additional selection model was performed using the Epistacy macro v. 2.0 (Holland, 1998) in SAS (SAS Institute Inc., Cary, NC, USA, SAS 9.4, 2002-2012), with modifications provided by Arnold Saxton. This model was included in the analysis as an effort to account for significant epistatic interactions that influence yield, fatty acids, protein and oil. A P value cutoff of 0.001 was chosen to select significant interactions. In order to limit the influence of duplicate interactions, markers were screened for variation among RILs using the 'findDupMarkers' function in the R 'qtl' package (Broman et al., 2003), with one marker randomly chosen from each duplicate set to remain for analysis. After removing duplicate markers, 6900 SNPs remained. Deviations due to these interactions for each RIL were then summed, divided by 6900, and added to the mean in order to predict expected performance.

Further predictions were estimated using combined output from the GS and Epistacy models. For these estimates, multiple regression was performed for each GS method, along with the 20 Epistacy interactions with largest effect against the phenotypic values using the REG procedure in SAS (SAS Institute

Inc., Cary, NC, USA, SAS 9.4, 2002-2012) in order to identify the combination of GS method and Epistacy interactions that would produce the highest adjusted R squared value. The predictions from these regressions were then used to produce the G-BLUP(Epi) and BayesB(Epi) selection methods. Inclusion of these methods in the analysis was done in order to account for large effect epistatic interactions in the GS models. Both G-BLUP(Epi) and BayesB(Epi) were evaluated for each trait except yield, because no combination of interactions added to the GS methods were able to produce a greater adjusted R squared value than the GS method alone.

The performance of each molecular breeding method (BayesB, G-BLUP, Epistacy, BayesB(Epi), and G-BLUP(Epi)), along with phenotypic selection (PS), were then evaluated in the 276 RIL population subset grown in 2013. To visualize the degree of relationship with the 2013 observed phenotypes, a regression was plotted for each selection against 2013 observed values in the R language and environment for statistical computing (R Core Team, 2015). Additionally, the Spearman correlations between each selection method with the 2013 observed phenotypes were compared to one another using the 'cocor.dep.groups.overlap' function in the 'cocor' package (Diedenhofen and Musch, 2015) in R using the Hittner et al. (2003) method for comparing dependent, overlapping correlations. Finally, 15% (41 RILs) high and low tail selections chosen using each selection method were evaluated for performance in the 2013 field season. These 15% tails were compared to each other and to the 15% high or low tail from the 2013 phenotypic rankings using 'estimate' and 'contrast' statements in SAS PROC GLIMMIX (SAS Institute Inc., Cary, NC, USA, SAS 9.4, 2002-2012). The model used for analysis was an RCBD, with RIL as the fixed term and location, rep(location), and RIL × location as random terms, and denominator degrees of freedom method set to residual.

Additional statistical analyses were performed in SAS PROC GLIMMIX (SAS Institute Inc., Cary, NC, USA, SAS 9.4, 2002-2012) using the terms from the model above, but with some adjustments to fixed and random terms. A model with RIL, location, and RIL × location as fixed terms and rep(location) as random was run in order to test for the significance of RIL × location interactions. Also, a model with no fixed terms and RIL, location, rep(location), and RIL × location as random terms was run in order to obtain the variance for each term. These variances were then used to estimate heritability on an entry means basis (Nyquist, 1991).

Results

Yield, fatty acids, protein, oil, maturity, and height were all measured in the E×W-50K 860 RIL soybean population grown in Knoxville, TN in 2010. Variability was observed among RILs for each trait, although this was not supported by statistical analysis since this was a single rep field test (Table 2.1). However, the non-replicated nature of the 2010 field test provided an opportunity to evaluate selection methods for advancement from a progeny row stage into multi-location replicated field trials. Prior to advancing into multi-location replicated field trials in 2013, the population was subset to 276 RILs based on maturity in order to eliminate variability due to different harvest dates. The simple statistics for the 2013 field test are displayed in Table 2.2, with each trait displaying a significant difference ($P < 0.05$) among RILs. While the fatty acid traits did not display the same extreme values as some recent studies (Pantalone et al., 2002; Pham et al., 2010; Bilyeu et al., 2011; Boersma et al., 2012; Gillman et al., 2014) they were still of interest for evaluating the whole genome selection methods in this study. Of the traits chosen for selection method evaluation, three different phenotyping methods were used; recorded weight (yield), GC (fatty acids), and NIRS (protein and oil). Additionally, these traits displayed a range of heritability values (Table 2.2), with yield as the lowest at 0.63, followed by NIRS traits (0.87), and finally GC traits (0.92-0.94), indicating differing degrees of gain from

selection. Thus, it was of interest to evaluate different selection methods for yield, fatty acids, protein, and oil.

For the six different selection methods chosen (PS, BayesB, G-BLUP, Epistacy, BayesB(Epi), and G-BLUP(Epi)), an initial comparison was available for BayesB and G-BLUP using a cross-validation approach. Following this approach, no advantage for BayesB or G-BLUP was observed ($P > 0.05$) for any of the traits (Table 2.3). The cross-validation correlations ranged from 0.41-0.55, with none of the differences for individual traits exceeding 0.02 (Table 2.3).

Regression plots are displayed for each trait and selection method (Figures 2.1-2.8) in order to visualize the relationship between the 2010 predictions and the 2013 observed phenotypes. For yield, each selection method displayed a weak relationship with the 2013 observed phenotypes, with R squared values ranging from 0.036 (BayesB) to 0.062 (Epistacy) (Figure 2.1). For the fatty acids; palmitic (Figure 2.2), stearic (Figure 2.3), oleic (Figure 2.4), linoleic (Figure 2.5), and linolenic (Figure 2.6) ranged in R squared values from 0.21-0.62, 0.41-0.60, 0.08-0.74, 0.19-0.69, and 0.13-0.65, respectively. The same R squared trend was observed in palmitic (Figure 2.2), oleic (Figure 2.4), linoleic (Figure 2.5), and linolenic (Figure 2.6); with Epistacy representing the lowest value and BayesB and G-BLUP tying for the highest value in each case. This trend in R squared values was not continued for stearic (Figure 2.3); with PS (0.41) displaying the lowest value and BayesB (0.60) displaying the highest. The R squared values for the NIRS traits were somewhat lower than those for the fatty acids, with values for protein (Figure 2.7) and oil (Figure 2.8) ranging from 0.13-0.31 and 0.16-0.34, respectively. For both protein and oil, Epistacy displayed the lowest R squared value. However, the highest R squared values differed between protein (Figure 2.7) and oil (Figure 2.8), with G-BLUP (0.31) having the highest value for protein and BayesB, BayesB(Epi), and G-BLUP(Epi) (0.34) all tied with the highest R squared values.

Spearman correlations were performed between the 2013 observed phenotypes and each individual selection method for all traits in order to understand the relationship between rankings. These correlations were then compared against one another using the Hittner et al. (2003) method for comparing dependent, overlapping correlations in order to test for differences between correlations (Tables 2.4-2.11). For yield (Table 2.4), there was only one difference in correlations, with G-BLUP (0.15) displaying a greater association to the 2013 phenotypic rankings than BayesB (0.13).

For the fatty acids (Tables 2.5-2.9), Epistacy displayed the least similarity in rank correlations with 2013 phenotypes, with lower values in every correlation comparison ($P < 0.05$) except stearic PS (0.67), in which there was no difference ($P > 0.05$) from stearic Epistacy (0.69) (Table 2.6). After Epistacy, PS displayed the least similarity in rankings to the 2013 phenotypes, with lower values ($P < 0.05$) than BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi) for each fatty acid (Tables 2.5-2.9). For palmitic (Table 2.5), stearic (Table 2.6), and oleic (Table 2.7) BayesB displayed higher correlation values ($P > 0.05$) with 2013 phenotypic rankings in more comparisons than any other method. The BayesB correlation (0.80) was higher than all other selection methods ($P < 0.05$) for palmitic (Table 2.5). For stearic (Table 2.6), only BayesB(Epi) (0.79) displayed no difference ($P > 0.05$) from BayesB (0.80); while for oleic (Table 2.7), both G-BLUP (0.87) and BayesB(Epi) (0.86) were not different ($P > 0.05$) from BayesB (0.87). For linoleic (Table 2.8), BayesB(Epi) (0.83) displayed higher correlations in more comparisons than any other method; exceeding ($P < 0.05$) PS (0.75), Epistacy (0.41), and G-BLUP(Epi) (0.83), but not differing from ($P > 0.05$) BayesB (0.84) and G-BLUP (0.84). Linolenic correlations (Table 2.9) with BayesB (0.80) and G-BLUP (0.80) displayed higher values in more comparisons than other methods; exceeding ($P < 0.05$) PS (0.67), Epistacy (0.39), and BayesB(Epi) (0.79), while not differing from ($P > 0.05$) each other or from G-BLUP(Epi) (0.79).

Protein (Table 2.10) and oil (Table 2.11) correlations with 2013 phenotypes displayed far fewer differences between selection methods than the fatty acid traits (Tables 2.5-2.9). For both protein (Table 2.10) and oil (Table 2.11), Epistacy displayed lower correlation values ($P < 0.05$) than all other selection methods. The only other difference between selection methods was protein (Table 2.10), in which G-BLUP (0.50) exceeded BayesB (0.50).

Contrasts are provided in order to compare 15% high and low tail selections for each method with all traits (Tables 2.12-2.19). For each trait, the 15% high and low tails for 2013 phenotypic rankings significantly differed ($P < 0.05$) in the direction of selection from all methods (Tables 2.12-2.19). Epistacy was the best method for yield (Table 2.12), outperforming ($P < 0.05$) BayesB and G-BLUP in both tails, and PS in the low tail. No other differences ($P > 0.05$) for high or low tail selection occurred in yield (Table 2.12).

In each fatty acid except stearic, Epistacy high and low tail selections were outperformed ($P < 0.05$) by all other methods (Tables 2.13, 2.15-17). In the high tail selections for palmitic (Table 2.13), BayesB(Epi) (129.98 g kg^{-1}) and G-BLUP(Epi) (129.86 g kg^{-1}) were the premier methods, exceeding all other methods ($P > 0.05$) and not differing ($P < 0.05$) from each other. However, for the palmitic low tail selections (Table 2.13), BayesB (111.46 g kg^{-1}) and G-BLUP (111.46 g kg^{-1}) outperformed all other methods ($P < 0.05$) while not differing from each other ($P > 0.05$). For stearic (Table 2.14), BayesB (44.06 g kg^{-1}) and G-BLUP (43.99 g kg^{-1}) outperformed all other methods ($P < 0.05$) in the high tail but were not different from each other ($P > 0.05$). The only other significant differences for stearic (Table 2.14) were that BayesB(Epi) outperformed ($P < 0.05$) G-BLUP(Epi) in both tails. For oleic (Table 2.15), BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi) outperformed ($P < 0.05$) PS in both tails while not differing from each other in the high tail ($P > 0.05$). However, BayesB (315.51 g

kg^{-1}) was lower than ($P < 0.05$) both BayesB(Epi) (317.00 g kg^{-1}) and G-BLUP(Epi) (317.29 g kg^{-1}), thus winning more comparisons than any other method for oleic (Table 2.14). For linoleic acid (Table 2.16), G-BLUP (433.62 g kg^{-1}) was the best method in the high tail, outperforming ($P < 0.05$) all others except PS (432.27), for which there was no difference ($P > 0.05$). This trend for linoleic (Table 2.16), did not continue in the low tail, with BayesB (359.95 g kg^{-1}) , G-BLUP (360.04 g kg^{-1}), BayesB(Epi) (359.50 g kg^{-1}), and G-BLUP(Epi) (359.19 g kg^{-1}) all outperforming ($P < 0.05$) PS (369.46 g kg^{-1}) but not differing from each other ($P > 0.05$). For linolenic (Table 2.17), BayesB and G-BLUP were the preferred methods, outperforming ($P < 0.05$) PS in both tails, BayesB(Epi), and G-BLUP(Epi) in the low tail, and not differing from each other ($P > 0.05$) in either tail.

Contrast statements for protein (Table 2.18) showed that BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi) were all equivalent as the best method; outperforming ($P < 0.05$) Epistacy in both tails, PS in the high tail, and not differing from each other ($P > 0.5$) in either tail. BayesB (223.47 g kg^{-1}) and BayesB(Epi) (223.64 g kg^{-1}) were the preferred methods for oil in the high tail (Table 2.19), displaying higher values ($P < 0.05$) than G-BLUP (223.29 g kg^{-1}), Epistacy (222.65 g kg^{-1}), and G-BLUP(Epi) (223.15 g kg^{-1}), while not differing ($P > 0.05$) from each other or from PS (223.26 g kg^{-1}). However, in the low tail for oil (Table 2.19) G-BLUP(Epi) (212.47 g kg^{-1}) was the best method, with lower values ($P < 0.05$) than BayesB (212.77 g kg^{-1}), G-BLUP (212.77 g kg^{-1}), and Epistacy (214.87 g kg^{-1}), while not differing from PS (212.76 g kg^{-1}) or BayesB(Epi) (212.60 g kg^{-1}).

Discussion

Recommendations for selection method based on Spearman correlations with 2013 observed phenotypes and contrasts between 15% tail selections vary depending on the trait. Yield differed from the other traits in this study for several

reasons, including; method of phenotyping (recorded weight), lower heritability (Table 2.2), and fewer selections methods (PS, BayesB, G-BLUP, and Epistacy). Soybean heritability for yield has previously been demonstrated to be lower than for protein and oil (Wiggins, 2012). These differences in heritability, along with the possibility of greatly influencing fatty acid traits based on few loci (Pantalone et al., 2002; Pham et al., 2010; Bilyeu et al., 2011; Boersma et al., 2012; Gillman et al., 2014) with no comparable studies for yield demonstrate the highly quantitative nature of soybean yield, and subsequently highlight the extreme challenge in making selections for yield improvement. As noted by Nakaya and Isobe (2012), GS methods for low heritability traits may be prone to limited success. This was certainly true in this study, with BayesB and G-BLUP both displaying low values ($R < 0.15$) when rank correlated with 2013 observed yield (Table 2.4). However it should be noted that both GS methods did not significantly differ from PS with regard to Spearman correlations or 15% tail selections (Tables 2.4, 2.12) for yield. Overall, Epistacy was the most influential selection method for soybean yield. Yield Spearman correlations for Epistacy did not differ from other methods ($P > 0.05$) (Table 2.4), yet outperformed ($P < 0.05$) or equaled ($P > 0.05$) other methods with regard to 15% tail selections (Table 2.12).

In contrast with yield, Epistacy was the least desirable method for each fatty acid. This could indicate that epistatic interactions are more influential for determining soybean yield in comparison with fatty acids. Adjustments to the Epistacy selection approach would be necessary in order to be a worthwhile selection strategy for fatty acids. A possible solution would be to relax the significance threshold for choosing which interactions could be included for selection, thus accounting for a greater degree of overall genomic effects. However, this could require some fine tuning; with all interactions between non-duplicated markers, approximately 23.8 million SNP interactions would have been in the prediction model for this study. Construction of a selection model using that much

information could quickly become unwieldy, particularly for studies with even more non-duplicated SNPs. Choosing SNPs based on haplotype could be another possible solution; thus limiting the total number of interactions while still providing useful information on significant interactions.

The fatty acids, phenotyped by GC, displayed the highest heritability values, ranging from 0.92-0.94 (Table 2.2). Consequently, these traits also proved to be easily selected for with all methods except Epistacy. Spearman correlations with 2013 phenotypes for PS, BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi) were never below 0.67 (stearic PS), while rising as high as 0.87 (oleic BayesB) (Tables 2.5-2.9). A few notable trends emerged when comparing the selection methods with one another. First, in all direct comparisons, PS was outperformed ($P < 0.05$) or was not different from ($P > 0.05$) BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi) (Tables 2.4-2.9, 2.13-2.17). Second, with few exceptions (Table 2.13 high tail), BayesB and G-BLUP outperformed ($P < 0.05$) or did not differ from ($P > 0.05$) BayesB(Epi) and G-BLUP(Epi) (Tables 2.4-2.9, 2.13-2.17). Finally, in the 15 direct comparisons between BayesB and G-BLUP, there were only three differences ($P < 0.05$) (Tables 2.4-2.9, 2.13-2.17). For palmitic (Table 2.5) and stearic (Table 2.6), BayesB displayed a higher correlation ($P < 0.05$) with the 2013 observed phenotypes. However, in the high tail selections for linoleic (Table 2.16), G-BLUP produced a greater prediction ($P < 0.05$) than BayesB. Thus, BayesB is the overall best method for predicting fatty acids with only a slight edge over G-BLUP, corresponding to the findings of Clark et al. (2011).

Protein and oil, phenotyped with NIRS, displayed higher heritability values than yield but lower than the fatty acids (Table 2.2). Similar to the fatty acids but in contrast with yield, Epistacy was the least effective selection method for protein and oil (Tables 2.10-2.11, 2.18-2.19), indicating that epistatic interactions are likely less influential for protein and oil than for yield. For the Spearman

correlations with 2013 phenotypes for protein and oil, the only significant difference ($P < 0.05$) that occurred was between BayesB (0.50) and G-BLUP (0.50) for protein (Tables 2.10-2.11). In the high tail selections for protein, BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi) all outperformed ($P < 0.05$) PS while not differing from each other ($P > 0.05$) (Table 2.18). In the contrasts for oil (Table 2.19), BayesB(Epi) was the only method that either outperformed ($P < 0.05$) or did not differ from ($P > 0.05$) all other methods (excluding Epistacy). Therefore, the preferred methods for protein and oil are G-BLUP and BayesB(Epi), respectively.

Conclusions

Breeding method evaluation is an important strategy in maximizing gain from selection. Given the importance of yield, fatty acids, protein, and oil in soybean production, it is necessary to determine the most useful methods for trait improvement. Additionally, it is also of interest to evaluate different selection methods from the progeny row stage, as this is a critical step in the breeding pipeline. After evaluating both phenotypic (PS) and molecular (BayesB, G-BLUP, Epistacy, BayesB(Epi), and G-BLUP(Epi)) breeding methods for yield, fatty acids, protein and oil, it was determined that there was no consensus method for maximum improvement for all traits. Instead, the preferred method differed based on the trait evaluated. Yield was perhaps the biggest outlier in terms of selection method recommendation. For yield, Epistacy was the preferred method; however for all other traits Epistacy was the least influential method. For the fatty acids, BayesB and G-BLUP were the best methods, with the slight overall edge going to BayesB. For the NIRS traits, the preferred method for protein was G-BLUP and for the preferred method for oil was BayesB(Epi). Notably, for each trait the preferred method was a molecular selection strategy. This provides important implications for how soybean breeders could maximize selections from the progeny row stage. Further research evaluating these methods in a wide range of pedigrees or with elite

germplasm would be useful in order to refine selection recommendations for breeders.

References

Allen, F.L., R. Johnson, R.C. Williams Jr., A.T. McClure, M. Newman, and P. Donald. 2011. Soybean variety performance tests in Tennessee. http://varietytrials.tennessee.edu/ (accessed 10 Feb. 2014).

Allen, F.L., R. Johnson, R.C. Williams Jr., A.T. McClure, M. Newman, H. Young-Kelly, and P. Donald. 2012. Soybean variety performance tests in Tennessee. http://varietytrials.tennessee.edu/ (accessed 10 Feb. 2014).

Allen, F.L., V.R. Sykes, R.C. Williams Jr., A.T. McClure, H. Young-Kelly, and P. Donald. 2013. Soybean variety performance tests in Tennessee. http://varietytrials.tennessee.edu/ (accessed 10 Feb. 2014).

Bernard, R.L., and C.R. Cremeens. 1988. Registration of 'Williams 82' soybean. Crop Sci. 28:1027–1028.

Bilyeu, K., J.D. Gillman, and A.R. LeRoy. 2011. Novel FAD3 mutant allele combinations produce soybeans containing 1% linolenic acid in the seed oil. Crop Sci. 51:259–264.

Boersma, J.G., J.D. Gillman, K.D. Bilyeu, G.R. Ablett, C. Grainger, and I. Rajcan. 2012. New mutations in a delta-9-stearoyl-acyl carrier protein desaturase gene associated with enhanced stearic acid levels in soybean seed. Crop Sci. 52:1736–1742.

Bolon, Y., W.J. Haun, W.W. Xu, D. Grant, M.G. Stacey, R.T. Nelson, D.J. Gerhardt, J.A. Jeddeloh, G. Stacey, G.J. Muehlbauer, J.H. Orf, S.L. Naeve, R.M. Stupar, and C.P. Vance. 2011. Phenotypic and genomic analyses of fast neutron mutant population resource in soybean. Plant Physiol. 156:240-253.

Brim, C.A. 1966. A modified pedigree method of selection in soybeans. Crop Sci. 6:220.

Broman, K.W., H. Wu, Ś. Sen, and G.A. Churchill. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890.

Browning, B.L., and S.R. Browning. 2009. A unified approach to genotype imputation and haplotype phase inference for large data sets of trios and unrelated individuals. Am. J. Hum. Genet. 84:210-223.

Browning, S.R., and B.L. Browning. 2007. Rapid and accurate haplotype phasing and missing data inference for whole genome association studies using localized haplotype clustering. Am. J. Hum. Genet. 81:1084-1097.

Clark, S.A., J.M. Hickey, and J.H.J. van der Werf. 2011. Different models of genetic variation and their effect on genomic evaluation. Genet. Sel. Evol. 43:18

Cober, E.R., S.R. Cianzio, V.R. Pantalone, and I. Rajcan. 2009. Soybean. In: J. Vollman and I. Rajcan, editors, Oil crops: Handbook of plant breeding, volume 4. Springer Science + Business Media LLC. p. 57-90.

Crossa, J., P. Pérez, J. Hickey, J. Burgeño, L. Ornella, J. Cerón-Rojas, X. Zhang, S. Dreisigacker, R. Babu, Y. Li, D. Bonnett, and K. Mathews. 2014. Genomic prediction in CYMMIT maize and wheat breeding programs. Heredity 112:48-60.

De los Campos, G., J.M. Hickey, R. Pong-Wong, H.D. Daetwyler, and M.P.L.Calus. 2013. Whole-genome regression and prediction methods applied to plant and animal breeding. Genetics 193:327-345.

Diedenhofen, B., and J. Musch. 2015. Cocor: a comprehensive solution for the statistical comparison of correlations. PLoS ONE 10:e0121945.

Diers, B.W., T.R. Cary, D.J. Thomas, A. Colgrove, and T. Niblack. 2010. Registration of ‘LD00-2817P’ germplasm line with resistance to soybean cyst nematode from PI 437654. J. Plant Regist. 4:141-144.

Diers, B.W., T.R. Cary, D.J. Thomas, and C.D. Nickell. 2006. Registration of ‘LD00-3309’ soybean. Crop Sci. 46:1384.

Fallen, B.D., K. Rainey, C.E. Sams, D.A.Kopsell, and V.R. Pantalone. 2012. Evaluation of agronomic and seed characteristics in elevated oleic acid soybean lines in the south-eastern US. J. Am. Oil Chem. Soc. 89:1333-1343.

Federal Register. 2015. Final determination regarding partially hydrogenated oils.

https://www.federalregister.gov/articles/2015/06/17/2015-14883/final-determination-regarding-partially-hydrogenated-oils (accessed 24 July 2015).

Fehr W.R., and C.E. Caviness. 1977. Stages of soybean development. Special Report, Agriculture and Home Economics Experiment Station, Iowa State University, 1977, issue 80, p 11.

Gillman, J.D., M.G. Stacy, Y. Cui, H.R. Berg, and G. Stacey. 2014. Deletions of the SACPD-C locus elevate seed stearic acid but also result in fatty acid and morphological alterations in nitrogen fixing nodules. BMC Plant Biol. 14:143.

Hittner, J. B., K. May, and N.C. Silver. 2003. A Monte Carlo evaluation of tests for comparing dependent correlations. J. Gen. Phychol. 130:149-168.

Holland, J.B. 1998. EPISTACY: A SAS program for detecting two-locus epistatic interactions using genetic marker information. J. Hered. 89:374–375.

Jonas, E., and D.J. de Koning. 2013. Does genomic selection have a future in plant breeding? Trends Biotechnol. 31:497-504.

Kinney, A.J. 1996. Development of genetically engineered soybean oils for food application. J. Food Lipids. 3:273-292.

Kinney, A.J., and T.E. Clemente. 2005. Modifying soybean oil for enhanced performance in biodiesel blends. Fuel Pro. Technol. 86:1137–1147.

Kris-Etherton, P.M., and S. Yu. 1997. Individual fatty acid effects on plasma lipids and lipoproteins: Human studies. Am. J. Clin. Nutr. 65:S1628–S1644.

Lee, J.D., K.D. Bilyeu, V.R. Pantalone, A.M. Gillen, Y.S. So, and J.G. Shannon. 2012. Environmental stability of oleic acid concentration in seed oil for soybean lines with FAD2-1A and FAD2-1B mutant genes. Crop Sci. 52:1290–1297.

Lillehammer, M., T.H.E. Meuwissen, and A.K. Sonesson. 2011. A comparison of dairy cattle breeding designs that use genomic selection. J. Dairy Sci. 94:493-500.

Meuwissen, T. 2007. Genomic selection: Marker assisted selectionon a genome wide scale. J. Anim. Breed. Genet. 124:321–322.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.

Morrison, M.J., E.K. Cober, M.F. Saleem, N.B. McLaughlin, J. Fregeau-Reid, B.L. Ma, W. Yan, and L. Woodrow. 2008. Changes in isoflavone concentration with 58 years of genetic improvement of short-season soybean cultivars in Canada. Crop Sci. 48:2201–2208.

Nakaya, A., and S.N. Isobe. 2012. Will genomic selection be a practical method for plant breeding? Ann. Bot. 110:1303-1316.

Nyquist, W.E. 1991. Estimation of heritability and prediction of selection response in plant populations. Crit. Rev. Plant Sci.10:235–322.

Ødegård, J., A.K. Sonesson, M.H. yazdi, and T.H.E. Meuwissen. 2009. Introgression of a major QTL from an inferior into a superior population using genomic selection. Genet. Sel. Evol. 41:38.

Pantalone, V.R., F.L. Allen, and D. Landau-Ellis. 2003. Registration of ‘5601T’ soybean. Crop Sci. 43:1123-1124.

Pantalone, V.R., F.L. Allen, and D. Landau-Ellis. 2004. Registration of ‘5002T’ soybean. Crop Sci. 44:1483-1484.

Pantalone, V.R., R.F. Wilson, W.P. Novitzky, and J.W. Burton. 2002. Genetic regulation of elevated stearic acid concentration in soybean oil. J. Am. Oil Chem. Soc. 79:543–553.

Panthee, D.R., V.R. Pantalone, C.E. Sams, A.M. Saxton, D.R. West, J.H. Orf, and A.S. Killam. 2006. Quantitative trait loci controlling sulfur containing amino acids, methionine and cysteine, in soybean seeds. Theor. Appl. Genet. 112:546–553.

Pérez, P., and G. de los Campos. 2014. Genome-wide regression and prediction with the BGLR statistical package. Genetics 198:483-495.

Pham, A.T., J.D. Lee, J.G. Shannon, and K.D. Bilyeu. 2010. Mutant alleles of

FAD2-1A and FAD2-1B combine to produce soybeans with the high oleic acid seed oil trait. BMC Plant Biol. 10:195.

Poland, J., J. Endelman, J. Dawson, J. Rutkoski, S. Wu, Y. Manes, S. Dreisigacker, J. Crossa, H. Sanchez-Villeda, M. Sorrells, and J.L. Jannink. 2012. Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Gen. 5:103–113.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Resende Jr., M.F.R., P. Muñoz, J.J. Acousta, G.F. Peter, J.M. Davis, D. Grattapaglia, M.D.V Resende, and M. Kirst. 2012. Accelerating the domestication of trees using genomic selection: Accuracy of prediction models across ages and environments. New Phytol. 193:617-624.

SAS Institute Inc. 2002-2012. Cary, NC, USA. SAS 9.4)

Schmutz, J., S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D.L. Hyten, Q. Song, J.J. Thelen, J. Cheng, et al. 2010. Genome sequence of the palaeopolyploid soybean. Nature 463:178–183.

Sitzenstock, F., F. Ytournel, A.R. Sharifi, D. Cavero, H. Täubert, R. Preisinger, and H. Simianer. 2013. Effeciency of genomic selection in an established commercial layer breeding program. Gent. Sel. Evol. 45:29.

Smith, T.J., and H.M. Camper. 1973. Registration of Essex Soybean (Reg. No. 97). Crop Sci. 13:495.

Song, Q., D.L. Hyten, G. Jia, C.V. Quigley, E.W. Fickus, R.L. Nelson, and P.B. Cregan. 2013. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS ONE 8:e54985.

Spencer, M.M., D. Landau-Ellis, E.J. Meyer, and V.R. Pantalone. 2004. Molecular markers associated with linolenic acid content in soybean. J. Am. Oil Chem. Soc. 81:559–562.

USDA, ARS, National Genetic Resources Program. Germplasm Resources Information Network - (GRIN) [Online Database]. National Germplasm

Resources Laboratory, Beltsville, Maryland. URL: http://www.ars-grin.gov.4/cgi-bin/npgs/html/index.pl?language=en (24 July 2015)

Wiggins, B.T. 2012. Heritability and genetic gain of seed protein, oil, and yield among RIL of soybean. M.S. thesis. Univ. of Tennessee, Knoxville, TN, USA.

Wilson, R.F. 2004. Seed composition. In: H.R. Boerma and J.E. Specht, editors, Soybeans: Improvement, production, and uses. 3rd ed. ASA, CSSA, and SSSA, Madison, WI. p. 621–678.

Wimmer, V., T. Albrecht, H.J. Auinger, and C.C. Schön. 2012. Synbreed: A framework for the analysis of genomic prediction using R. Bioinformatics. 28:2086-2087.

Yaklich, R.W., B. Vinyard, M. Camp, and S. Douglass. 2002. Analysis of seed protein and oil from soybean northern and southern region uniform tests. Crop Sci. 42:1504–1515.

Appendix B-Chapter 2 Tables and Figures

Tables

Table 2.1 Simple statistics for soybean population E×W-50K consisting of 860 F5 derived RILs planted in single rep plots in 2010 in Knoxville, TN. This dataset was used to make performance predictions for traits of interest in a subset of the population (276 RILs) grown in replicated field trials in 2013 at three locations (Knoxville, TN; Springfield, TN; and Milan, TN).

Trait	Essex	Williams 82	min	mean	max	std. dev.†
Maturity (Julian)	278.00	262.00	251.00	270.19	288.00	7.34
Height (cm)	53.34	60.96	25.40	78.51	132.08	19.99
Yield (kg ha^{-1})	2548.78	1566.93	685.95	2137.53	3591.15	528.26
Palmitic (g kg^{-1} seed oil)	107.20	100.30	90.50	106.64	165.00	9.28
Stearic (g kg^{-1} seed oil)	48.50	44.00	32.50	42.36	79.90	4.95
Oleic (g kg^{-1} seed oil)	233.50	237.30	158.40	242.50	353.00	27.69
Linoleic (g kg^{-1} seed oil)	534.40	551.30	436.40	535.57	601.10	22.03
Linolenic (g kg^{-1} seed oil)	76.40	67.10	53.90	72.94	116.60	7.33
Protein (g kg^{-1} seed dry weight)	430.46	417.01	366.32	412.91	459.54	16.18
Oil (g kg^{-1} seed dry weight)	217.82	232.76	200.46	225.98	247.36	7.37

† std. deviation of LSMEANs

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.2 Simple statistics for soybean population E×W-50K subset consisting of 276 F5 derived RILs planted in replicated field trials at three locations in 2013 (Knoxville, TN; Springfield, TN; and Milan, TN). Information from this dataset was compared with performance predictions for traits of interest in the full population (860 RILs) grown in 2010 in single rep plots planted at Knoxville, TN.

Trait	Genotype P value	GxE P value	Essex	Williams 82	min	mean	max	std. dev.†	LSD value	h^{2}‡
Maturity (Julian)	***	***	272.22	262.56	259.42	270.42	276.89	2.79	3.53	0.79
Height (cm)	***	***	75.64	93.98	37.82	89.08	133.49	18.81	11.71	0.95
Yield (kg ha^{-1})	***	***	3588.91	3002.94	1371.60	3222.92	4087.60	395.33	663.42	0.63
Palmitic (g kg^{-1} seed oil)	***	NS	125.68	110.57	99.58	121.19	138.73	7.56	5.67	0.93
Stearic (g kg^{-1} seed oil)	***	***	42.12	38.68	34.21	40.68	50.95	3.08	2.19	0.94
Oleic (g kg^{-1} seed oil)	***	***	350.93	408.64	279.10	360.40	508.58	44.06	29.75	0.94
Linoleic (g kg^{-1} seed oil)	***	***	404.83	377.41	299.75	404.42	455.26	30.83	22.60	0.93
Linolenic (g kg^{-1} seed oil)	***	**	76.44	64.83	53.52	73.32	93.40	8.14	6.53	0.92
Protein (g kg^{-1} seed dry weight)	***	***	423.63	421.61	376.27	410.47	444.02	11.99	12.16	0.87
Oil (g kg^{-1} seed dry weight)	***	***	211.96	227.36	200.54	218.33	238.35	5.92	6.03	0.87

**Significant at 0.01 probability level
***Significant at 0.001 probability level
NS, Not significant at 0.05 probability level
E×W-50K = soybean population with parental lines Essex and Williams 82
† std. deviation of LSMEANs
‡ heritability calculated using entry means basis (Nyquist, 1991)

Table 2.3 Comparison of cross-validations for G-BLUP and BayesB methods of GS for soybean population E×W-50K consisting of 860 F5 derived RILs grown in 2010 at Knoxville, TN. Cross-validations were replicated 50 times for each trait. In each rep, a randomly chosen 1/5 of the population had phenotypic data removed (test set), while phenotypic and genotypic information were retained for the remaining 4/5 of the population (training set). The values displayed for G-BLUP and BayesB are the mean Pearson correlation coefficients for the predicted and observed values in the test set. For each trait there was no statistical difference (P > 0.05) between G-BLUP and BayesB methods.

	G-BLUP	BayesB	P value
Yield	0.43	0.44	NS
Palmitic	0.50	0.48	NS
Stearic	0.55	0.55	NS
Oleic	0.55	0.53	NS
Linoleic	0.53	0.52	NS
Linolenic	0.42	0.41	NS
Protein	0.55	0.55	NS
Oil	0.48	0.49	NS

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.4 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for soybean yield in population E×W-50K subset consisting of 276 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), and E (2010 Epistacy). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Yield	Correlation Differences	**13:P** 0.18	**13:B** 0.13	**13:G** 0.15	**13:E** 0.23
13:P	0.18	\	0.05	0.03	-0.05
13:B	0.13	NS	\	-0.01	-0.10
13:G	0.15	NS	***	\	-0.08
13:E	0.23	NS	NS	NS	\

Spearman Correlations	**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**
2013 Pheno	\	0.18	0.13	0.15	0.23
PS	**	\	0.39	0.41	0.30
BayesB	*	***	\	1.00	0.29
G-BLUP	*	***	***	\	0.29
Epistacy	***	***	***	***	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.5 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for palmitic acid in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Palmitic	Correlation Differences	**13:P** 0.69	**13:B** 0.80	**13:G** 0.79	**13:E** 0.51	**13:B(E)** 0.78	**13:G(E)** 0.78
13:P	0.69	\	-0.11	-0.11	0.18	-0.09	-0.09
13:B	0.80	***	\	0.00	0.28	0.02	0.02
13:G	0.79	***	**	\	0.28	0.02	0.02
13:E	0.51	***	***	***	\	-0.27	-0.26
13:B(E)	0.78	***	*	NS	***	\	0.00
13:G(E)	0.78	**	**	*	***	**	\

Spearman Correlations	**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2013 Pheno	\	0.69	0.80	0.79	0.51	0.78	0.78
PS	***	\	0.73	0.72	0.44	0.73	0.73
BayesB	***	***	\	1.00	0.67	0.98	0.98
G-BLUP	***	***	***	\	0.68	0.98	0.98
Epistacy	***	***	***	***	\	0.66	0.66
BayesB(Epi)	***	***	***	***	***	\	1.00
G-BLUP(Epi)	***	***	***	***	***	***	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.6 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for stearic acid in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Stearic	Correlation Differences	**13:P** 0.67	**13:B** 0.80	**13:G** 0.76	**13:E** 0.69	**13:B(E)** 0.79	**13:G(E)** 0.76
13:P	0.67	\	-0.13	-0.10	-0.02	-0.12	-0.09
13:B	0.80	***	\	0.03	0.11	0.01	0.04
13:G	0.76	***	***	\	0.07	-0.02	0.00
13:E	0.69	NS	***	**	\	-0.10	-0.07
13:B(E)	0.79	***	NS	**	***	\	0.03
13:G(E)	0.76	***	***	NS	**	***	\

Spearman Correlations	**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2013 Pheno	\	0.67	0.80	0.76	0.69	0.79	0.76
PS	***	\	0.81	0.82	0.70	0.83	0.84
BayesB	***	***	\	0.99	0.78	0.98	0.97
G-BLUP	***	***	***	\	0.77	0.97	0.98
Epistacy	***	***	***	***	\	0.77	0.76
BayesB(Epi)	***	***	***	***	***	\	0.99
G-BLUP(Epi)	***	***	***	***	***	***	\

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.7 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for oleic acid in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Oleic	Correlation Differences	**13:P** 0.78	**13:B** 0.87	**13:G** 0.87	**13:E** 0.28	**13:B(E)** 0.86	**13:G(E)** 0.86
13:P	0.78	\	-0.09	-0.09	0.49	-0.08	-0.08
13:B	0.87	***	\	0.00	0.59	0.01	0.01
13:G	0.87	***	NS	\	0.58	0.01	0.01
13:E	0.28	***	***	***	\	-0.58	-0.58
13:B(E)	0.86	***	NS	NS	***	\	0.00
13:G(E)	0.86	***	*	NS	***	NS	\

Spearman Correlations	**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2013 Pheno	\	0.78	0.87	0.87	0.28	0.86	0.86
PS	***	\	0.82	0.81	0.34	0.82	0.82
BayesB	***	***	\	1.00	0.38	0.99	0.99
G-BLUP	***	***	***	\	0.38	0.99	0.99
Epistacy	***	***	***	***	\	0.37	0.37
BayesB(Epi)	***	***	***	***	***	\	1.00
G-BLUP(Epi)	***	***	***	***	***	***	\

*Significant at 0.05 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.8 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for linoleic acid in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Linoleic	Correlation Differences	**13:P** 0.75	**13:B** 0.84	**13:G** 0.84	**13:E** 0.41	**13:B(E)** 0.83	**13:G(E)** 0.83
13:P	0.75	\	-0.10	-0.10	0.33	-0.09	-0.09
13:B	0.84	***	\	0.00	0.43	0.01	0.01
13:G	0.84	***	NS	\	0.43	0.01	0.01
13:E	0.41	***	***	***	\	-0.42	-0.42
13:B(E)	0.83	***	NS	NS	***	\	0.00
13:G(E)	0.83	***	NS	NS	***	**	\

Spearman Correlations	**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2013 Pheno	\	0.75	0.84	0.84	0.41	0.83	0.83
PS	***	\	0.81	0.81	0.45	0.82	0.83
BayesB	***	***	\	1.00	0.50	0.98	0.98
G-BLUP	***	***	***	\	0.50	0.98	0.98
Epistacy	***	***	***	***	\	0.50	0.50
BayesB(Epi)	***	***	***	***	***	\	1.00
G-BLUP(Epi)	***	***	***	***	***	***	\

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.9 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for linolenic acid in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Linolenic	Correlation Differences	**13:P** 0.67	**13:B** 0.80	**13:G** 0.80	**13:E** 0.39	**13:B(E)** 0.79	**13:G(E)** 0.79
13:P	0.67	\	-0.14	-0.14	0.28	-0.12	-0.12
13:B	0.80	***	\	0.00	0.42	0.01	0.01
13:G	0.80	***	NS	\	0.42	0.01	0.01
13:E	0.39	***	***	***	\	-0.41	-0.41
13:B(E)	0.79	***	*	*	***	\	0.00
13:G(E)	0.79	***	NS	NS	***	NS	\

Spearman Correlations	**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2013 Pheno	\	0.67	0.80	0.80	0.39	0.79	0.79
PS	***	\	0.74	0.73	0.54	0.74	0.73
BayesB	***	***	\	1.00	0.56	0.99	0.99
G-BLUP	***	***	***	\	0.56	0.99	0.99
Epistacy	***	***	***	***	\	0.55	0.55
BayesB(Epi)	***	***	***	***	***	\	1.00
G-BLUP(Epi)	***	***	***	***	***	***	\

*Significant at 0.05 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.10 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for protein in soybean population E×W-50K subset consisting of 271 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Protein	Correlation Differences	**13:P** 0.48	**13:B** 0.50	**13:G** 0.50	**13:E** 0.31	**13:B(E)** 0.50	**13:G(E)** 0.50
13:P	0.48	\	-0.02	-0.02	0.17	-0.02	-0.02
13:B	0.50	NS	\	0.00	0.19	0.00	-0.01
13:G	0.50	NS	*	\	0.19	0.00	-0.01
13:E	0.31	***	***	***	\	-0.19	-0.20
13:B(E)	0.50	NS	NS	NS	***	\	-0.01
13:G(E)	0.50	NS	NS	NS	***	NS	\

Spearman Correlations	**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2013 Pheno	\	0.48	0.50	0.50	0.31	0.50	0.50
PS	***	\	0.67	0.68	0.60	0.67	0.68
BayesB	***	***	\	1.00	0.51	0.99	0.99
G-BLUP	***	***	***	\	0.52	0.99	0.99
Epistacy	***	***	***	***	\	0.51	0.51
BayesB(Epi)	***	***	***	***	***	\	1.00
G-BLUP(Epi)	***	***	***	***	***	***	\

*Significant at 0.05 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.11 Comparison of Spearman correlations between 2013 phenotypes and 2010 predictions for oil in soybean population E×W-50K subset consisting of 271 F5 derived RILs. The estimates listed are as follows: 13 (2013 phenotype), P (2010 PS), B (2010 BayesB), G (2010 G-BLUP), E (2010 Epistacy), B(E) (2010 BayesB(Epi)), and G(E) (2010 G-BLUP(Epi)). Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Oil	Correlation Differences	**13:P** 0.59	**13:B** 0.56	**13:G** 0.56	**13:E** 0.39	**13:B(E)** 0.57	**13:G(E)** 0.57
13:P	0.59	\	0.03	0.04	0.20	0.02	0.02
13:B	0.56	NS	\	0.00	0.17	-0.01	-0.01
13:G	0.56	NS	NS	\	0.16	-0.02	-0.01
13:E	0.39	***	***	***	\	-0.18	-0.18
13:B(E)	0.57	NS	NS	NS	***	\	0.00
13:G(E)	0.57	NS	NS	NS	***	NS	\

Spearman Correlations	**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2013 Pheno	\	0.59	0.56	0.56	0.39	0.57	0.57
PS	***	\	0.76	0.75	0.58	0.78	0.77
BayesB	***	***	\	1.00	0.71	0.98	0.98
G-BLUP	***	***	***	\	0.72	0.98	0.98
Epistacy	***	***	***	***	\	0.72	0.72
BayesB(Epi)	***	***	***	***	***	\	1.00
G-BLUP(Epi)	***	***	***	***	***	***	\

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.12 Yield contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 276 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, and G-BLUP) selection methods.

15% High		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**
Yield kg ha^{-1}		3761.98	3284.64	3250.55	3264.97	3367.12
2013 Pheno	3761.98	\	477.34	511.43	497.01	394.86
PS	3284.64	***	\	34.09	19.67	-82.48
BayesB	3250.55	***	NS	\	-14.42	-116.57
G-BLUP	3264.97	***	NS	NS	\	-102.15
Epistacy	3367.12	***	NS	*	*	\
15% Low		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**
Yield kg ha^{-1}		2565.89	3104.31	3131.86	3130.75	2996.45
2013 Pheno	2565.89	\	-538.42	-565.97	-564.86	-430.56
PS	3104.31	***	\	-27.55	-26.44	107.86
BayesB	3131.86	***	NS	\	1.11	135.41
G-BLUP	3130.75	***	NS	NS	\	134.30
Epistacy	2996.45	***	*	***	***	\

*Significant at 0.05 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.13 Palmitic acid contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Palmitic g kg^{-1} seed oil		131.96	129.06	129.45	129.51	126.17	129.98	129.86
2013 Pheno	131.96	\	2.90	2.51	2.45	5.79	1.98	2.10
PS	129.06	***	\	-0.39	-0.45	2.89	-0.92	-0.80
BayesB	129.45	***	NS	\	-0.06	3.28	-0.53	-0.41
G-BLUP	129.51	***	NS	NS	\	3.34	-0.47	-0.35
Epistacy	126.17	***	***	***	***	\	-3.81	-3.69
BayesB(Epi)	129.98	***	**	**	**	***	\	0.12
G-BLUP(Epi)	129.86	***	**	*	*	***	NS	\
15% low		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Palmitic g kg^{-1} seed oil		108.78	112.98	111.46	111.46	114.75	112.25	112.25
2013 Pheno	108.78	\	-4.20	-2.68	-2.68	-5.97	-3.47	-3.47
PS	112.98	***	\	1.52	1.52	-1.77	0.73	0.73
BayesB	111.46	***	***	\	0.00	-3.29	-0.79	-0.79
G-BLUP	111.46	***	***	NS	\	-3.29	-0.79	-0.79
Epistacy	114.75	***	***	***	***	\	2.50	2.50
BayesB(Epi)	112.25	***	*	***	***	***	\	0.00
G-BLUP(Epi)	112.25	***	*	***	***	***	NS	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.14 Stearic acid contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Stearic g kg^{-1} seed oil		45.72	43.61	44.06	43.99	43.59	43.76	43.64
2013 Pheno	45.72	\	2.10	1.65	1.72	2.12	1.95	2.08
PS	43.61	***	\	-0.45	-0.38	0.02	-0.15	-0.02
BayesB	44.06	***	***	\	0.07	0.47	0.30	0.43
G-BLUP	43.99	***	***	NS	\	0.40	0.23	0.36
Epistacy	43.59	***	NS	***	***	\	-0.17	-0.05
BayesB(Epi)	43.76	***	NS	***	***	NS	\	0.13
G-BLUP(Epi)	43.64	***	NS	***	***	NS	**	\
15% low		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Stearic g kg^{-1} seed oil		36.50	37.74	37.67	37.62	37.75	37.57	37.65
2013 Pheno	36.50	\	-1.24	-1.17	-1.12	-1.25	-1.07	-1.15
PS	37.74	***	\	0.07	0.12	-0.01	0.17	0.09
BayesB	37.67	***	NS	\	0.05	-0.08	0.10	0.02
G-BLUP	37.62	***	NS	NS	\	-0.13	0.05	-0.03
Epistacy	37.75	***	NS	NS	NS	\	0.18	0.10
BayesB(Epi)	37.57	***	NS	NS	NS	NS	\	-0.08
G-BLUP(Epi)	37.65	***	NS	NS	NS	NS	*	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.15 Oleic acid contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Oleic g kg^{-1} seed oil	435.23	435.23	413.85	423.55	423.55	379.74	424.16	424.16
2013 Pheno	435.23	\	21.38	11.68	11.68	55.49	11.07	11.07
PS	413.85	***	\	-9.70	-9.70	34.11	-10.31	-10.31
BayesB	423.55	***	***	\	0.00	43.81	-0.61	-0.61
G-BLUP	423.55	***	***	NS	\	43.81	-0.61	-0.61
Epistacy	379.74	***	***	***	***	\	-44.42	-44.42
BayesB(Epi)	424.16	***	***	NS	NS	***	\	0.00
G-BLUP(Epi)	424.16	***	***	NS	NS	***	NS	\
15% low		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Oleic g kg^{-1} seed oil	305.25	305.25	321.43	315.51	316.43	341.96	317.00	317.29
2013 Pheno	305.25	\	-16.18	-10.26	-11.18	-36.71	-11.75	-12.04
PS	321.43	***	\	5.92	5.00	-20.53	4.43	4.14
BayesB	315.51	***	***	\	-0.92	-26.45	-1.49	-1.78
G-BLUP	316.43	***	**	NS	\	-25.53	-0.57	-0.86
Epistacy	341.96	***	***	***	***	\	24.96	24.67
BayesB(Epi)	317.00	***	**	*	NS	***	\	-0.29
G-BLUP(Epi)	317.29	***	*	*	NS	***	NS	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.16 Linoleic acid contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Linoleic g kg^{-1} seed oil		442.80	432.27	432.73	433.62	421.61	432.84	432.74
2013 Pheno	442.80	\	10.53	10.07	9.18	21.19	9.96	10.06
PS	432.27	***	\	-0.46	-1.35	10.66	-0.57	-0.47
BayesB	432.73	***	NS	\	-0.89	11.12	-0.11	-0.01
G-BLUP	433.62	***	NS	*	\	12.01	0.78	0.88
Epistacy	421.61	***	***	***	***	\	-11.23	-11.13
BayesB(Epi)	432.84	***	NS	NS	**	***	\	0.10
G-BLUP(Epi)	432.74	***	NS	NS	*	***	NS	\
15% low		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Linoleic g kg^{-1} seed oil		351.68	369.46	359.95	360.04	382.84	359.50	359.19
2013 Pheno	351.68	\	-17.78	-8.27	-8.36	-31.16	-7.82	-7.51
PS	369.46	***	\	9.51	9.42	-13.38	9.96	10.27
BayesB	359.95	***	***	\	-0.09	-22.89	0.45	0.76
G-BLUP	360.04	***	***	NS	\	-22.80	0.54	0.85
Epistacy	382.84	***	***	***	***	\	23.34	23.65
BayesB(Epi)	359.50	***	***	NS	NS	***	\	0.31
G-BLUP(Epi)	359.19	***	***	NS	NS	***	NS	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.17 Linolenic acid contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Linolenic g kg^{-1} seed oil		85.30	81.50	82.94	82.94	79.46	82.82	82.82
2013 Pheno	85.30	\	3.80	2.36	2.36	5.84	2.48	2.48
PS	81.50	***	\	-1.44	-1.44	2.04	-1.33	-1.33
BayesB	82.94	***	***	\	0.00	3.48	0.12	0.12
G-BLUP	82.94	***	***	NS	\	3.48	0.12	0.12
Epistacy	79.46	***	***	***	***	\	-3.36	-3.36
BayesB(Epi)	82.82	***	***	NS	NS	***	\	0.00
G-BLUP(Epi)	82.82	***	***	NS	NS	***	NS	\
15% low		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Linolenic g kg^{-1} seed oil		60.47	63.29	62.74	62.74	70.19	63.23	63.23
2013 Pheno	60.47	\	-2.82	-2.27	-2.27	-9.72	-2.76	-2.76
PS	63.29	***	\	0.55	0.55	-6.90	0.06	0.06
BayesB	62.74	***	NS	\	0.00	-7.46	-0.49	-0.49
G-BLUP	62.74	***	NS	NS	\	-7.46	-0.49	-0.49
Epistacy	70.19	***	***	***	***	\	6.96	6.96
BayesB(Epi)	63.23	***	NS	***	***	***	\	0.00
G-BLUP(Epi)	63.23	***	NS	***	***	***	NS	\

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.18 Protein contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 271 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Protein g kg^{-1} seed dw		429.29	418.66	421.95	421.95	416.06	421.81	421.85
2013 Pheno	429.29	\	10.63	7.34	7.34	13.23	7.48	7.44
PS	418.66	***	\	-3.29	-3.29	2.60	-3.15	-3.19
BayesB	421.95	***	***	\	0.00	5.89	0.14	0.10
G-BLUP	421.95	***	***	NS	\	5.89	0.14	0.10
Epistacy	416.06	***	***	***	***	\	-5.75	-5.79
BayesB(Epi)	421.81	***	***	NS	NS	***	\	-0.04
G-BLUP(Epi)	421.85	***	***	NS	NS	***	NS	\
15% low		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Protein g kg^{-1} seed dw		392.29	399.89	399.19	399.19	405.79	399.50	399.25
2013 Pheno	392.29	\	-7.60	-6.90	-6.90	-13.50	-7.21	-6.96
PS	399.89	***	\	0.70	0.70	-5.90	0.39	0.64
BayesB	399.19	***	NS	\	0.00	-6.60	-0.31	-0.06
G-BLUP	399.19	***	NS	NS	\	-6.60	-0.31	-0.06
Epistacy	405.79	***	***	***	***	\	6.29	6.54
BayesB(Epi)	399.50	***	NS	NS	NS	***	\	0.25
G-BLUP(Epi)	399.25	***	NS	NS	NS	***	NS	\

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 2.19 Oil contrasts of 15% (41 RILs) high and low tail selections from 2010 predictions as measured in 2013 from soybean population E×W-50K subset consisting of 271 F5 derived RILs. The 2013 high and low tail selections are included for comparison. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Oil g kg^{-1} seed dw		227.58	223.26	223.47	223.29	222.65	223.64	223.15
2013 Pheno	227.58	\	4.32	4.11	4.29	4.93	3.94	4.43
PS	223.26	***	\	-0.21	-0.03	0.61	-0.38	0.11
BayesB	223.47	***	NS	\	0.18	0.82	-0.17	0.32
G-BLUP	223.29	***	NS	*	\	0.64	-0.35	0.14
Epistacy	222.65	***	NS	*	NS	\	-0.99	-0.50
BayesB(Epi)	223.64	***	NS	NS	*	**	\	0.49
G-BLUP(Epi)	223.15	***	NS	**	NS	NS	***	\
15% low		**2013 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Oil g kg^{-1} seed dw		209.49	212.76	212.77	212.77	214.87	212.60	212.47
2013 Pheno	209.49	\	-3.27	-3.28	-3.28	-5.38	-3.11	-2.98
PS	212.76	***	\	-0.01	-0.01	-2.11	0.16	0.29
BayesB	212.77	***	NS	\	0.00	-2.10	0.17	0.30
G-BLUP	212.77	***	NS	NS	\	-2.10	0.17	0.30
Epistacy	214.87	***	***	***	***	\	2.27	2.40
BayesB(Epi)	212.60	***	NS	NS	NS	***	\	0.13
G-BLUP(Epi)	212.47	***	NS	*	*	***	NS	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Figures

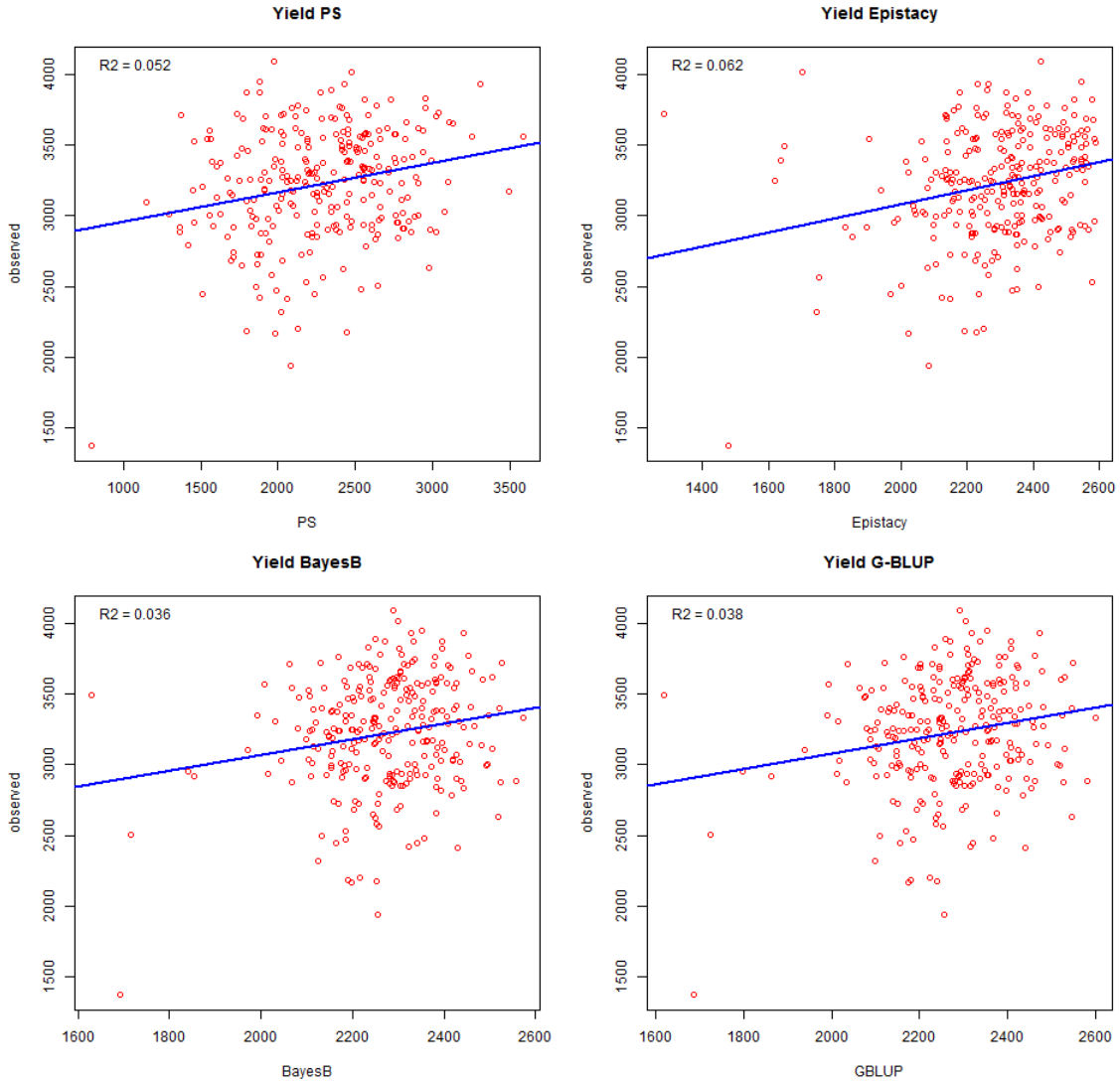

Figure 2.1 Yield (kg ha^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 276 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, and G-BLUP) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes.

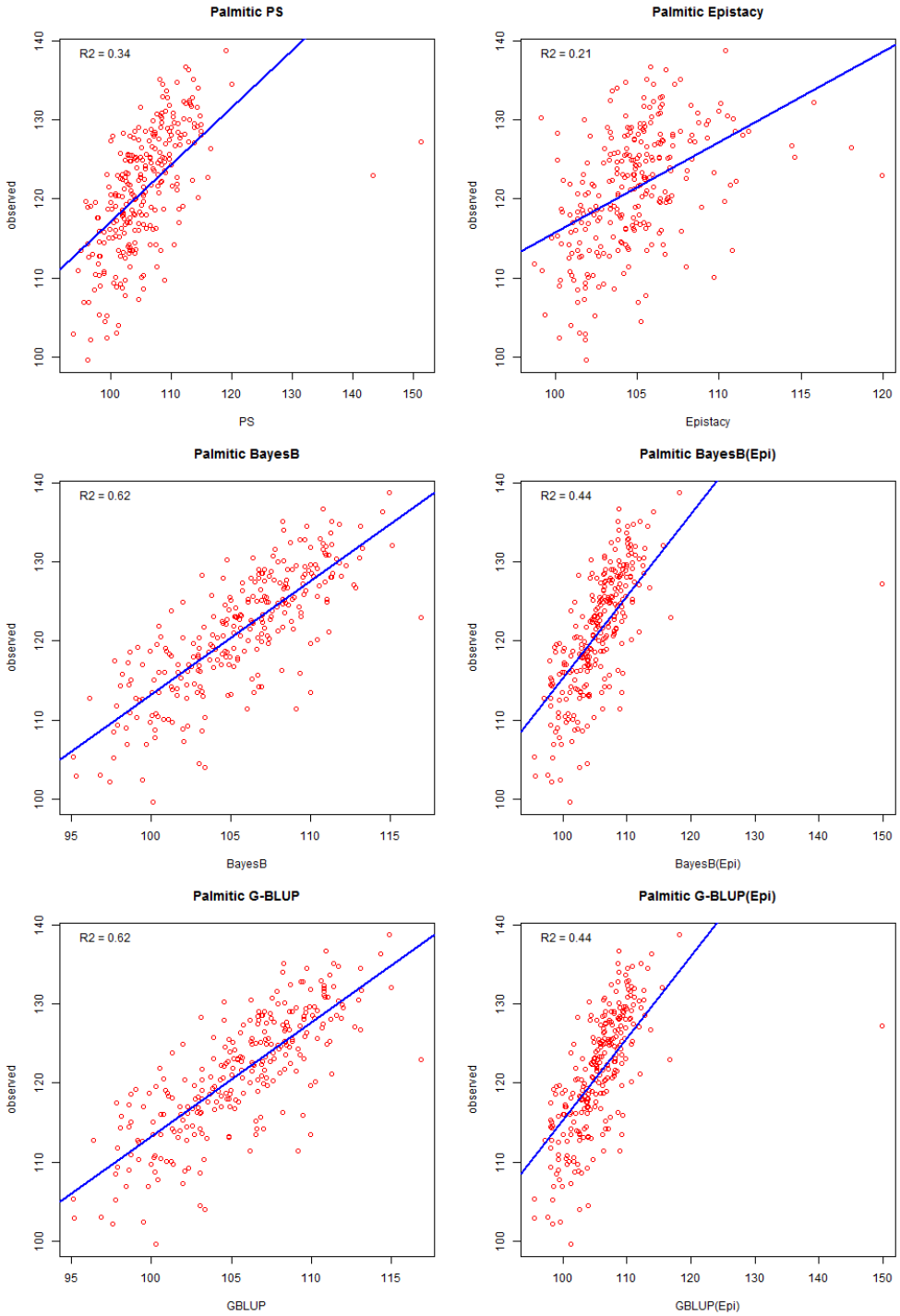

Figure 2.2 Palmitic acid (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes.

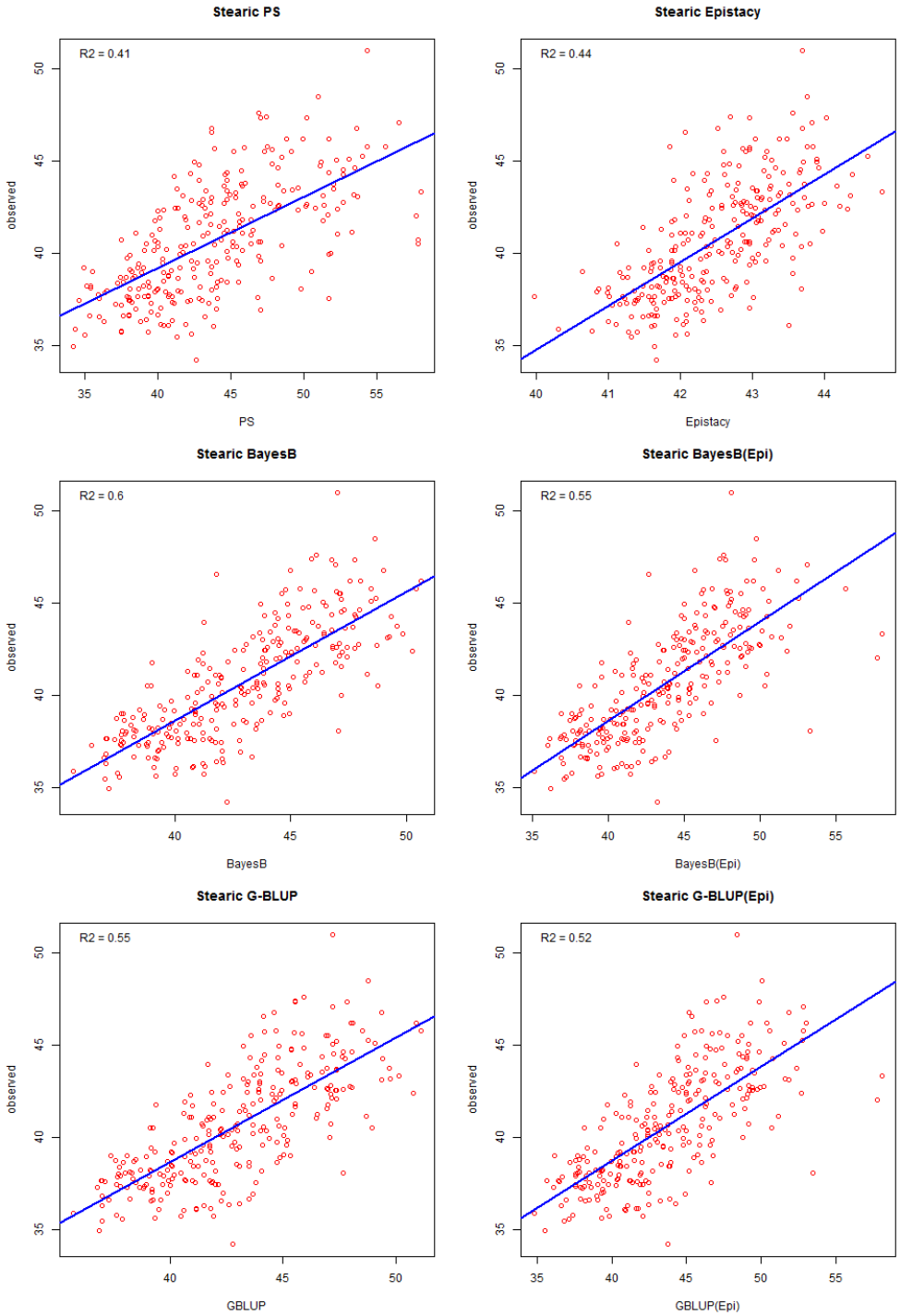

Figure 2.3 Stearic acid (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes.

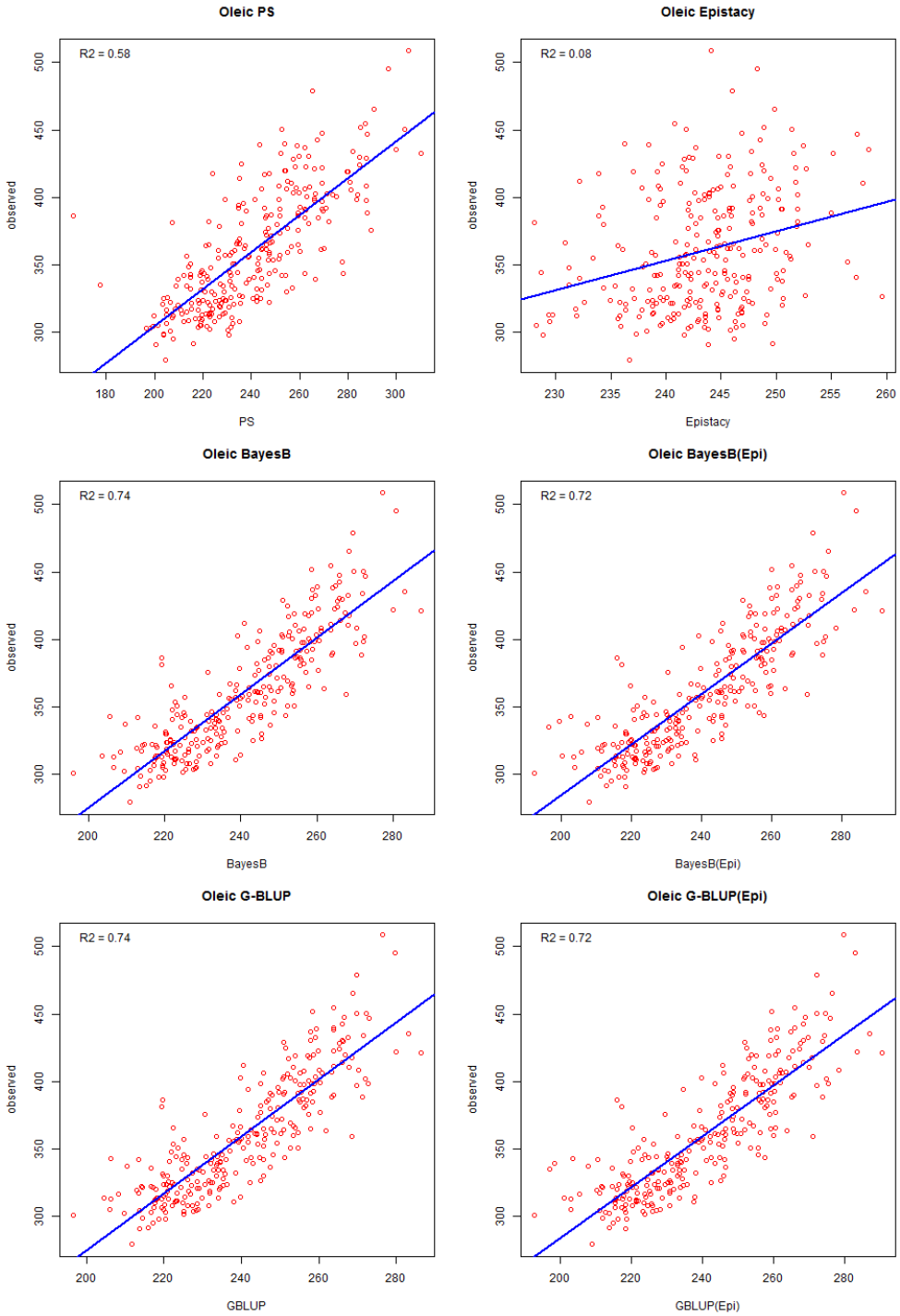

Figure 2.4 Oeic acid (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes.

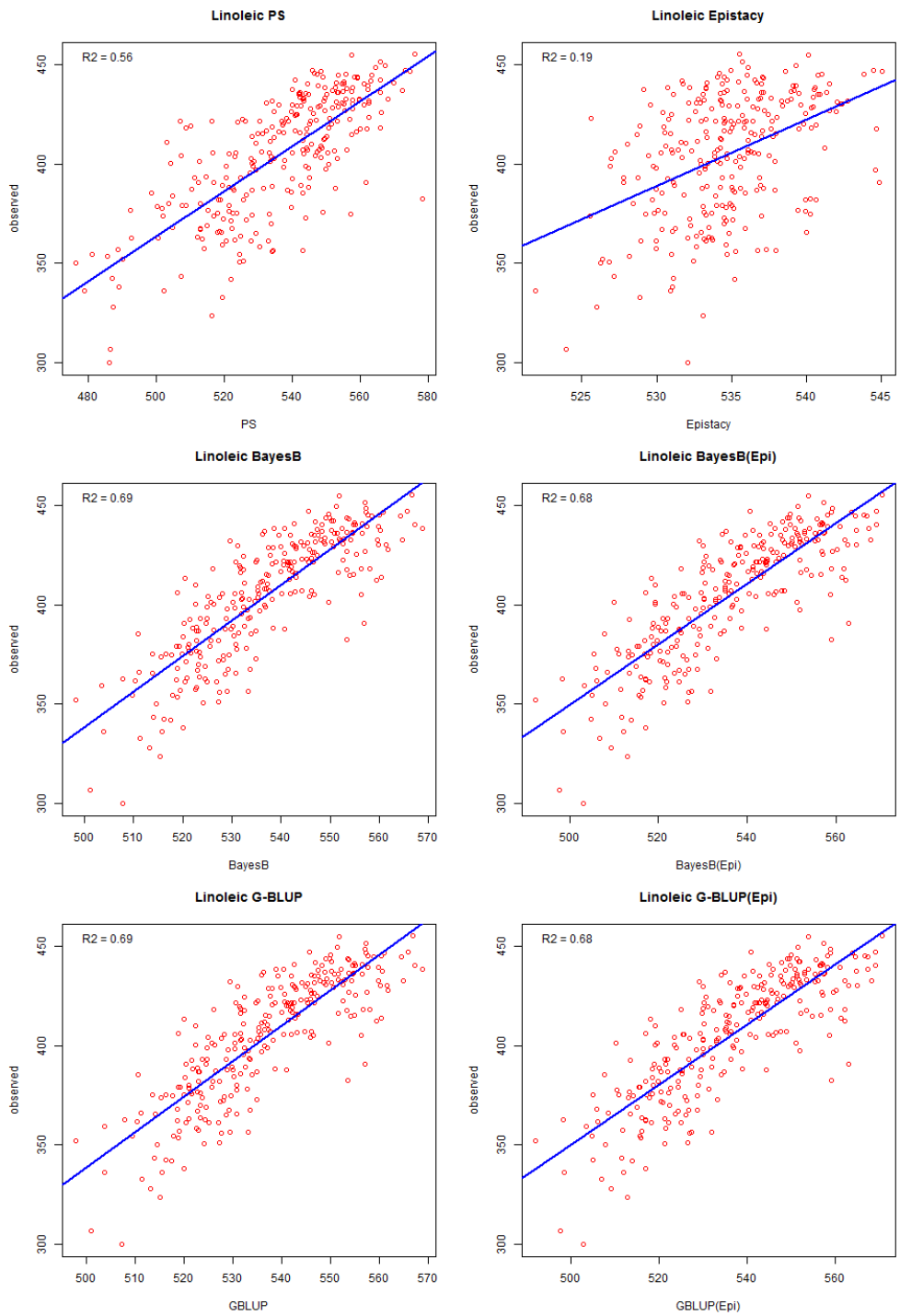

Figure 2.5 Linoleic acid (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes.

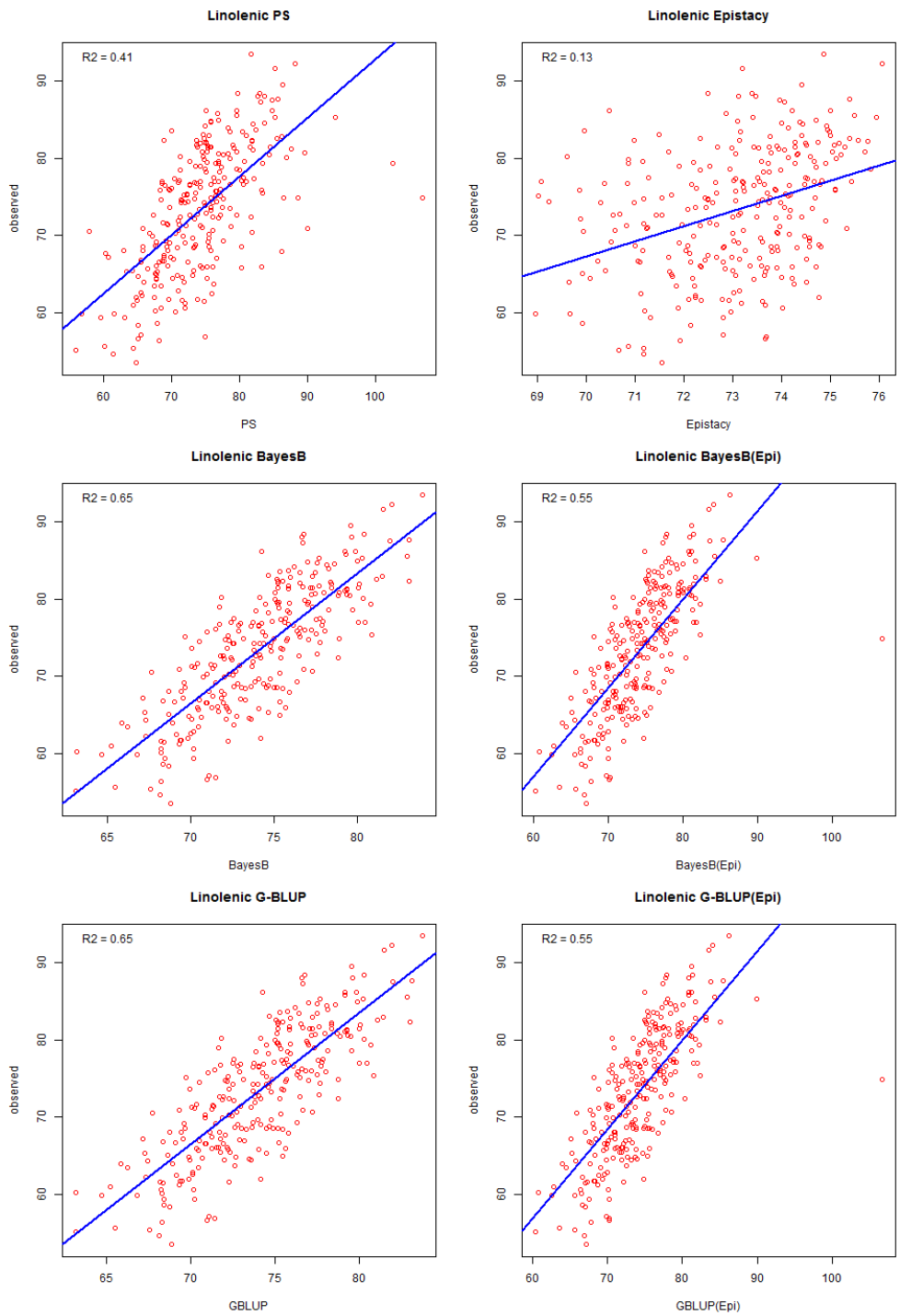

Figure 2.6 Linolenic acid (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 275 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes.

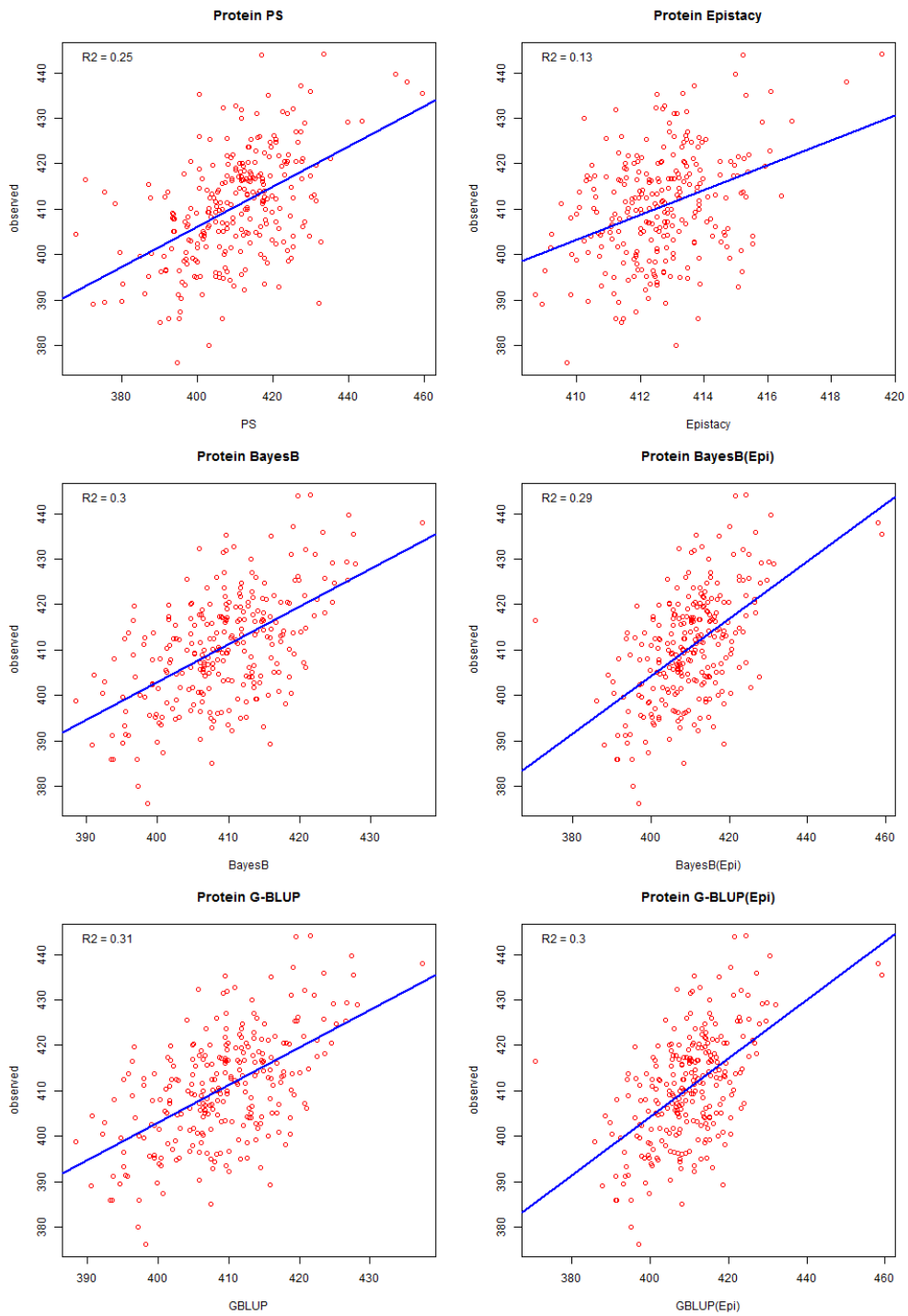

Figure 2.7 Protein (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 271 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes.

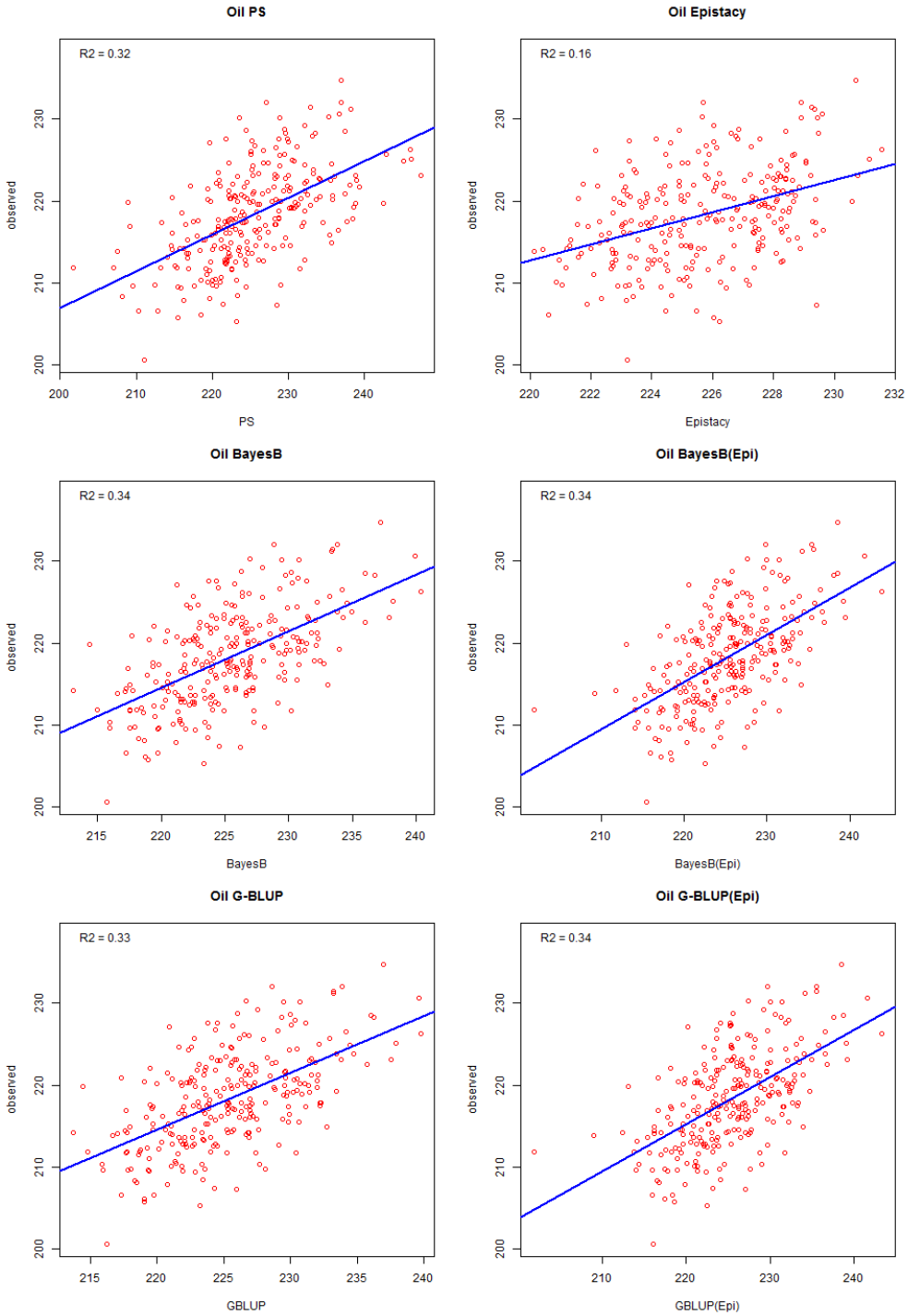

Figure 2.8 Oil (g kg^{-1}) performance comparisons between 2010 predictions (X-axis) and 2013 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 271 F5 derived RILs. The 2010 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2013 observed phenotypes.

CHAPTER 3

MIXED RESULTS BETWEEN PHENOTYPIC AND MOLECULAR BREEDING METHODS FOR SOYBEAN QUANTITATIVE TRAITS PREDICTED FROM REPLICATED FIELD TRIALS

Abstract

In order to achieve the best results in plant breeding, it is necessary to compare different selection strategies for prediction performance in targeted traits. Yield, fatty acids, protein, and oil are all commercially important traits in soybean [Glycine max (L.) Merrill] that merit selection strategy comparison. Since each of these traits displays quantitative inheritance, it is of interest to evaluate breeding methods that can account for a broad range of genetic effects. Along with phenotypic selection (PS), the molecular breeding methods chosen for this study were BayesB, G-BLUP, Epistacy, BayesB(Epi), and G-BLUP(Epi). These methods were evaluated in a soybean population consisting of 276 F5 derived recombinant inbred lines (RILs), which was genotyped with 17,236 polymorphic SNPs using the SoySNP50K BeadChip. Each RIL was grown in a single plot in 2010 in Knoxville, Tennessee; followed by a replicated, multi-location field trial in 2013. The phenotypic data from 2010 and 2013 were analyzed together and combined with the genotypic data in order to make predictions with the methods mentioned above. A subset of 203 RILs from this population was then grown in multi-location, replicated field trials in 2014 in order to compare each selection method using Spearman correlations and 15% tail selection contrasts for comparison. For yield, PS was the best selection strategy by a wide margin. Also, yield differed from all other traits in the performance of Epistacy; which was typically the least successful method, but for yield it was the best molecular method. While the fatty acids and NIRS traits displayed variability in which selection method(s) were preferred, a common theme was that there was little difference between PS and GS (BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi)) selection methods; with either strategy being useful for making improvements depending on the goals and resources of the researchers making the selections.

Introduction

Soybean [Glycine max (L.) Merrill] is a globally important crop produced for a wide range of purposes. Primary components of soybean seed that contribute to its high value are protein (~400 g kg^{-1}) and oil (~200 g kg^{-1}). Simultaneous improvement of protein and oil is quite challenging, as there is strong historical evidence of a negative correlation between these two traits (Yaklich et al., 2002). Further, oil and yield share a positive relationship, while protein and yield have a negative relationship (Morrison et al., 2008). Because of this, increases in soybean oil and yield must be sought after while maintaining adequate protein levels (Cober et al., 2009).

Five primary fatty acids exist in soybean oil. These fatty acids are palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2) and linolenic (18:3); typically occurring in relative concentrations of 100, 40, 220, 540, and 100 g kg^{-1} of total lipids, respectively (Wilson, 2004). Adjusting relative fatty acid concentrations within soybean oil is gaining importance, particularly with the Food and Drug Administration (FDA) recently banning partially hydrogenated oils (PHOs) in all food products; considered to be no longer Generally Recognized as Safe (GRAS) (https://www.federalregister.gov/articles/2015/06/17/2015-14883/final-determination-regarding-partially-hydrogenated-oils). Because of this, a major initiative in fatty acid improvement is to reduce linolenic acid (< 30 g kg^{-1}), thereby reducing the need for hydrogenation, a process that creates PHOs in soybean oil.

Another major goal of soybean fatty acid improvement is to increase monounsaturated oleic acid (> 800 g kg^{-1}). When compared with saturated fatty acids, oleic acid has been shown to lower cholesterol when consumed by humans (Kris-Etherton and Yu, 1997). Additionally, increasing the concentration of oleic acid in soybean oil results in improved oxidative stability; which leads to

increased shelf life of soybean oil food products (Kinney, 1996) and biodiesel (Kinney and Clemente, 2005; Fallen et al, 2012).

Improving soybean fatty acids has been the focus of much recent research (Pantalone et al., 2002; Pham et al., 2010; Bilyeu et al., 2011; Boersma et al., 2012; Gillman et al., 2014). Coinciding with these efforts is the need to evaluate the relative utility of various selection strategies for fatty acids; as well as for yield, protein, and oil. Evaluating selection methods that can account for a broad range of genetic effects would be useful, as each of these traits displays quantitative inheritance. While targeted goals for oleic acid and linolenic acid have been achieved using relatively few loci (Pham et al., 2010; Bilyeu et al., 2011), there is still concern for oleic acid that environmental variation may result in levels that are below 800 g kg^{-1} (Lee et al., 2012; Fallen et al., 2012). In such cases, it would be useful to evaluate breeding methods that can act as a fine-tuning approach using a broad range of genetic effects in order to provide more consistent results. In addition to fatty acids, such methods would be useful to explore for yield, protein, and oil.

A worthwhile method to explore would be genomic selection (GS), which accounts for genetic effects across the entire genome (Nakaya and Isobe, 2012). First described by Meuwissen et al. (2001), GS is the simultaneous selection of thousands of markers densely spread across the entire genome; with any gene affecting the trait of interest expected to occur in linkage disequilibrium with some markers (Meuwissen, 2007). Many studies have evaluated the utility of GS in animal and plant breeding with some success (Ødegård et al., 2009; Lillehammer et al., 2010; Resende et al., 2012; Poland et al., 2012; Sitzenstock et al., 2013; Crossa et al., 2014). Since high potential for GS has been demonstrated using a cross-validation approach, there is a need to evaluate its accuracy over multiple generations in crop studies (Jonas and de Koning, 2013). Thus, the purpose of this research will be to evaluate various GS strategies in comparison with other

molecular breeding strategies and with phenotypic selection (PS) for relative value in selections made from replicated field tests for soybean yield, fatty acids, protein, and oil.

Materials and Methods

Plant Materials

A population of 860 recombinant inbred lines (RILs) was developed for evaluation of selection methods. The parental lines used were 'Essex' (Smith and Camper, 1973) and 'Williams 82' (Bernard and Cremeens, 1988), with the population hereafter known as E×W-50K. Essex (southern cultivar) and Williams 82 (northern cultivar) were chosen as parents to sample the genetic diversity between different breeding groups. Descriptions of these two cultivars illustrate some of the genetic differences, with Essex in maturity group (MG) V, with a determinate growth habit, purple flower, and gray pubescence; while Williams 82 is in MG III, with indeterminate growth habit, white flower, and tawny pubescence. Seed from the parental lines was obtained from the USDA soybean germplasm collection (www.ars-grin.gov). A random single plant of both Essex and Williams 82 was intentionally selfed for two generations to provide highly homozygous parental lines to be crossed for RIL development. The development of the E×W-50K population is further described in Chapter 2.

In 2010 each of the 860 RILs were planted in a single plot consisting of two adjacent rows 6.1 m in length, with the rows spaced 0.8 m apart. In addition to the RILs and parents, four checks from relevant MGs were included in the 2010 field test: 'LD00-3309' (MG IV-early) (Diers et al., 2006), 'IA4004' (MG IV-early), '5002T' (MG V-early) (Pantalone et al., 2004), and '5601T' (MG V-mid) (Pantalone et al., 2003).

A subset of this population (276 RILs) ranging in maturity from MG IV-mid to IV-late was chosen for advancement into replicated field trials planted in 2013. The field test design was a randomized complete block design (RCBD) with three replications per environment at three environments (Knoxville, TN; Springfield, TN; and Milan, TN), representative of the eco-geographic regions of East, Middle, and West Tennessee, respectively. In addition to the RILs and parents, three maturity checks were included: 'LD00-3309' (MG IV-early), 'LD00-2817P' (MG IV-mid) (Diers et al., 2010), and 'Ellis' (MG IV-late).

Using data from the combined 2010 and 2013 growing seasons, the maturity range in Julian calendar date was 260.56-276.37 (Table 3.1) (note: all tables and figures in this chapter are located in Appendix C). To reduce the maturity range, a new subset of 203 RIL ranging in maturity from MG IV-mid to IV-late was chosen for advancement into replicated field trials planted in 2014. The field design, locations, and checks were consistent from 2013 to 2014 with one exception; 'IA4005' (MG IV-early) was added as an additional maturity check. For each of the three field seasons, flower color was determined at the R2 growth stage; and pubescence color, plant height, and maturity were determined at the R8 growth stage (Fehr and Caviness, 1977). For each field season, plots were harvested at maturity. Yield was measured in kg ha^{-1} after adjusting the plot weight to 13% moisture.

Seed Quality Trait Detection

Fatty acid measurements for palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid were performed in the same manner as described in Chapter 2. Briefly, gas chromatography was used to analyze seed from each plot from the 2010, 2013, and 2014 field tests with a procedure described by Spencer et al. (2004) using a Hewlett Packard HP 6890 series gas chromatograph (Agilent Technologies, Santa Clara, CA) system. This gas chromatograph was equipped with a 7683 auto sampler, a 7673 flame ionization detector, and an immobilized

30 m x 0.53 mm inner diameter Agilent DB-23 capillary column with 0.5 µm fused stationery phase. Estimates for fatty acids were obtained as percentage of seed oil, and converted to g kg^{-1} seed oil.

Near infrared reflectance spectroscopy (NIRS) was used to obtain protein and oil estimates using the same methods described in chapter 2. Briefly, approximately 25 g of seed from each plot in the 2010 growing season was uniformly ground for 20 sec in a Knifetec 1095 Sample Mill (FOSS Tecator, Hoganas, Sweden). These ground samples were analyzed using the NIRS instrument (NIR 6500, FOSS North America) as described by Panthee et al. (2006), except that in this study samples were scanned using updated ISIscan software v. 2.85. In collaboration with the University of Minnesota, each plot from the 2013 season was scanned as a whole bean sample using a Perten DA 7200 Diode Array (Perten, Hägersten, Sweden) NIRS instrument. The calibration equations were developed from a cooperative effort between Perten and University of Minnesota (Bolon et al., 2011). Plots from the 2014 growing season were scanned using the same procedure as the 2013 plots with a Perten DA 7250 Diode Array NIRS instrument. For each NIRS analysis, values for protein and oil concentration were adjusted to g kg^{-1} on a dry weight basis.

SNP Genotyping and Marker Cleanup

A more detailed description of the SNP genotyping and marker cleanup can be found in Chapter 2. Briefly, DNA samples were collected from leaf tissue for each RIL and analyzed using the Illumina Infinium beadchip SoySNP50K (Song et al., 2013) in 2009 at the Soybean Genomics Laboratory at the USDA Beltsville Agricultural Research Center (USDA-ARS) in Beltsville, MD. The marker positions from the genotyping analysis were taken from the Williams 82 whole genome sequence (WGS) (Glyma 1.01) (Schmutz et al., 2010). Markers that were missing in ≥ 5% (43 RILs) of the population were dropped using the ‘synbreed’ package (Wimmer et al., 2012) in the R language and environment for

statistical computing (R Core Team, 2015), with 16,718 polymorphic SNPs remaining for analysis. Missing SNPs were then imputed using default settings in the Beagle Genetic Analysis Software Package v. 3.3.1 (Browning and Browning 2007; 2009) via the 'synbreed' package (Wimmer et al., 2012) in the R language and environment for statistical computing (R Core Team, 2015). Finally, a screening for potential genotyping errors was performed using default settings in the new version of the 'calc.errorlod' function within the 'qtl' package (Broman et al., 2003) in R, with no errors detected.

Selection Methods and Statistical Analysis

As in chapter 2, genomic selections (GS) were performed with the 'BGLR' package (Pérez and de los Campos, 2014) in the R language and environment for statistical computing (R Core Team, 2015). In the combined analysis 276 RILs planted in multi-location, replicated field tests in 2010 and 2013 (10_13) were used to generate predictions for fatty acids (palmitic, stearic, oleic, linoleic, and linolenic), protein, and oil. Predictions for yield were generated in the same manner with the exception that only data from the Knoxville location was used, as there was no significant difference ($P > 0.05$) for yield among RILs in the multi-location field test (Table 3.2). As this population segregates at the E1 (maturity), E3 (maturity), and Dt1 (growth habit) loci, SNPs located within or adjacent to (< 5 kb) each locus (ss715593840, ss715635705, and ss715635423 or if missing ss715635422, respectively) were used to determine the expected parental allele. Four exceptions were made at the Dt1 locus, with updated growth habit calls made based on field observation. Loci for E1, E3, and Dt1 were included as covariates in the GS models in order to minimize any associated noise in variability.

Genomic best linear unbiased predictor (G-BLUP) and BayesB (Meuwissen et al., 2001) were the GS models chosen for analysis. The reason for choosing G-BLUP is that it is easily implemented, and thus is a commonly used method;

while BayesB was chosen because it is also a common method, and it has ability to depart from the infinitesimal model as well as preference in comparison with Bayesian penalized regression models (de los Campos et al., 2013). The three covariates listed above (E1, E3, and Dt1) were included as fixed effects for both GS and BayesB, along with 40,000 iterations, a burn in of 10,000, and heritability estimates from the 10_13 analysis (Table 3.1); otherwise the default settings were used. Cross-validations were performed in the same manner as in chapter 2, with 50 replications for each trait. For each rep, phenotypic data was removed from 1/5 of the population (test set), while phenotypic and genotypic information remained for the remaining 4/5 of the population (training set). The Pearson correlation coefficients were compared for BayesB and G-BLUP using a paired-t test (Pérez and de los Campos, 2014).

Following the procedure described in chapter 2, an additional selection model was performed using the Epistacy macro v. 2.0 (Holland, 1998) in SAS (SAS Institute Inc., Cary, NC, USA, SAS 9.4, 2002-2012), with modifications provided by Arnold Saxton. The reason for including this model in the analysis was to account for significant epistatic interactions that influence yield, fatty acids, protein or oil. A P value cutoff of 0.001 was chosen to identify significant interactions. Markers were screened for variation among RILs using the 'findDupMarkers' function in the R 'qtl' package (Broman et al., 2003), with one randomly chosen marker from each duplicate set remaining for analysis. After duplicate markers were removed, 4233 SNPs remained. Deviations due to these interactions were then summed, divided by 4233, and added to the mean for predicted performance for each RIL.

Additional predictions using combined output from the GS and Epistacy models were estimated using the GS predictions, along with the 20 Epistacy interactions with largest effect. Through multiple regression analysis against the phenotypic values using PROC REG in SAS (SAS Institute Inc., Cary, NC, USA, SAS 9.4,

2002-2012), combinations of GS method and Epistacy interactions with the highest adjusted R squared value were chosen. The resulting predictions from these regressions were then used to produce the G-BLUP(Epi) and BayesB(Epi) selection methods. These methods were included in the analysis as an effort to account for large effect epistatic interactions in the GS models.

Performance for each molecular breeding method (BayesB, G-BLUP, Epistacy, BayesB(Epi), and G-BLUP(Epi)), along with phenotypic selection (PS), was then evaluated in the 203 RIL population subset grown in 2014. A regression was plotted for each selection against 2014 observed values in the R language and environment for statistical computing (R Core Team, 2015) in order to visualize the degree of relationship with the 2014 observed phenotypes. Additionally, Spearman correlations between each selection method with the 2014 observed phenotypes were compared using the 'cocor.dep.groups.overlap' function in the 'cocor' package (Diedenhofen and Musch, 2015) in R using the using the Hittner et al. (2003) method for comparing dependent, overlapping correlations. Finally, the performance of 15% (30 RILs) high and low tail selections chosen from each selection method was assessed in the 2014 field season using 'estimate' and 'contrast' statements in SAS PROC GLIMMIX (SAS Institute Inc., Cary, NC, USA, SAS 9.4, 2002-2012). An RCBD model was used for this analysis, with RIL as the fixed term and location, rep(location), and RIL × location as random terms, and denominator degrees of freedom method set to residual. An exception occurred for yield in the 2014 analysis, as there was only one location. Thus, the model terms were RIL (fixed) and rep (random). This problem did not occur for the 10_13 yield analysis, as each year was treated as a different location.

Further statistical analyses were performed in SAS PROC GLIMMIX (SAS Institute Inc., Cary, NC, USA, SAS 9.4, 2002-2012) with some adjustments to fixed and random terms. A model with RIL, location, and RIL × location as fixed terms and rep(location) as random was run (except in yield 2014) in order to test

for significant RIL × location interactions. Additionally, a model with no fixed terms and RIL, location, rep(location), and RIL × location as random terms was run for all analyses except yield 2014 in order to obtain the variance for each term. The corresponding analysis for yield 2014 included no fixed terms and RIL and rep as random terms. These variance estimates were then used to calculate heritability on an entry means basis (Nyquist, 1991).

Results

Fatty acids, protein, oil, maturity, and height all had significant differences ($P < 0.05$) among RILs for the combined 10_13 data analysis (Table 3.1) and the 2014 analysis (Table 3.2). However, there were no differences ($P > 0.05$) among RILs for yield in the 2014 multi-location analysis even with estimates ranging from 2321.39-4022.70 kg ha^{-1} (Table 3.2). Therefore, analyses from the Knoxville only datasets were used for predictions in 10_13 (Table 3.1) and validations in 2014 (Table 3.2). A possible reason for 2014 yield differences in Knoxville, but not across locations, is the highly significant ($P < 0.001$) RIL × location (genotype × environment) interaction (Table 3.2). This reason is further supported by the extreme difference in heritability between Knoxville only (0.69) and multi-location (rounded to 0.00) (Table 3.2), indicating that most of the variation for yield in the multi-location analysis was due to non-genic effects. The heritability estimates from the 10_13 analysis used for prediction were lowest for yield (0.43), followed by protein (0.87) and oil (0.88), and then fatty acids (0.92-0.94) (Table 3.1). With the range of heritability estimates indicating different levels of gain from selection, it was of interest to evaluate various selection methods for relative utility for these important soybean traits.

For the BayesB and G-BLUP selection methods, an initial comparison was available using cross-validation. For the eight traits evaluated, BayesB was favored ($P < 0.05$) for palmitic, stearic, oleic, linolenic, and protein; G-BLUP was favored ($P < 0.05$) for yield; and there was no difference between methods ($P >$

0.05) for linoleic and oil (Table 3.3). Cross-validations with G-BLUP and BayesB were quite high (> 0.71) for all traits except yield (< 0.34) (Table 3.3).

Regression plots for each trait and selection method are displayed (Figures 3.1-3.8) in order to visualize the degree of relatedness between the 10_13 predictions and the 2014 observed phenotypes. For all traits, Epistacy had the lowest R squared value (Figures 3.1-3.8). However, this is somewhat misleading for yield, as each of the R squared values are quite close to zero (< 0.02) (Figure 3.1), indicating that any of the selection methods would have very limited success. PS was the selection method with the highest R squared value for palmitic (0.89), stearic (0.76), oleic (0.93), linoleic (0.92), and oil (0.81) (Figures 3.2-3.5, 3.8). This trend did not continue for linolenic and protein. BayesB, G-BLUP, BayesB(Epi), and GBLUP(Epi) all tied for the highest R squared value (0.90) for linolenic; while G-BLUP, BayesB(Epi), and GBLUP(Epi) all achieved the highest R squared value (0.81) for protein (Figures 3.5-3.6).

Spearman correlations are provided to better understand the association between rankings for 2014 observed phenotypes and each 10_13 selection method (Tables 3.4-3.11). The Hittner et al. (2003) method for comparing dependent, overlapping correlations was used to determine which methods match the 2014 observed phenotypes most closely in rank. For yield, none of the selection method correlations with 2014 phenotypic rankings significantly differed (P > 0.05) from zero (Table 3.4). However, when comparing these correlations with each other, differences were observed. PS (0.13) was convincingly the best method, with a significantly higher correlation (P < 0.05) than other methods; indeed the only positive correlation with 2014 phenotypic rankings (Table 3.4). The G-BLUP correlation (-0.12) with 2014 phenotypic rankings was lower than (P < 0.05) all other methods except Epistacy (-0.01), for which there was no difference (P > 0.05) (Table 3.4).

For all of the fatty acids, Epistacy had a lower ($P < 0.05$) correlation with 2014 phenotypic rankings than any other selection method. Similar to yield, PS was the most closely correlated with 2014 phenotypic rankings for palmitic, stearic, oleic, and linoleic; exceeding ($P < 0.05$) or tying ($P > 0.05$) all other selection methods for these traits (Tables 3.5-3.8). However, this trend was reversed for linolenic, with BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi) all higher correlation values ($P < 0.05$) than PS while not differing ($P > 0.05$) from each other (Table 3.9).

As with the fatty acids, Epistacy had a lower ($P < 0.05$) correlation with 2014 phenotypic rankings than any other selection method for protein and oil. Along with yield, palmitic, stearic, oleic, and linoleic, PS was the most closely correlated with 2014 phenotypic rankings for protein and oil; exceeding ($P < 0.05$) or tying ($P > 0.05$) all other selection methods (Tables 3.10-3.11). Yet PS (0.88) was not a clear favorite for protein, not differing ($P > 0.05$) from BayesB (0.89), G-BLUP (0.89), BayesB(Epi) (0.89), or G-BLUP(Epi) (0.89) (Table 3.10).

Contrasts are provided in order to compare 15% high and low tail selections for each method with all traits (Tables 3.12-3.19). With four exceptions, the 15% high and low tails from the 2014 phenotypic rankings were significantly ($P < 0.05$) higher than the high tails and lower than the low tails for all traits using each selection method (Tables 3.12-3.19). These four exceptions that did not differ ($P > 0.05$) from the 2014 phenotypic ranking tail selections were PS and BayesB(Epi) in the low tail for palmitic (Table 3.13), PS in the low tail for oleic (Table 3.15), and PS in the high tail for linoleic (Table 3.16).

Based on tail selections, PS was clearly the preferred method for yield, outperforming ($P < 0.05$) all other methods except for Epistacy in the high tail ($P < 0.05$) (Table 3.12). Epistacy (2980.92 kg ha^{-1}) showed the best results of any of the molecular methods, producing greater results ($P < 0.05$) than BayesB

(2907.16 kg ha^{-1}), G-BLUP (2845.94 kg ha^{-1}), and G-BLUP(Epi) (2878.75 kg ha^{-1}) in the high tail, and acting as the only molecular method that was not overcome by PS in both head to head comparisons (Table 3.12). Additionally, while the high and low tails were not compared statistically, it is worth noting that Epistacy was the only molecular method with a numerically higher value in the high tail when compared with the low tail (Table 3.12). The poorest performing of all of the selection methods for yield was G-BLUP; losing ($P < 0.05$) or not differing ($P > 0.05$) from other methods in each comparison (Table 3.12).

For the fatty acids, Epistacy was consistently the poorest performing selection method, being outcompeted ($P < 0.05$) in every tail comparison except for the low tail of linoleic, in which it did not differ ($P > 0.05$) from PS (Tables 3.13-3.17). Selecting the best method for fatty acids from the tail comparisons was not as straightforward. For palmitic, PS, BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi) all outperformed ($P < 0.05$) Epistacy in both tails while not differing ($P > 0.05$) from each other (Table 3.13). It should be noted that both PS (112.42 g kg^{-1}) and BayesB(Epi) (112.53 g kg^{-1}) did not differ ($P > 0.05$) in the low tail from the 2014 phenotypic low tail (Table 3.13), setting these two methods apart as the most promising for palmitic based on tail selections. For stearic, PS, G-BLUP, and G-BLUP(Epi) each outperformed ($P < 0.05$) or did not differ ($P > 0.05$) from other selection methods (Table 3.14). Of these three methods G-BLUP has a slight edge as the preferred method ($P < 0.05$) in 4 of 10 comparisons, while PS and BayesB(Epi) were only preferred ($P < 0.05$) in 3 of 10 comparisons (Table 3.14). For oleic, PS, BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi) all outperformed ($P < 0.05$) Epistacy in both tails while not differing ($P > 0.05$) from each other (Table 3.15). However, PS was the only method that was not outperformed ($P < 0.05$) by both 2014 phenotypic tails; with no difference occurring in the low tail ($P > 0.05$) (Table 3.15). For linoleic, PS, BayesB, and BayesB(Epi) each outperformed ($P < 0.05$) or did not differ ($P > 0.05$) from other selection methods (Table 3.16). While BayesB and BayesB(Epi) each were

preferred ($P < 0.05$) in 3 of 10 comparisons, PS was only preferred ($P < 0.05$) in 1 of 10 comparisons (Table 3.16). Additionally, while PS was the only method to match ($P > 0.05$) the 2014 phenotypic high tail, it was also the only method that did not outperform ($P > 0.05$) Epistacy in the low tail (Table 3.16). For linolenic, PS, G-BLUP, BayesB(Epi), and G-BLUP(Epi) each outperformed ($P < 0.05$) or did not differ ($P > 0.05$) from other selection methods (Table 3.17). Of these methods, BayesB(Epi) is given slight preference by winning ($P < 0.05$) 3 of 10 comparisons, while the PS, G-BLUP, and G-BLUP(Epi) each won ($P < 0.05$) only 2 of 10 comparisons (Table 3.17).

Similar to the fatty acids, Epistacy was the lowest performing method in tail selections for protein and oil; being outcompeted ($P < 0.05$) in every comparison with other methods (Tables 3.18-3.19). For the remaining selection methods (PS, BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi)), there were no differences ($P > 0.05$) in any comparisons with each other for protein or oil (Tables 3.18-3.19).

Discussion

As in chapter 2, there was no consensus among traits for best overall selection method based on Spearman correlations with 2014 observed phenotypes and contrasts between 15% tail selections. Yield displayed several inconsistencies in comparison with the other traits in the study. One example is that yield is the only trait for which none of the molecular breeding methods were reasonably competitive with PS. Another is that for yield, Epistacy was the preferred molecular method, while for every other trait Epistacy was the poorest performing method. This could indicate that epistatic interactions produce a much larger effect on yield than on other traits in this study. Additionally, the large difference in heritability estimate between yield and each of the other traits is worth noting (Tables 3.1-3.2). In agreement with this study, previous research has shown that protein and oil have higher estimates for heritability than yield (Wiggins, 2012).

As previously discussed, selection validations for yield were not possible using the 2014 multi-location analysis due to lack of genetic variation. Since yield was the trait with the lowest heritability (Tables 3.1-3.2), it is reasonable to expect that GS breeding methods (BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi)) would not perform well (Nakaya and Isobe 2012). Thus, PS was overwhelmingly the preferred method for soybean yield when selected from replicated field studies based on the results from this study.

In contrast with yield, no method was unanimously preferred for the fatty acids. The clearest trend was that Epistacy was the least desirable method for each fatty acid. In order to be a worthwhile selection strategy for fatty acids, adjustments to the Epistacy selection approach would be necessary. One possibility would be to relax the significance threshold used to choose which interactions could be included for selection, thus accounting for a greater degree of overall genomic effects. However, this could require some fine tuning; if all interactions between non-duplicated markers were to be included, approximately 8.9 million SNP interactions would have been in the prediction model. Using that much information to construct a selection model could quickly become unwieldly, particularly for studies with even more non-duplicated SNPs. Another possibility could be to choose SNPs based on haplotype; thus limiting the total number of interactions while still providing useful information on significant interactions.

Using results for both the Spearman correlations and the tail contrasts, PS was the preferred method for both oleic and linoleic; the two largest components of soybean oil. For palmitic, no clear separation for preferred method was observed between PS and BayesB(Epi). Notably, both of these methods matched ($P < 0.05$) the 2014 phenotypic 15% low tail, indicating that selections from these methods for decreased palmitic would be quite useful to breeders looking to reduce that trait. As with palmitic, BayesB(Epi) proved to be a useful method for linolenic, displaying a slight edge over the other breeding methods. However, in

contrast with palmitic, PS was determined to be the least useful method for linolenic. Stearic differed from all of the other fatty acids, with G-BLUP slightly out-competing the other breeding methods. Summarizing the results for all of the fatty acids, it could be argued that PS is slightly more effective than the GS breeding methods. Although perhaps a more useful summary would be that PS and GS breeding methods in the replicated field trial stage are mostly inseparable; with either being useful for improving fatty acids depending on the goals and resources of the researchers making the selections.

As in the fatty acids, Epistacy was the poorest performing method for protein and oil after evaluating each method based on Spearman correlations and tail contrasts. For protein, PS, BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi) all appear to be equivalent methods of selection. However, PS was the preferred method for oil, followed closely by the GS methods.

As opposed to the results from chapter 2 based on progeny row selections, the combined GS and Epistacy approaches (BayesB(Epi) and G-BLUP(Epi)) appeared to have more utility in selections made from replicated field trials. A simple explanation for this is that the regression prediction from SNP interactions and GS models was chosen based on adjusted R squared values in relation to phenotypes in the year of selection; therefore, training the GS model to be more similar to the PS model than it previously was. As the PS model performed better than BayesB and G-BLUP for several of the traits, it seems reasonable to assume that the performance of BayesB(Epi) and G-BLUP(Epi) would also be improved. This theory is supported for palmitic, with PS and BayesB(Epi) as the two leading methods. Yet, linolenic provides contradictory evidence, with BayesB(Epi) as the best method and PS as the worst method other than Epistacy.

Since the cross-validations were available as an additional comparison between BayesB and G-BLUP, it seems useful to discuss which method is preferred overall. In the direct comparisons for Spearman correlations, BayesB was favored ($P < 0.05$) once, while there was no difference seven times ($P > 0.05$) (Tables 3.4-3.11). For the direct contrast comparisons, BayesB was favored ($P < 0.05$) three times, G-BLUP was favored ($P < 0.05$) once, and there was no difference ($P > 0.05$) twelve times (Tables 3.12-3.19). Adding in the favorable results for BayesB from the cross-validations (Table 3.3), these results add support to the findings of Clark et al. (2011) that BayesB is a slightly preferred over G-BLUP as a GS method.

Conclusions

Evaluating different selection methods is a necessary process in order to determine the most useful methods for trait improvement. Additionally, for important soybean traits such as yield, fatty acids, protein, and oil, it is useful to evaluate the performance of different breeding methods when making selections from replicated field trials. Following evaluation of both phenotypic (PS) and molecular (BayesB, G-BLUP, Epistacy, BayesB(Epi), and G-BLUP(Epi)) selection strategies, it was determined that the relative utility of each breeding method varied based on the trait. For many of the traits (yield, palmitic, oleic, linoleic, and oil), PS was either the best or tied for the best method. This differs strongly from the progeny row selections in chapter 2, in which PS was outperformed by at least one molecular strategy for every trait.

No other trait than yield had such a clearly preferred method, with PS as the best selection strategy by a wide margin. Yield also differed from all other traits in the performance of Epistacy; which was typically the least successful method, but for yield it was the best molecular method. While the fatty acids and NIRS traits displayed variability in which selection method or methods were preferred, a common theme was that there was not much difference between PS and GS

(BayesB, G-BLUP, BayesB(Epi), and G-BLUP(Epi)) selection methods. In summary, PS and GS breeding methods for fatty acids, protein, and oil both have utility; with either being useful for making improvements depending on the goals and resources of the researchers making the selections. Further research evaluating these selection methods in a wide range of pedigrees or with elite germplasm would be useful for breeders seeking to incorporate these strategies into their programs.

References

Bernard, R.L., and C.R. Cremeens. 1988. Registration of ‘Williams 82’ soybean. Crop Sci. 28:1027–1028.

Bilyeu, K., J.D. Gillman, and A.R. LeRoy. 2011. Novel FAD3 mutant allele combinations produce soybeans containing 1% linolenic acid in the seed oil. Crop Sci. 51:259–264.

Boersma, J.G., J.D. Gillman, K.D. Bilyeu, G.R. Ablett, C. Grainger, and I. Rajcan. 2012. New mutations in a delta-9-stearoyl-acyl carrier protein desaturase gene associated with enhanced stearic acid levels in soybean seed. Crop Sci. 52:1736–1742.

Bolon, Y., W.J. Haun, W.W. Xu, D. Grant, M.G. Stacey, R.T. Nelson, D.J. Gerhardt, J.A. Jeddeloh, G. Stacey, G.J. Muehlbauer, J.H. Orf, S.L. Naeve, R.M. Stupar, and C.P. Vance. 2011. Phenotypic and genomic analyses of fast neutron mutant population resource in soybean. Plant Physiol. 156:240-253.

Broman, K.W., H. Wu, Ś. Sen, and G.A. Churchill. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890.

Browning, B.L., and S.R. Browning. 2009. A unified approach to genotype imputation and haplotype phase inference for large data sets of trios and unrelated individuals. Am. J. Hum. Genet. 84:210-223.

Browning, S.R., and B.L. Browning. 2007. Rapid and accurate haplotype phasing and missing data inference for whole genome association studies using localized haplotype clustering. Am. J. Hum. Genet. 81:1084-1097.

Clark, S.A., J.M. Hickey, and J.H.J. van der Werf. 2011. Different models of genetic variation and their effect on genomic evaluation. Genet. Sel. Evol. 43:18

Cober, E.R., S.R. Cianzio, V.R. Pantalone, and I. Rajcan. 2009. Soybean. In: J. Vollman and I. Rajcan, editors, Oil crops: Handbook of plant breeding, volume 4. Springer Science + Business Media LLC. p. 57-90.

Crossa, J., P. Pérez, J. Hickey, J. Burgeño, L. Ornella, J. Cerón-Rojas, X. Zhang, S. Dreisigacker, R. Babu, Y. Li, D. Bonnett, and K. Mathews. 2014.

Genomic prediction in CYMMIT maize and wheat breeding programs. Heredity 112:48-60.

De los Campos, G., J.M. Hickey, R. Pong-Wong, H.D. Daetwyler, and M.P.L.Calus. 2013. Whole-genome regression and prediction methods applied to plant and animal breeding. Genetics 193:327-345.

Diedenhofen, B., and J. Musch. 2015. Cocor: a comprehensive solution for the statistical comparison of correlations. PLoS ONE 10:e0121945.

Diers, B.W., T.R. Cary, D.J. Thomas, A. Colgrove, and T. Niblack. 2010. Registration of 'LD00-2817P' germplasm line with resistance to soybean cyst nematode from PI 437654. J. Plant Regist. 4:141-144.

Diers, B.W., T.R. Cary, D.J. Thomas, and C.D. Nickell. 2006. Registration of 'LD00-3309' soybean. Crop Sci. 46:1384.

Fallen, B.D., K. Rainey, C.E. Sams, D.A.Kopsell, and V.R. Pantalone. 2012. Evaluation of agronomic and seed characteristics in elevated oleic acid soybean lines in the south-eastern US. J. Am. Oil Chem. Soc. 89:1333-1343.

Federal Register. 2015. Final determination regarding partially hydrogenated oils. https://www.federalregister.gov/articles/2015/06/17/2015-14883/final-determination-regarding-partially-hydrogenated-oils (accessed 24 July 2015).

Fehr W.R., and C.E. Caviness. 1977. Stages of soybean development. Special Report, Agriculture and Home Economics Experiment Station, Iowa State University, 1977, issue 80, p 11.

Gillman, J.D., M.G. Stacy, Y. Cui, H.R. Berg, and G. Stacey. 2014. Deletions of the SACPD-C locus elevate seed stearic acid but also result in fatty acid and morphological alterations in nitrogen fixing nodules. BMC Plant Biol. 14:143.

Hittner, J. B., K. May, and N.C. Silver. 2003. A Monte Carlo evaluation of tests for comparing dependent correlations. J. Gen. Phychol. 130:149-168.

Holland, J.B. 1998. EPISTACY: A SAS program for detecting two-locus epistatic interactions using genetic marker information. J. Hered. 89:374–375.

Jonas, E., and D.J. de Koning. 2013. Does genomic selection have a future in plant breeding? Trends Biotechnol. 31:497-504.

Kinney, A.J. 1996. Development of genetically engineered soybean oils for food application. J. Food Lipids. 3:273-292.

Kinney, A.J., and T.E. Clemente. 2005. Modifying soybean oil for enhanced performance in biodiesel blends. Fuel Pro. Technol. 86:1137–1147.

Kris-Etherton, P.M., and S. Yu. 1997. Individual fatty acid effects on plasma lipids and lipoproteins: Human studies. Am. J. Clin. Nutr. 65:S1628–S1644.

Lee, J.D., K.D. Bilyeu, V.R. Pantalone, A.M. Gillen, Y.S. So, and J.G. Shannon. 2012. Environmental stability of oleic acid concentration in seed oil for soybean lines with FAD2-1A and FAD2-1B mutant genes. Crop Sci. 52:1290–1297.

Lillehammer, M., T.H.E. Meuwissen, and A.K. Sonesson. 2011. A comparison of dairy cattle breeding designs that use genomic selection. J. Dairy Sci. 94:493-500.

Meuwissen, T. 2007. Genomic selection: Marker assisted selectionon a genome wide scale. J. Anim. Breed. Genet. 124:321–322.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.

Morrison, M.J., E.K. Cober, M.F. Saleem, N.B. McLaughlin, J. Fregeau-Reid, B.L. Ma, W. Yan, and L. Woodrow. 2008. Changes in isoflavone concentration with 58 years of genetic improvement of short-season soybean cultivars in Canada. Crop Sci. 48:2201–2208.

Nakaya, A., and S.N. Isobe. 2012. Will genomic selection be a practical method for plant breeding? Ann. Bot. 110:1303-1316.

Nyquist, W.E. 1991. Estimation of heritability and prediction of selection response in plant populations. Crit. Rev. Plant Sci.10:235–322.

Ødegård, J., A.K. Sonesson, M.H. yazdi, and T.H.E. Meuwissen. 2009. Introgression of a major QTL from an inferior into a superior population using genomic selection. Genet. Sel. Evol. 41:38.

Pantalone, V.R., F.L. Allen, and D. Landau-Ellis. 2003. Registration of ‘5601T’ soybean. Crop Sci. 43:1123-1124.

Pantalone, V.R., F.L. Allen, and D. Landau-Ellis. 2004. Registration of ‘5002T’ soybean. Crop Sci. 44:1483-1484.

Pantalone, V.R., R.F. Wilson, W.P. Novitzky, and J.W. Burton. 2002. Genetic regulation of elevated stearic acid concentration in soybean oil. J. Am. Oil Chem. Soc. 79:543–553.

Panthee, D.R., V.R. Pantalone, C.E. Sams, A.M. Saxton, D.R. West, J.H. Orf, and A.S. Killam. 2006. Quantitative trait loci controlling sulfur containing amino acids, methionine and cysteine, in soybean seeds. Theor. Appl. Genet. 112:546–553.

Pérez, P., and G. de los Campos. 2014. Genome-wide regression and prediction with the BGLR statistical package. Genetics 198:483-495.

Pham, A.T., J.D. Lee, J.G. Shannon, and K.D. Bilyeu. 2010. Mutant alleles of FAD2-1A and FAD2-1B combine to produce soybeans with the high oleic acid seed oil trait. BMC Plant Biol. 10:195.

Poland, J., J. Endelman, J. Dawson, J. Rutkoski, S. Wu, Y. Manes, S. Dreisigacker, J. Crossa, H. Sanchez-Villeda, M. Sorrells, and J.L. Jannink. 2012. Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Gen. 5:103–113.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Resende Jr., M.F.R., P. Muñoz, J.J. Acousta, G.F. Peter, J.M. Davis, D. Grattapaglia, M.D.V Resende, and M. Kirst. 2012. Accelerating the domestication of trees using genomic selection: Accuracy of prediction models across ages and environments. New Phytol. 193:617-624.

SAS Institute Inc. 2002-2012. Cary, NC, USA. SAS 9.4)

Schmutz, J., S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D.L. Hyten, Q. Song, J.J. Thelen, J. Cheng, et al. 2010. Genome sequence of the palaeopolyploid soybean. Nature 463:178–183.

Sitzenstock, F., F. Ytournel, A.R. Sharifi, D. Cavero, H. Täubert, R. Preisinger, and H. Simianer. 2013. Effeciency of genomic selection in an established commercial layer breeding program. Gent. Sel. Evol. 45:29.

Smith, T.J., and H.M. Camper. 1973. Registration of Essex Soybean (Reg. No. 97). Crop Sci. 13:495.

Song, Q., D.L. Hyten, G. Jia, C.V. Quigley, E.W. Fickus, R.L. Nelson, and P.B. Cregan. 2013. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS ONE 8:e54985.

Spencer, M.M., D. Landau-Ellis, E.J. Meyer, and V.R. Pantalone. 2004. Molecular markers associated with linolenic acid content in soybean. J. Am. Oil Chem. Soc. 81:559–562.

USDA, ARS, National Genetic Resources Program. Germplasm Resources Information Network - (GRIN) [Online Database]. National Germplasm Resources Laboratory, Beltsville, Maryland. URL: http://www.ars-grin.gov.4/cgi-bin/npgs/html/index.pl?language=en (24 July 2015)

Wiggins, B.T. 2012. Heritability and genetic gain of seed protein, oil, and yield among RIL of soybean. M.S. thesis. Univ. of Tennessee, Knoxville, TN, USA.

Wilson, R.F. 2004. Seed composition. In: H.R. Boerma and J.E. Specht, editors, Soybeans: Improvement, production, and uses. 3rd ed. ASA, CSSA, and SSSA, Madison, WI. p. 621–678.

Wimmer, V., T. Albrecht, H.J. Auinger, and C.C. Schön. 2012. Synbreed: A framework for the analysis of genomic prediction using R. Bioinformatics. 28:2086-2087.

Yaklich, R.W., B. Vinyard, M. Camp, and S. Douglass. 2002. Analysis of seed protein and oil from soybean northern and southern region uniform tests.

Crop Sci. 42:1504–1515.

Appendix C-Chapter 3 Tables and Figures

Tables

Table 3.1 Simple statistics for soybean population E×W-50K subset consisting of 276 F5 derived RILs from the combined 2010 and 2013 field seasons. This dataset was used to make performance predictions for traits of interest in a subset of the population (203 RILs) grown in replicated field trials in 2014 at three locations (Knoxville, TN; Springfield, TN; and Milan, TN). Yield predictions were made using data from only the Knoxville, TN location (displayed in bold text).

Trait	Genotype P value	GxE P value	Essex	Williams 82	min	mean	max	std. dev.†	LSD value	h^{2}‡
Maturity (Julian)	***	***	273.06	262.50	260.56	270.49	276.37	2.60	3.10	0.82
Height (cm)	***	***	71.50	88.32	36.92	86.71	130.90	17.60	11.42	0.95
Yield (kg ha^{-1})	***	***	3342.67	2698.42	1206.34	2991.62	3748.34	356.57	590.76	0.64
Yield (kg ha^{-1}); Knoxville, TN	***	NS	2835.76	2469.62	1172.55	2629.79	3477.58	322.61	628.08	0.43
Palmitic (g kg^{-1} seed oil)	***	NS	121.45	107.29	97.28	117.24	134.33	7.11	5.27	0.93
Stearic (g kg^{-1} seed oil)	***	***	43.29	39.69	35.16	41.38	51.70	3.14	2.41	0.93
Oleic (g kg^{-1} seed oil)	***	***	321.40	371.90	255.25	330.66	467.79	40.94	27.78	0.94
Linoleic (g kg^{-1} seed oil)	***	***	437.45	415.85	339.66	437.25	486.81	28.98	20.71	0.94
Linolenic (g kg^{-1} seed oil)	***	*	76.52	65.21	54.97	73.48	92.06	7.73	6.05	0.92
Protein (g kg^{-1} seed dry weight)	***	***	424.56	420.97	378.76	410.40	442.57	11.43	11.47	0.87
Oil (g kg^{-1} seed dry weight)	***	***	213.52	228.86	202.72	220.04	240.05	5.76	5.69	0.88

*Significant at 0.05 probability level
***Significant at 0.001 probability level
NS, Not significant at 0.05 probability level
E×W-50K = soybean population with parental lines Essex and Williams 82
† std. deviation of LSMEANs
‡ heritability calculated using entry means basis (Nyquist, 1991)

Table 3.2 Simple statistics for soybean population E×W-50K subset consisting of 203 F5 derived RILs, planted in replicated field trials at three locations in 2014 (Knoxville, TN; Springfield, TN; and Milan, TN). Information from this dataset was compared with performance predictions for traits of interest in the larger subset (276 RILs) from the combined 2010 and 2013 field seasons. Yield comparisons were made using data from only the Knoxville, TN location (displayed in bold text).

Trait	Genotype P value	GxE P value	Essex	Williams 82	min	mean	max	std. dev.†	LSD value	$h^{2‡}$
Maturity (Julian)	***	***	274.89	267.53	271.00	273.77	278.22	1.15	2.20	0.53
Height (cm)	***	***	80.15	83.26	57.86	91.67	125.00	16.38	11.36	0.94
Yield (kg ha^{-1})	NS	***	3244.88	2194.39	2321.39	3061.43	4022.70	317.81	888.08	0.00
Yield (kg ha^{-1}); Knoxville, TN	***	NA	3109.18	1544.26	1849.48	2989.48	4108.38	390.86	612.58	0.69
Palmitic (g kg^{-1} seed oil)	***	***	125.95	122.91	98.90	123.40	137.84	6.97	5.52	0.92
Stearic (g kg^{-1} seed oil)	***	***	44.61	41.25	35.67	41.60	51.49	2.96	3.25	0.85
Oleic (g kg^{-1} seed oil)	***	***	331.11	326.52	269.69	339.65	490.99	37.90	32.81	0.90
Linoleic (g kg^{-1} seed oil)	***	***	422.77	434.13	318.30	422.64	466.31	26.02	22.84	0.90
Linolenic (g kg^{-1} seed oil)	***	***	75.56	75.09	52.16	72.71	93.06	7.91	6.04	0.93
Protein (g kg^{-1} seed dry weight)	***	***	443.22	441.36	404.73	433.29	462.09	10.79	12.01	0.84
Oil (g kg^{-1} seed dry weight)	***	***	205.50	221.68	191.90	212.15	228.04	5.42	5.42	0.87

***Significant at 0.001 probability level
NS, Not significant at 0.05 probability level
NA, Not applicable
E×W-50K = soybean population with parental lines Essex and Williams 82
† std. deviation of LSMEANs
‡ heritability calculated using entry means basis (Nyquist, 1991)

Table 3.3 Comparison of cross-validations for G-BLUP and BayesB methods of GS for soybean population E×W-50K subset consisting of 276 F5 derived RILs; the combined 2010 and 2013 datasets were used for analysis. All traits were analyzed with the multi-location dataset except yield, which was only analyzed with data from the Knoxville, TN location. Cross-validations were replicated 50 times for each trait. In each rep, a randomly chosen 1/5 of the population had phenotypic data removed (test set), while phenotypic and genotypic information were retained for the remaining 4/5 of the population (training set). The values displayed for G-BLUP and BayesB are the mean Pearson correlation coefficients for the predicted and observed values in the test set. Listed P values indicate if there is a statistically significant difference between GS methods. Methods with higher cross-validation correlation values are displayed with bold text.

	G-BLUP	BayesB	P value
Yield	**0.33**	0.33	*
Palmitic	0.86	**0.86**	***
Stearic	0.84	**0.87**	***
Oleic	0.86	**0.86**	*
Linoleic	0.85	0.85	NS
Linolenic	0.86	**0.87**	***
Protein	0.76	**0.76**	***
Oil	0.72	0.72	NS

*Significant at 0.05 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population w ith parental lines Essex and Williams 82

Table 3.4 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for soybean yield from only the Knoxville, TN location in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Yield	Correlation Differences	**14:P** 0.13	**14:B** -0.08	**14:G** -0.12	**14:E** -0.01	**14:B(E)** -0.04	**14:G(E)** -0.09
14:P	0.13	\	0.22	0.25	0.15	0.18	0.22
14:B	-0.08	***	\	0.04	-0.07	-0.04	0.00
14:G	-0.12	***	***	\	-0.11	-0.08	-0.03
14:E	-0.01	**	NS	NS	\	0.03	0.07
14:B(E)	-0.04	***	**	***	NS	\	0.04
14:G(E)	-0.09	***	NS	*	NS	***	\

Spearman Correlations	**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2014 Pheno	\	0.13	-0.08	-0.12	-0.01	-0.04	-0.09
PS	NS	\	0.74	0.72	0.72	0.76	0.73
BayesB	NS	***	\	1.00	0.67	0.98	0.98
G-BLUP	NS	***	***	\	0.67	0.97	0.98
Epistacy	NS	***	***	***	\	0.66	0.65
BayesB(Epi)	NS	***	***	***	***	\	0.99
G-BLUP(Epi)	NS	***	***	***	***	***	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.5 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for palmitic acid in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Palmitic	Correlation Differences	**14:P** 0.94	**14:B** 0.93	**14:G** 0.93	**14:E** 0.80	**14:B(E)** 0.93	**14:G(E)** 0.93
14:P	0.94	\	0.01	0.01	0.14	0.01	0.01
14:B	0.93	NS	\	0.00	0.13	0.00	0.00
14:G	0.93	NS	NS	\	0.13	0.00	0.00
14:E	0.80	***	***	***	\	-0.13	-0.13
14:B(E)	0.93	NS	NS	NS	***	\	0.00
14:G(E)	0.93	NS	NS	NS	***	NS	\

Spearman Correlations	**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2014 Pheno	\	0.94	0.93	0.93	0.80	0.93	0.93
PS	***	\	0.97	0.97	0.84	0.97	0.97
BayesB	***	***	\	1.00	0.86	1.00	1.00
G-BLUP	***	***	***	\	0.86	1.00	1.00
Epistacy	***	***	***	***	\	0.86	0.86
BayesB(Epi)	***	***	***	***	***	\	1.00
G-BLUP(Epi)	***	***	***	***	***	***	\

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.6 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for stearic acid in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Stearic	Correlation Differences	**14:P** 0.88	**14:B** 0.87	**14:G** 0.88	**14:E** 0.74	**14:B(E)** 0.87	**14:G(E)** 0.88
14:P	0.88	\	0.01	0.01	0.14	0.01	0.01
14:B	0.87	NS	\	0.00	0.13	0.00	0.00
14:G	0.88	NS	NS	\	0.14	0.01	0.00
14:E	0.74	***	***	***	\	-0.13	-0.14
14:B(E)	0.87	NS	NS	NS	***	\	-0.01
14:G(E)	0.88	NS	NS	NS	***	NS	\

Spearman Correlations	**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2014 Pheno	\	0.88	0.87	0.88	0.74	0.87	0.88
PS	***	\	0.96	0.98	0.83	0.97	0.98
BayesB	***	***	\	0.99	0.83	1.00	0.99
G-BLUP	***	***	***	\	0.83	0.99	1.00
Epistacy	***	***	***	***	\	0.83	0.83
BayesB(Epi)	***	***	***	***	***	\	0.99
G-BLUP(Epi)	***	***	***	***	***	***	\

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.7 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for oleic acid in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Oleic	Correlation Differences	**14:P** 0.96	**14:B** 0.95	**14:G** 0.95	**14:E** 0.82	**14:B(E)** 0.95	**14:G(E)** 0.95
14:P	0.96	\	0.01	0.01	0.14	0.01	0.01
14:B	0.95	*	\	0.00	0.13	0.00	0.00
14:G	0.95	*	NS	\	0.13	0.00	0.00
14:E	0.82	***	***	***	\	-0.13	-0.13
14:B(E)	0.95	*	NS	NS	***	\	0.00
14:G(E)	0.95	*	NS	NS	***	NS	\

Spearman Correlations	**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2014 Pheno	\	0.96	0.95	0.95	0.82	0.95	0.95
PS	***	\	0.97	0.97	0.83	0.97	0.97
BayesB	***	***	\	1.00	0.85	1.00	1.00
G-BLUP	***	***	***	\	0.85	1.00	1.00
Epistacy	***	***	***	***	\	0.85	0.85
BayesB(Epi)	***	***	***	***	***	\	1.00
G-BLUP(Epi)	***	***	***	***	***	***	\

*Significant at 0.05 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.8 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for linoleic acid in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Linoleic	Correlation Differences	**14:P** 0.96	**14:B** 0.94	**14:G** 0.94	**14:E** 0.82	**14:B(E)** 0.94	**14:G(E)** 0.94
14:P	0.96	\	0.01	0.02	0.14	0.01	0.01
14:B	0.94	***	\	0.00	0.12	0.00	0.00
14:G	0.94	***	NS	\	0.12	0.00	0.00
14:E	0.82	***	***	***	\	-0.12	-0.12
14:B(E)	0.94	**	NS	NS	***	\	0.00
14:G(E)	0.94	**	NS	NS	***	NS	\

Spearman Correlations	**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2014 Pheno	\	0.96	0.94	0.94	0.82	0.94	0.94
PS	***	\	0.98	0.97	0.84	0.97	0.97
BayesB	***	***	\	1.00	0.86	1.00	1.00
G-BLUP	***	***	***	\	0.86	1.00	1.00
Epistacy	***	***	***	***	\	0.86	0.86
BayesB(Epi)	***	***	***	***	***	\	1.00
G-BLUP(Epi)	***	***	***	***	***	***	\

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.9 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for linolenic acid in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Linolenic	Correlation Differences	**14:P** 0.94	**14:B** 0.95	**14:G** 0.95	**14:E** 0.85	**14:B(E)** 0.95	**14:G(E)** 0.95
14:P	0.94	\	-0.01	-0.01	0.09	-0.01	-0.01
14:B	0.95	*	\	0.00	0.11	0.00	0.00
14:G	0.95	*	NS	\	0.11	0.00	0.00
14:E	0.85	***	***	***	\	-0.11	-0.11
14:B(E)	0.95	*	NS	NS	***	\	0.00
14:G(E)	0.95	*	NS	NS	***	NS	\

Spearman Correlations	**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2014 Pheno	\	0.94	0.95	0.95	0.85	0.95	0.95
PS	***	\	0.97	0.97	0.88	0.98	0.98
BayesB	***	***	\	1.00	0.89	1.00	1.00
G-BLUP	***	***	***	\	0.89	1.00	1.00
Epistacy	***	***	***	***	\	0.89	0.89
BayesB(Epi)	***	***	***	***	***	\	1.00
G-BLUP(Epi)	***	***	***	***	***	***	\

*Significant at 0.05 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.10 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for protein in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Protein	Correlation Differences	**14:P** 0.88	**14:B** 0.89	**14:G** 0.89	**14:E** 0.76	**14:B(E)** 0.89	**14:G(E)** 0.89
14:P	0.88	\	-0.01	-0.02	0.12	-0.01	-0.01
14:B	0.89	NS	\	0.00	0.13	0.00	0.00
14:G	0.89	NS	NS	\	0.13	0.00	0.00
14:E	0.76	***	***	***	\	-0.13	-0.13
14:B(E)	0.89	NS	NS	NS	***	\	0.00
14:G(E)	0.89	NS	NS	NS	***	NS	\

Spearman Correlations	**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2014 Pheno	\	0.88	0.89	0.89	0.76	0.89	0.89
PS	***	\	0.95	0.95	0.83	0.96	0.96
BayesB	***	***	\	1.00	0.83	0.99	0.99
G-BLUP	***	***	***	\	0.84	0.99	0.99
Epistacy	***	***	***	***	\	0.83	0.84
BayesB(Epi)	***	***	***	***	***	\	1.00
G-BLUP(Epi)	***	***	***	***	***	***	\

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.11 Comparison of Spearman correlations for 2014 phenotypes with predictions from combined 2010 and 2013 datasets for oil in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The estimates listed are as follows: 14 (2014 phenotype), P (10_13 PS), B (10_13 BayesB), G (10_13 G-BLUP), E (10_13 Epistacy), B(E) (10_13 BayesB(Epi)), and G(E) (10_13 G-BLUP(Epi)), with 10_13 as the combined analysis from the 2010 and 2013 field seasons. Additionally, Spearman correlations between selections methods are displayed, with higher values indicating greater similarity between rankings.

Oil	Correlation Differences	**14:P** 0.88	**14:B** 0.86	**14:G** 0.86	**14:E** 0.69	**14:B(E)** 0.86	**14:G(E)** 0.86
14:P	0.88	\	0.02	0.02	0.19	0.02	0.02
14:B	0.86	NS	\	0.00	0.17	0.00	0.01
14:G	0.86	NS	NS	\	0.17	0.00	0.00
14:E	0.69	***	***	***	\	-0.17	-0.17
14:B(E)	0.86	*	NS	NS	***	\	0.00
14:G(E)	0.86	*	NS	NS	***	NS	\

Spearman Correlations	**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
2014 Pheno	\	0.88	0.86	0.86	0.69	0.86	0.86
PS	***	\	0.94	0.94	0.75	0.94	0.94
BayesB	***	***	\	1.00	0.75	1.00	1.00
G-BLUP	***	***	***	\	0.75	1.00	1.00
Epistacy	***	***	***	***	\	0.75	0.75
BayesB(Epi)	***	***	***	***	***	\	1.00
G-BLUP(Epi)	***	***	***	***	***	***	\

*Significant at 0.05 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.12 Yield contrasts from only the Knoxville, TN location of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Yield kg ha^{-1}		3606.90	2980.74	2907.16	2845.94	2980.92	2905.02	2878.75
2014 Pheno	3606.90	\	626.16	699.74	760.96	625.98	701.88	728.15
PS	2980.74	***	\	73.58	134.80	-0.18	75.72	101.99
BayesB	2907.16	***	*	\	61.22	-73.76	2.14	28.41
G-BLUP	2845.94	***	***	***	\	-134.98	-59.08	-32.81
Epistacy	2980.92	***	NS	*	***	\	75.90	102.17
BayesB(Epi)	2905.02	***	*	NS	***	NS	\	26.27
G-BLUP(Epi)	2878.75	***	**	**	*	**	*	\
15% low		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Yield kg ha^{-1}		2381.00	2831.03	2970.32	3005.85	2950.93	2960.36	2996.68
2014 Pheno	2381.00	\	-450.03	-589.32	-624.85	-569.93	-579.36	-615.68
PS	2831.03	***	\	-139.29	-174.82	-119.90	-129.33	-165.65
BayesB	2970.32	***	**	\	-35.53	19.39	9.96	-26.36
G-BLUP	3005.85	***	***	*	\	54.92	45.49	9.17
Epistacy	2950.93	***	**	NS	NS	\	-9.43	-45.75
BayesB(Epi)	2960.36	***	**	NS	NS	NS	\	-36.32
G-BLUP(Epi)	2996.68	***	***	NS	NS	NS	*	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.13 Palmitic acid contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Palmitic g kg^{-1} seed oil	133.86	133.86	133.06	132.94	132.94	131.39	133.19	133.19
2014 Pheno	133.86	\	0.80	0.92	0.92	2.47	0.67	0.67
PS	133.06	***	\	0.12	0.12	1.67	-0.13	-0.13
BayesB	132.94	***	NS	\	0.00	1.55	-0.25	-0.25
G-BLUP	132.94	***	NS	NS	\	1.55	-0.25	-0.25
Epistacy	131.39	***	***	***	***	\	-1.80	-1.80
BayesB(Epi)	133.19	**	NS	NS	NS	***	\	0.00
G-BLUP(Epi)	133.19	**	NS	NS	NS	***	NS	\
15% low		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Palmitic g kg^{-1} seed oil	112.20	112.20	112.42	112.76	112.76	114.72	112.53	112.59
2014 Pheno	112.20	\	-0.22	-0.56	-0.56	-2.52	-0.33	-0.39
PS	112.42	NS	\	-0.34	-0.34	-2.30	-0.11	-0.17
BayesB	112.76	**	NS	\	0.00	-1.96	0.23	0.17
G-BLUP	112.76	**	NS	NS	\	-1.96	0.23	0.17
Epistacy	114.72	***	***	***	***	\	2.19	2.13
BayesB(Epi)	112.53	NS	NS	NS	NS	***	\	-0.06
G-BLUP(Epi)	112.59	*	NS	NS	NS	***	NS	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.14 Stearic acid contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Stearic g kg^{-1} seed oil		46.56	45.68	45.49	45.69	44.86	45.39	45.65
2014 Pheno	46.56	\	0.88	1.07	0.87	1.70	1.17	0.91
PS	45.68	***	\	0.19	-0.01	0.82	0.29	0.03
BayesB	45.49	***	NS	\	-0.20	0.63	0.10	-0.16
G-BLUP	45.69	***	NS	*	\	0.83	0.30	0.04
Epistacy	44.86	***	***	**	***	\	-0.53	-0.79
BayesB(Epi)	45.39	***	*	NS	**	*	\	-0.26
G-BLUP(Epi)	45.65	***	NS	NS	NS	***	**	\
15% low		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Stearic g kg^{-1} seed oil		37.42	38.18	38.03	38.03	38.73	38.15	38.03
2014 Pheno	37.42	\	-0.76	-0.61	-0.62	-1.31	-0.73	-0.62
PS	38.18	***	\	0.15	0.14	-0.55	0.03	0.14
BayesB	38.03	**	NS	\	-0.01	-0.70	-0.12	-0.01
G-BLUP	38.03	**	NS	NS	\	-0.70	-0.11	0.00
Epistacy	38.73	***	*	**	**	\	0.59	0.70
BayesB(Epi)	38.15	***	NS	*	NS	**	\	0.11
G-BLUP(Epi)	38.03	**	NS	NS	NS	**	NS	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.15 Oleic acid contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Oleic g kg^{-1} seed oil	400.66	400.66	396.33	396.52	396.25	387.82	395.88	396.20
2014 Pheno	400.66	\	4.33	4.14	4.41	12.84	4.78	4.46
PS	396.33	**	\	-0.19	0.08	8.51	0.45	0.13
BayesB	396.52	**	NS	\	0.27	8.70	0.64	0.32
G-BLUP	396.25	**	NS	NS	\	8.43	0.37	0.05
Epistacy	387.82	***	***	***	***	\	-8.06	-8.38
BayesB(Epi)	395.88	**	NS	NS	NS	***	\	-0.32
G-BLUP(Epi)	396.20	**	NS	NS	NS	***	NS	\
15% low		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Oleic g kg^{-1} seed oil	291.32	291.32	294.34	295.78	295.78	301.68	295.57	295.57
2014 Pheno	291.32	\	-3.02	-4.46	-4.46	-10.36	-4.25	-4.25
PS	294.34	NS	\	-1.44	-1.44	-7.34	-1.23	-1.23
BayesB	295.78	*	NS	\	0.00	-5.90	0.21	0.21
G-BLUP	295.78	*	NS	NS	\	-5.90	0.21	0.21
Epistacy	301.68	***	**	**	**	\	6.11	6.11
BayesB(Epi)	295.57	*	NS	NS	NS	**	\	0.00
G-BLUP(Epi)	295.57	*	NS	NS	NS	**	NS	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.16 Linoleic acid contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Linoleic g kg^{-1} seed oil	456.21	456.21	454.45	453.26	452.50	448.22	453.39	452.63
2014 Pheno	456.21	\	1.76	2.95	3.71	7.99	2.82	3.58
PS	454.45	NS	\	1.19	1.95	6.23	1.06	1.82
BayesB	453.26	*	NS	\	0.76	5.04	-0.13	0.63
G-BLUP	452.50	**	NS	*	\	4.28	-0.89	-0.13
Epistacy	448.22	***	***	***	**	\	-5.17	-4.41
BayesB(Epi)	453.39	*	NS	NS	NS	***	\	0.76
G-BLUP(Epi)	452.63	**	NS	NS	NS	**	*	\
15% low		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Linoleic g kg^{-1} seed oil	380.47	380.47	384.07	383.60	383.58	386.45	383.60	383.58
2014 Pheno	380.47	\	-3.60	-3.13	-3.11	-5.98	-3.13	-3.11
PS	384.07	**	\	0.47	0.49	-2.38	0.47	0.49
BayesB	383.60	**	NS	\	0.02	-2.85	0.00	0.02
G-BLUP	383.58	**	NS	NS	\	-2.87	-0.02	0.00
Epistacy	386.45	***	NS	*	*	\	2.85	2.87
BayesB(Epi)	383.60	**	NS	NS	NS	*	\	0.02
G-BLUP(Epi)	383.58	**	NS	NS	NS	*	NS	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.17 Linolenic acid contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Linolenic g kg^{-1} seed oil	84.84	84.03	83.66	83.82	82.62	83.95	83.90	
2014 Pheno	84.84	\	0.80	1.18	1.01	2.22	0.89	0.94
PS	84.03	**	\	0.37	0.21	1.42	0.09	0.13
BayesB	83.66	***	NS	\	-0.16	1.04	-0.29	-0.24
G-BLUP	83.82	***	NS	NS	\	1.21	-0.12	-0.07
Epistacy	82.62	***	***	**	***	\	-1.33	-1.28
BayesB(Epi)	83.95	**	NS	*	NS	***	\	0.05
G-BLUP(Epi)	83.90	**	NS	NS	NS	***	NS	\
15% low		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Linolenic g kg^{-1} seed oil	60.85	61.55	61.32	61.32	63.06	61.32	61.32	
2014 Pheno	60.85	\	-0.70	-0.47	-0.47	-2.20	-0.47	-0.47
PS	61.55	*	\	0.23	0.23	-1.50	0.23	0.23
BayesB	61.32	*	NS	\	0.00	-1.74	0.00	0.00
G-BLUP	61.32	*	NS	NS	\	-1.74	0.00	0.00
Epistacy	63.06	***	***	***	***	\	1.74	1.74
BayesB(Epi)	61.32	*	NS	NS	NS	***	\	0.00
G-BLUP(Epi)	61.32	*	NS	NS	NS	***	NS	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.18 Protein contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Protein g kg^{-1} seed dw	449.67	449.03	448.15	448.15	445.65	447.88	447.97	
2014 Pheno	449.67	\	1.64	1.52	1.52	4.02	1.79	1.70
PS	448.03	**	\	-0.12	-0.12	2.38	0.15	0.06
BayesB	448.15	**	NS	\	0.00	2.50	0.27	0.18
G-BLUP	448.15	**	NS	NS	\	2.50	0.27	0.18
Epistacy	445.65	***	***	***	***	\	-2.23	-2.32
BayesB(Epi)	447.88	**	NS	NS	NS	**	\	-0.09
G-BLUP(Epi)	447.97	**	NS	NS	NS	***	NS	\
15% low		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Protein g kg^{-1} seed dw	416.10	418.31	418.46	418.54	420.88	418.79	418.79	
2014 Pheno	416.10	\	-2.21	-2.36	-2.44	-4.78	-2.69	-2.69
PS	418.31	***	\	-0.15	-0.23	-2.57	-0.48	-0.48
BayesB	418.46	***	NS	\	-0.08	-2.42	-0.33	-0.33
G-BLUP	418.54	***	NS	NS	\	-2.34	-0.25	-0.25
Epistacy	420.88	***	***	***	***	\	2.09	2.09
BayesB(Epi)	418.79	***	NS	NS	NS	**	\	0.00
G-BLUP(Epi)	418.79	***	NS	NS	NS	**	NS	\

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 3.19 Oil contrasts of 15% (30 RILs) high and low tail selections from the 2010 and 2013 combined analysis (10_13) predictions as measured in 2014 from soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 2014 high and low tail selections are included for comparison. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods.

15% high		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Oil g kg^{-1} seed dw		220.79	220.04	219.75	219.75	219.11	219.75	219.75
2014 Pheno	220.79	\	0.75	1.04	1.04	1.68	1.04	1.04
PS	220.04	**	\	0.29	0.29	0.93	0.29	0.29
BayesB	219.75	***	NS	\	0.00	0.64	0.00	0.00
G-BLUP	219.75	***	NS	NS	\	0.64	0.00	0.00
Epistacy	219.11	***	**	*	*	\	-0.64	-0.64
BayesB(Epi)	219.75	***	NS	NS	NS	*	\	0.00
G-BLUP(Epi)	219.75	***	NS	NS	NS	*	NS	\
15% low		**2014 Pheno**	**PS**	**BayesB**	**G-BLUP**	**Epistacy**	**BayesB(Epi)**	**G-BLUP(Epi)**
Oil g kg^{-1} seed dw		203.78	204.37	204.29	204.36	205.72	204.38	204.38
2014 Pheno	203.78	\	-0.59	-0.51	-0.58	-1.94	-0.60	-0.60
PS	204.37	*	\	0.08	0.01	-1.35	-0.01	-0.01
BayesB	204.29	*	NS	\	-0.07	-1.43	-0.09	-0.09
G-BLUP	204.36	*	NS	NS	\	-1.36	-0.02	-0.02
Epistacy	205.72	***	***	***	***	\	1.34	1.34
BayesB(Epi)	204.38	*	NS	NS	NS	***	\	0.00
G-BLUP(Epi)	204.38	*	NS	NS	NS	***	NS	\

*Significant at 0.05 probability level

**Significant at 0.01 probability level

***Significant at 0.001 probability level

NS, Not significant at 0.05 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

Figures

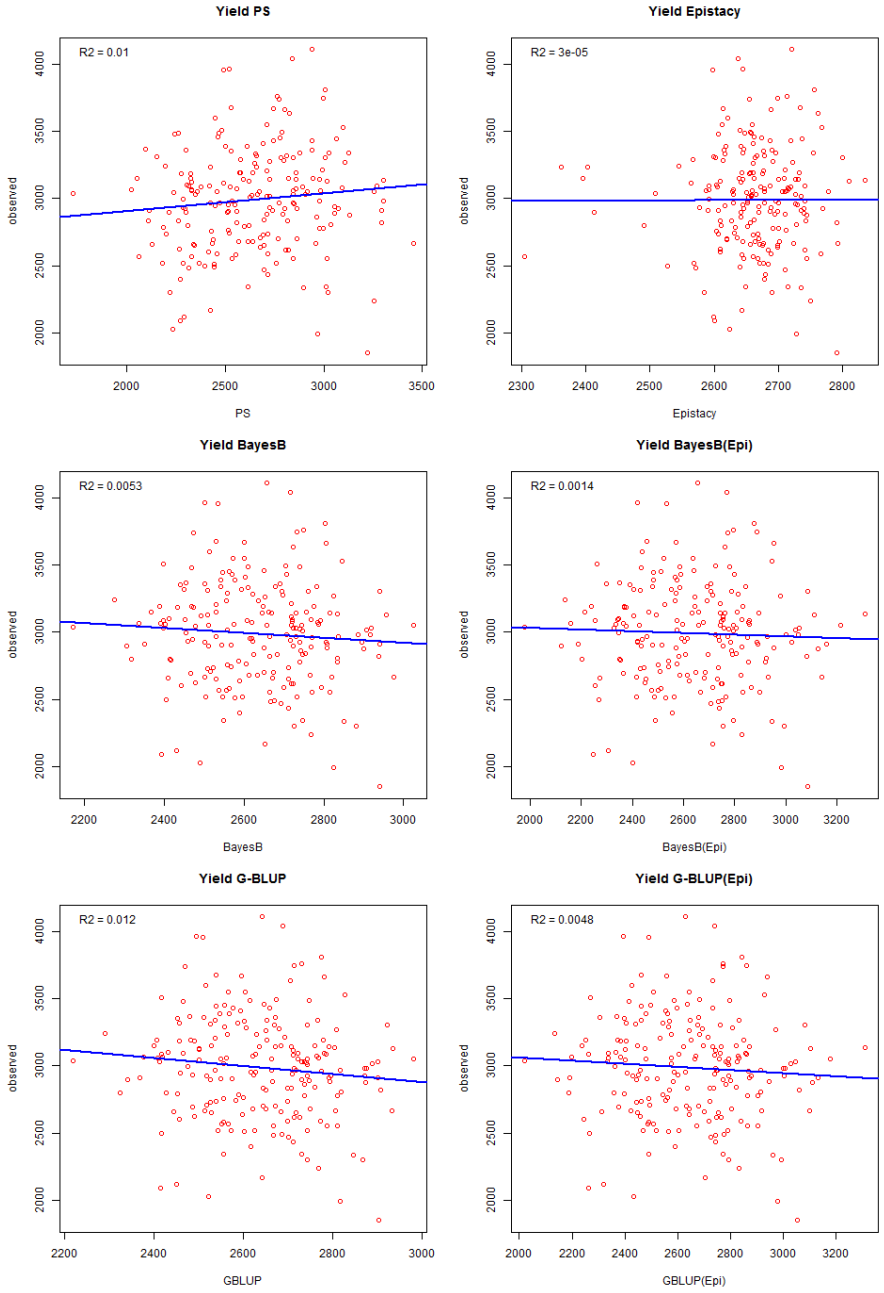

Figure 3.1 Yield (kg ha^{-1}) performance comparisons from only the Knoxville, TN location between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes.

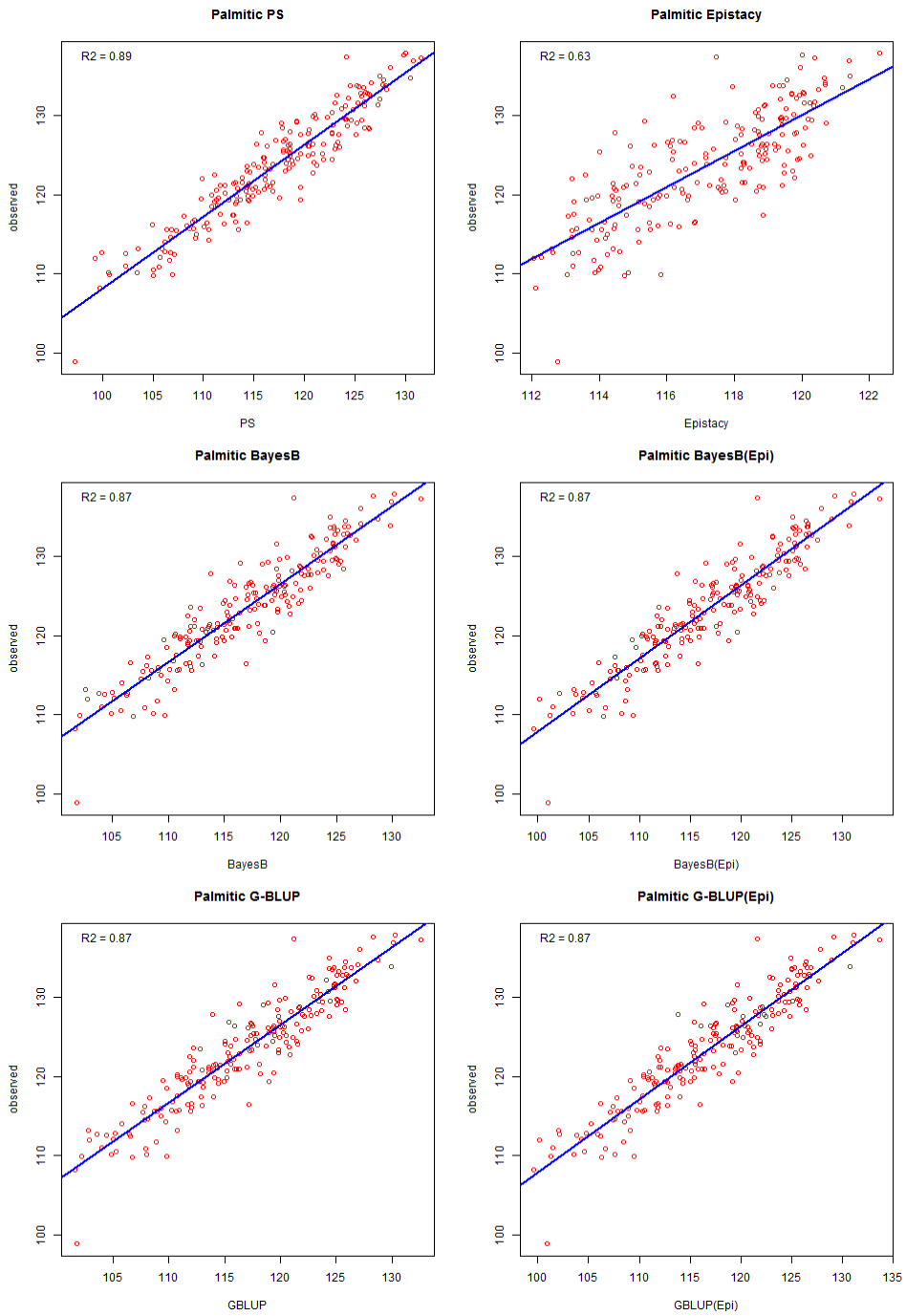

Figure 3.2 Palmitic acid (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes.

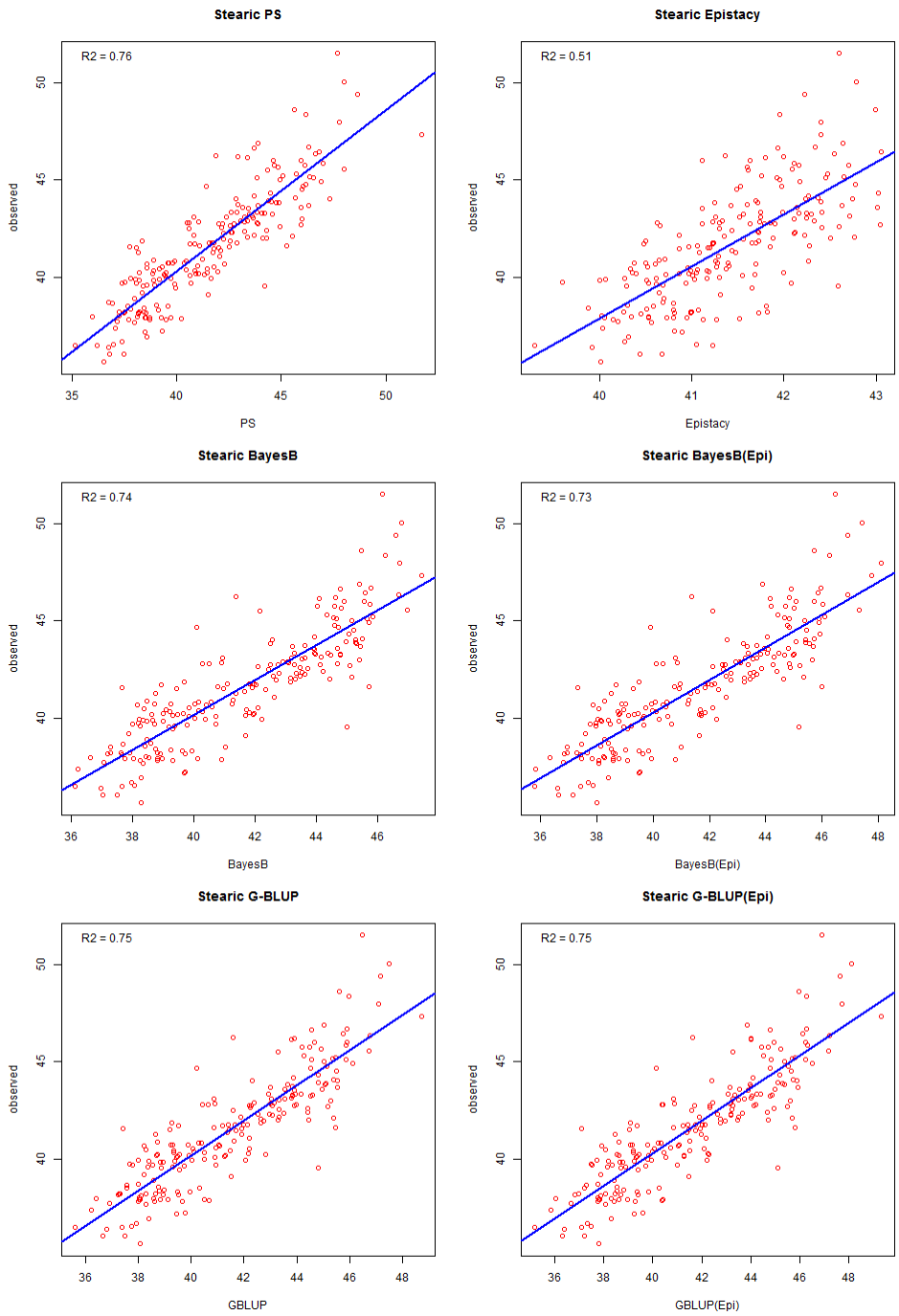

Figure 3.3 Stearic acid (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes.

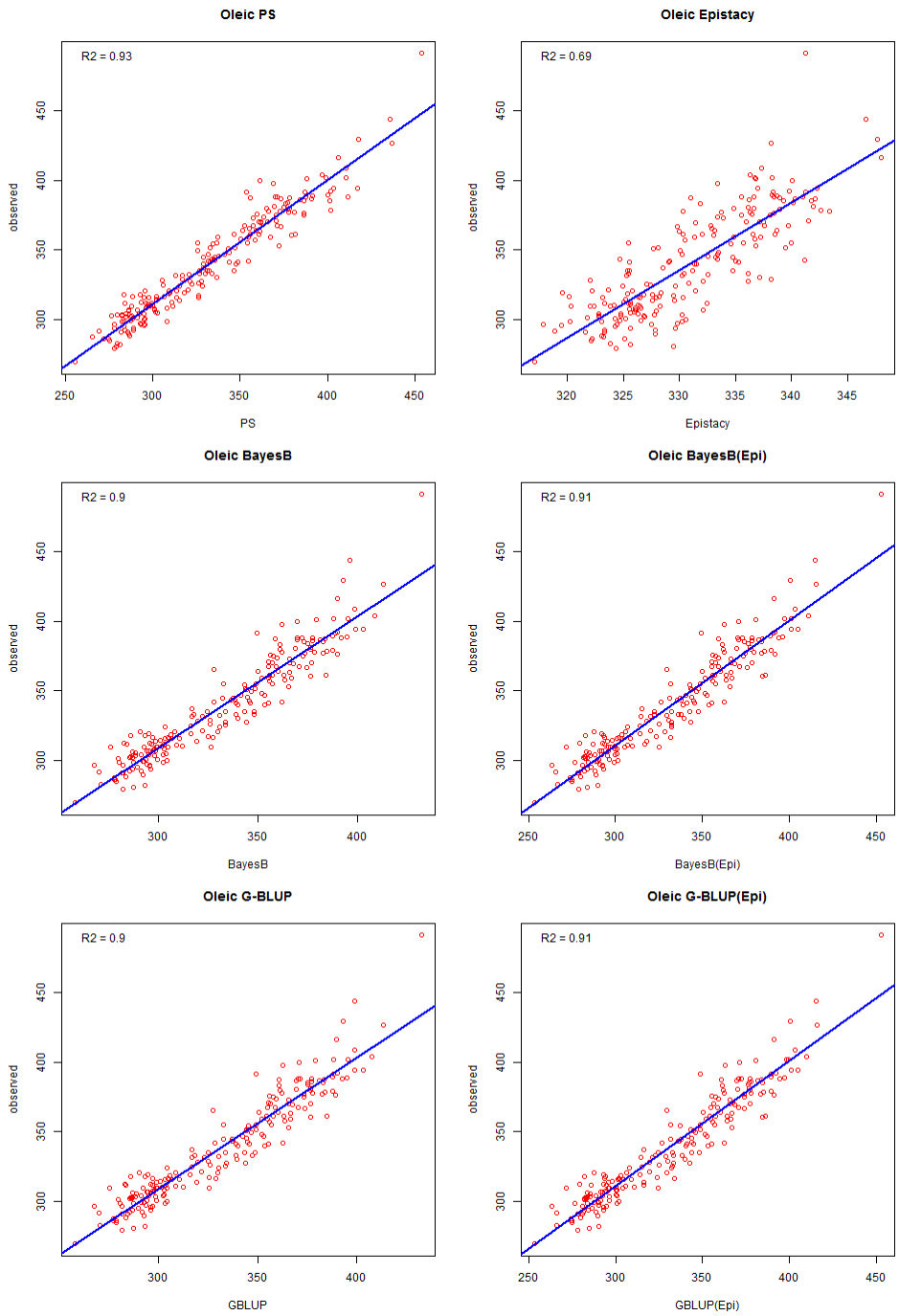

Figure 3.4 Oeic acid (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes.

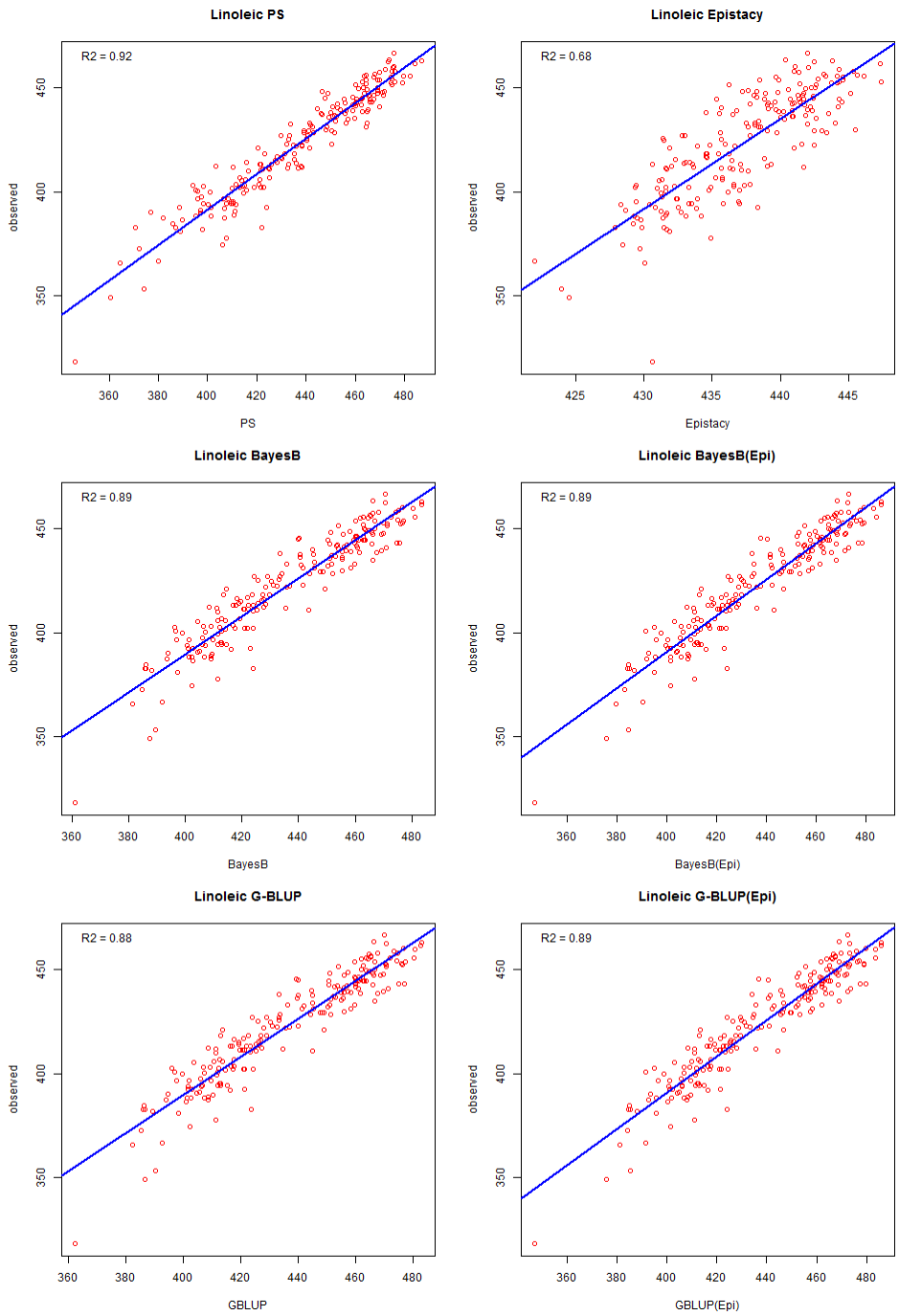

Figure 3.5 Linoleic acid (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes.

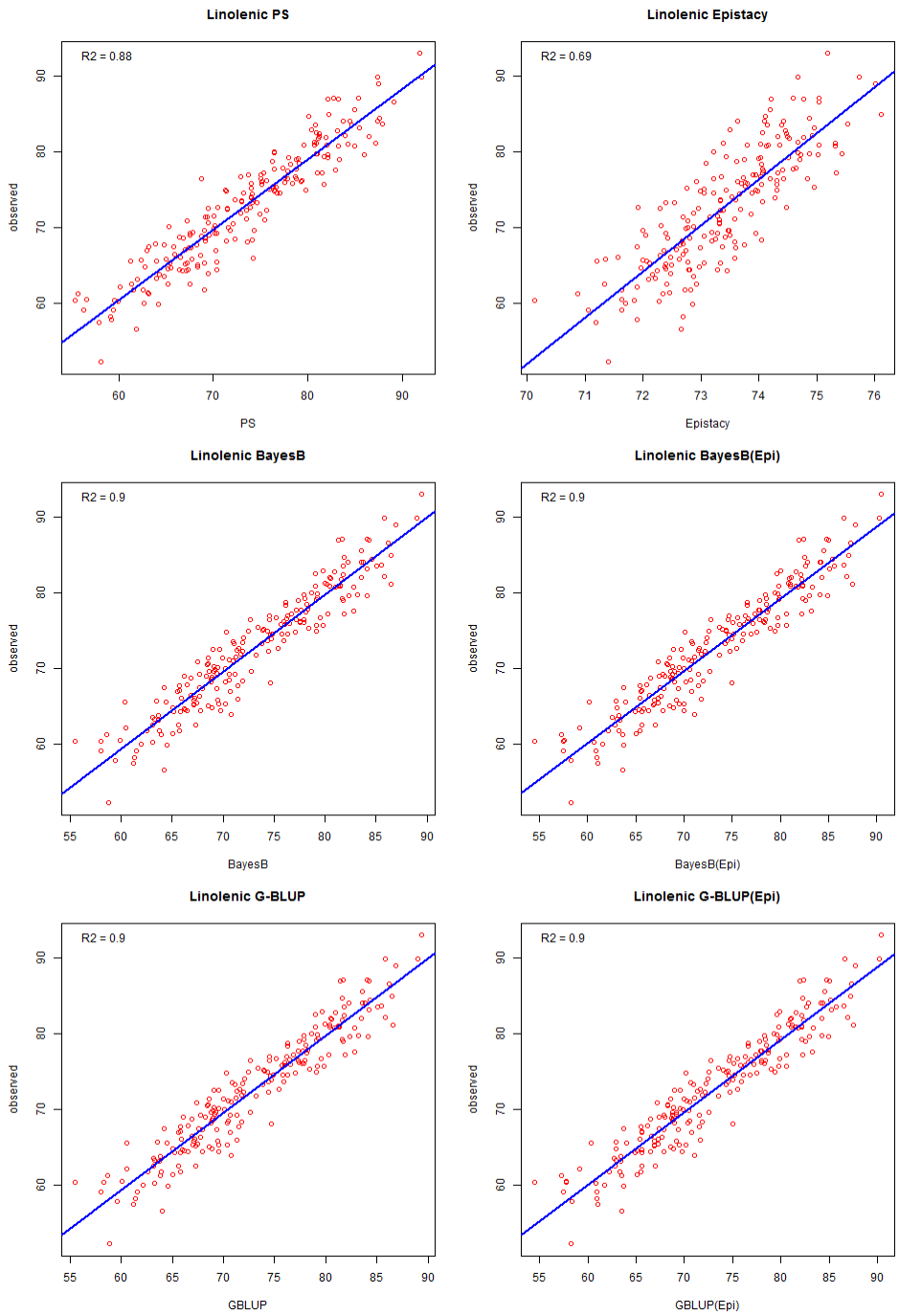

Figure 3.6 Linolenic acid (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes.

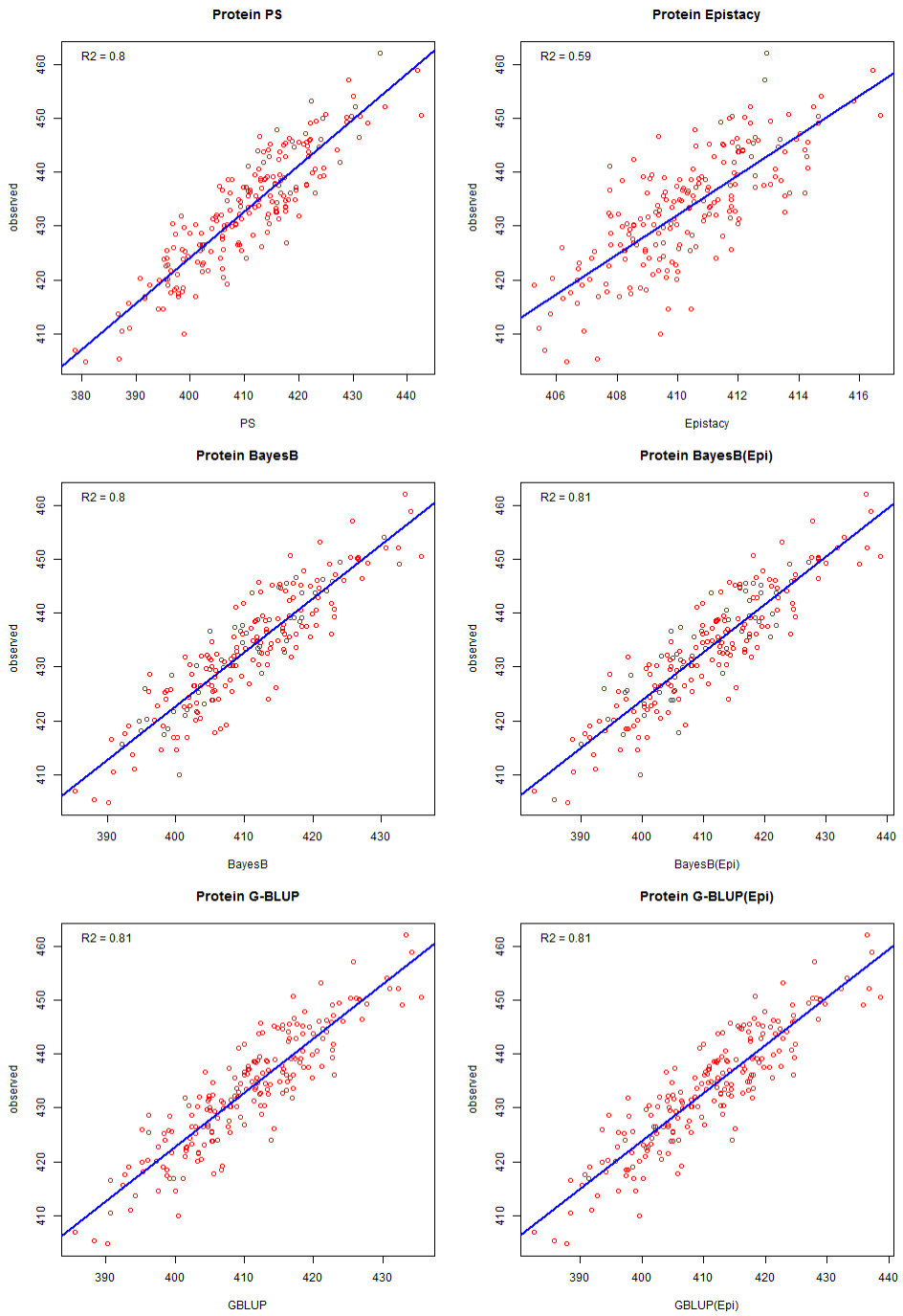

Figure 3.7 Protein (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes.

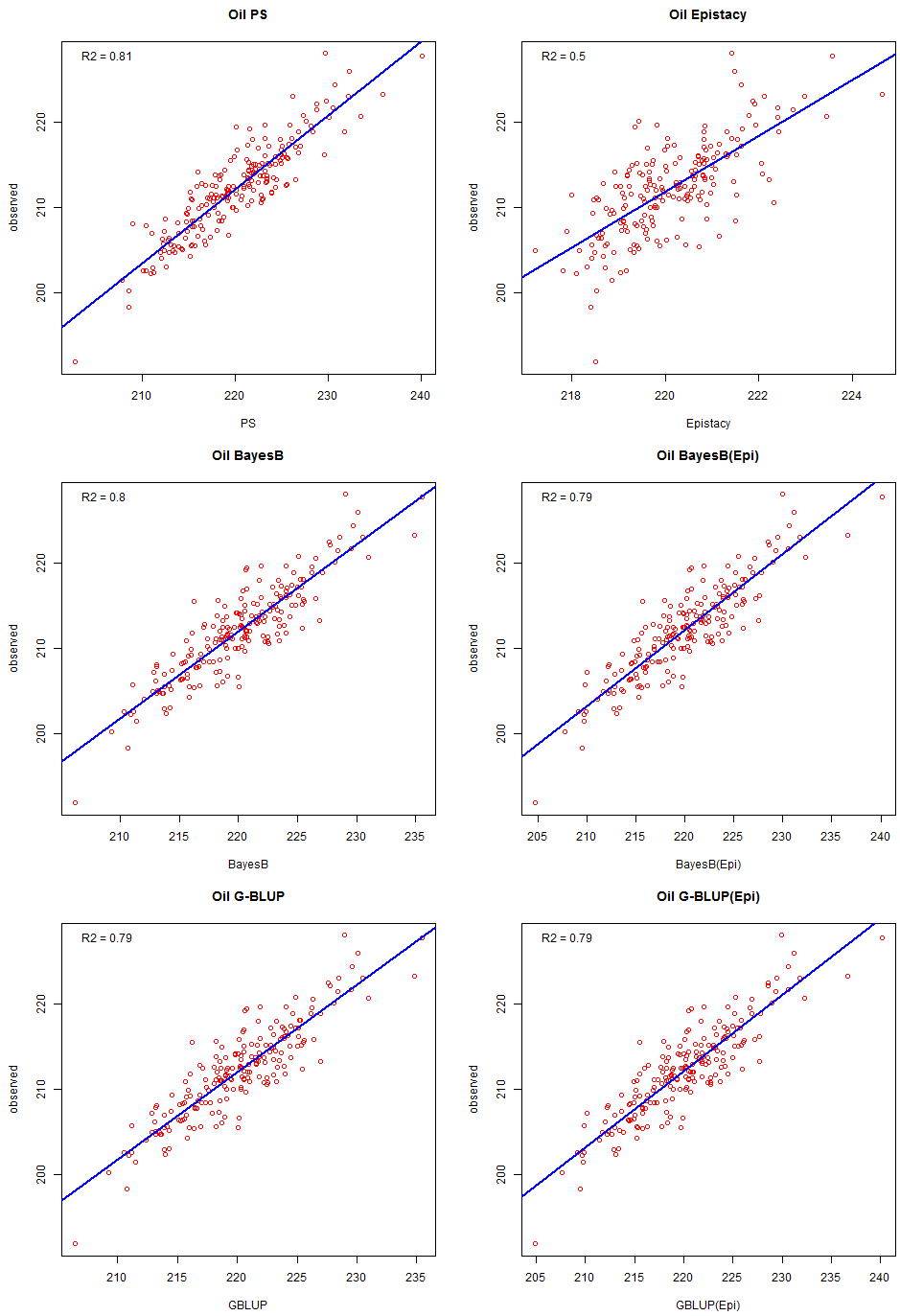

Figure 3.8 Oil (g kg^{-1}) performance comparisons between the 2010 and 2013 combined analysis (10_13) predictions (X-axis) and 2014 phenotypes (Y-axis) in soybean population E×W-50K subset consisting of 203 F5 derived RILs. The 10_13 predictions were estimated with phenotypic (PS) and molecular (Epistacy, BayesB, BayesB(Epi), G-BLUP, and G-BLUP(Epi)) selection methods. Predictions with higher R2 were more closely related to 2014 observed phenotypes.

CHAPTER 4

IDENTIFYING AND EXPLORING SIGNIFICANT GENOMIC REGIONS FOR SOYBEAN YIELD, FATTY ACIDS, PROTEIN, AND OIL

Abstract

Soybean [Glycine max (L.) Merrill] yield, fatty acids, protein, and oil are commercially important traits. Due to this, it is critical to develop an improved understanding of significant genomic regions governing these traits. To accomplish this, a soybean population consisting of 203 F5 derived recombinant inbred lines (RILs) was developed and genotyped with 17,236 polymorphic SNPs using the SoySNP50K BeadChip. Each RIL was grown in a single plot at Knoxville, TN in 2010; followed by replicated, multi-location field trials in 2013 and 2014. The phenotypic data from 2010, 2013, and 2014 were analyzed together and combined with the genotypic data in order to detect quantitative trait loci (QTL) influencing these traits. A total of 29 QTLs were detected [yield (1), palmitic acid (6), stearic acid (1), oleic acid (4), linoleic acid (3), linolenic acid (5), protein (7), and oil (2)]. Of these, the QTL for stearic acid stands out as the most influential (R^2 = 0.52). Additionally, the stearic QTL, along with two protein QTLs are excellent candidates for confirmed status; while four QTLs (two each for linolenic acid and protein) are strong candidates for positional confirmations. Many of the genes with amino acid changes in close proximity to the fatty acid QTLs are involved in biological processes for fatty acids and/or lipids; thus giving them high potential as candidate genes. Similarly, genes with amino acid changes in genomic regions near yield, protein, and oil were plentiful and may contribute to some of the variation observed in these traits. All of the traits except yield displayed pleiotropic effects with other traits in this study. Overall, the findings from this research contribute new information to the genetic understanding of soybean yield, fatty acids, protein and oil. This understanding will be useful in making trait improvements.

Introduction

Soybean [Glycine max (L.) Merrill] is a prominent crop grown throughout much of the world for many purposes. Seed protein (~400 g kg^{-1}) and oil (~200 g kg^{-1}) are primary components of soybean that contribute to its high value. These traits are

common targets for research seeking to improve the value of soybean. Additionally, with the Food and Drug Administration (FDA) recently banning partially hydrogenated oils (PHOs) in all food products (https://www.federalregister.gov/articles/2015/06/17/2015-14883/final-determination-regarding-partially-hydrogenated-oils), improving the fatty acid profile of soybean has become a major objective.

The five primary fatty acids in soybean are palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2) and linolenic (18:3); typically occurring in relative concentrations of 100, 40, 220, 540, and 100 g kg^{-1} of total lipids, respectively (Wilson, 2004). Due to the FDA ban on PHOs, a major initiative in fatty acid improvement is to reduce linolenic acid (< 30 g kg^{-1}), thereby reducing the need for hydrogenation, a process that creates partially hydrogenated soybean oil. Increasing the monounsaturated oleic acid (> 800 g kg^{-1}) is another major goal, as it has been shown to lower cholesterol in comparison with saturated fatty acids when consumed by humans (Kris-Etherton and Yu, 1997). Further, the oxidative stability of soybean oil is improved by increasing the concentration of oleic acid; leading to increased shelf life for soybean oil food products (Kinney, 1996) and biodiesel (Kinney and Clemente, 2005; Fallen et al, 2012).

Many recent studies have focused on improving soybean fatty acids (Pantalone et al., 2002; Pham et al., 2010; Bilyeu et al., 2011; Boersma et al., 2012; Gillman et al., 2014). However, there is still a need for further improvement. For stearic acid, competitively yielding breeding lines have yet to achieve the targeted goal of (> 200 g kg^{-1}) of total oil (Gillman et al., 2014). For oleic acid there is concern that environmental variation may result in levels that are below 800 g kg^{-1} (Lee et al., 2012; Fallen et al., 2012). Breeding with small effect modifier QTLs for fatty acids in addition to major QTLs may help achieve greater stability (Hyten et al., 2004b). Additionally, determining the genetic origins and pleiotropic effects of such QTLs would be useful (Cardinal et al., 2014). Thus, the objectives of this

research were to identify QTLs influencing soybean yield, fatty acids, protein, and oil; and to examine these QTLs for pleiotropic effects and candidate genes.

Materials and Methods

Plant Materials

A population of 860 recombinant inbred lines (RILs) with both genotypic and phenotypic data derived from parental lines 'Essex' and 'Williams 82' (hereafter known as E×W-50K) was created for QTL detection. Essex is a southern cultivar in maturity group (MG) V with a determinate growth habit, purple flower, and gray pubescence (Smith and Camper, 1973); while Williams 82 is a northern cultivar in MG III with indeterminate growth habit, white flower, and tawny pubescence (Bernard and Cremeens, 1988). Seed for both Essex and Williams 82 was obtained from the USDA soybean germplasm collection (www.ars-grin.gov). In order to provide highly homozygous parental lines to be crossed for RIL development, a random single plant of each parental line was intentionally selfed for two generations. The development and field testing of this population are described in more detail in Chapters 2 and 3, respectively.

Briefly, each of the 860 F5 derived RILs developed through single seed descent (Brim, 1966) were grown in single replicate plots in 2010 in Knoxville, TN. A subset of this population (276 RILs) ranging in maturity from MG IV-mid to IV-late was chosen for advancement into replicated field trials planted in 2013. The field test design was a randomized complete block design (RCBD) with three replications per environment at three environments (Knoxville, TN; Springfield, TN; and Milan, TN), representative of the eco-geographic regions of East, Middle, and West Tennessee, respectively. Using data from the combined 2010 and 2013 growing seasons, a new subset of 203 RIL ranging in maturity from MG IV-mid to IV-late was chosen for advancement into replicated field trials planted in 2014. The field design and locations were consistent from 2013 to 2014. For

each of the three field seasons, flower color was determined at the R2 growth stage; and pubescence color, plant height, and maturity were determined at the R8 growth stage (Fehr and Caviness, 1977). For each field season, plots were harvested at maturity. Yield was measured in kg ha^{-1} after adjusting the plot weight to 13% moisture.

Seed Quality Trait Detection

Fatty acid measurements for palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid were performed with the same techniques described in Chapter 2. Briefly, seeds from each plot from the 2010, 2013, and 2014 field tests were analyzed by gas chromatography following the procedure described by Spencer et al. (2004). A Hewlett Packard HP 6890 series gas chromatograph (Agilent Technologies, Santa Clara, CA) system equipped with a 7683 auto sampler, a 7673 flame ionization detector, and an immobilized 30 m x 0.53 mm inner diameter Agilent DB-23 capillary column with 0.5 µm fused stationery phase was used for analysis. Fatty acids were initially estimated as a percentage of seed oil, but were converted to g kg^{-1} seed oil.

To obtain protein and oil estimates, near infrared reflectance spectroscopy (NIRS) was performed in the same manner as described in chapter 2. Briefly, approximately 25 g seed samples from each plot in 2010 were uniformly ground for 20 sec in a Knifetec 1095 Sample Mill (FOSS Tecator, Hoganas, Sweden). The NIRS instrument (NIR 6500, FOSS North America) was used for analysis as described by Panthee et al. (2006a), except that samples in this study were scanned using updated ISIscan software v. 2.85. Each plot from the 2013 season was scanned as a whole bean sample using a Perten DA 7200 Diode Array (Perten, Hägersten, Sweden) NIRS instrument in collaboration with the University of Minnesota, with calibration equations developed through a cooperative effort between Perten and University of Minnesota (Bolon et al., 2011). Plots from the 2014 growing season were scanned with the same

procedure as the 2013 plots with a Perten DA 7250 Diode Array NIRS instrument. Values for protein and oil concentration were adjusted to g kg^{-1} of total seed on a dry weight basis for each NIRS analysis.

SNP Genotyping and Marker Cleanup

Chapter 2 provides a more detailed description of the SNP genotyping and marker cleanup. Briefly, DNA samples were collected from leaf tissue for each RIL in 2009 and analyzed with the Illumina Infinium beadchip SoySNP50K (Song et al., 2013) at the Soybean Genomics Laboratory at the USDA Beltsville Agricultural Research Center (USDA-ARS) in Beltsville, MD. Marker positions from this analysis were obtained from the Williams 82 whole genome sequence (WGS) (Glyma 1.01) (Schmutz et al., 2010). A total of 17,236 polymorphic markers were detected in this population. After dropping SNPs with missing data in ≥ 5% (43 RILs) of the population using the 'synbreed' package (Wimmer et al., 2012) in the R language and environment for statistical computing (R Core Team, 2015), 16,718 polymorphic SNPs remained for analysis. Imputations using default settings in the Beagle Genetic Analysis Software Package v. 3.3.1 (Browning and Browning 2007; 2009) via the 'synbreed' package (Wimmer et al., 2012) in R were used to address any remaining missing marker data. Finally, the 'calc.errorlod' function within the 'qtl' package (Broman et al., 2003) in R was performed in order to screen for potential genotyping errors, with none being detected.

Statistical Analysis and QTL Detection

An RCBD model was performed in SAS PROC GLIMMIX (SAS Institute Inc., Cary, NC, USA, SAS 9.4, 2002-2012) in order to estimate least squares means (LSMEANS) for the 203 RILs from the combined 2010, 2013, and 2014 datasets. The different years for the analysis were treated as separate environments, resulting in seven total environments and 19 total reps. The fixed term for the model was RIL; with environment, rep(environment), and RIL × environment as

random terms, and denominator degrees of freedom method set to residual. Additional statistical analyses were run using the terms from the model above, but with some adjustments to fixed and random terms. A model with RIL, environment, and RIL × environment as fixed terms and rep(environment) as random was run in order to test for the significance of RIL × environment interactions. Also, a model with no fixed terms and RIL, environment, rep(environment), and RIL × environment as random terms was run for all analyses in order to obtain the variance for each term. These variances were then used to estimate heritability on an entry means basis (Nyquist, 1991).

The LSMEANS from the 203 RILs for yield, fatty acids, protein and oil were combined with the 16,718 polymorphic SNPs and used to detect quantitative trait loci (QTL) for these traits with the 'qtl' package (Broman et al., 2003) in the R language and environment for statistical computing (R Core Team, 2015). Rather than estimate a linkage map based on the recombination within this population, the Glyma 1.01 (Schmutz et al., 2010) SNP positions were used; with base pair position scaled down to $2e^{-6}$ as approximated cenitMorgan (cM) position. Since this population segregates at the E1 (maturity), E3 (maturity), and Dt1 (growth habit) loci, SNPs located within or adjacent to (< 5 kb) E1 (ss715593840) and E3 (ss715635705); and field calls on growth habit for Dt1 were used to predict the expected parental allele. Interval mapping (IM) was the primary method of detection so that the Dt1 locus could be included as an additive covariate. The E1 and E3 loci were included as covariates in preliminary analyses, but dropped due to insignificant ($P > 0.05$) contributions to QTL models. A possible explanation for this insignificance is the continued narrowing of the population through maturity selection, leading to a greater uniformity at the E1 and E3 loci among the 203 remaining RILs.

In addition to IM, QTL estimates with composite interval mapping (CIM) were used as a safe guard against 'ghost' QTL (Martinez and Curnow, 1992). The

CIM analysis was not used as the primary method because it did not include an option to add the Dt1 locus as a covariate in the ‘qtl’ package (Broman, 2003). For both IM and CIM, step size was one cM and method was Haley-Knott regression; otherwise default settings were used. In the IM procedure, 10,000 permutations were performed for each trait in order to determine significance thresholds of 1% and 5% (Churchill and Doerge, 1994). The closest SNP to each QTL from the IM procedure was used for ‘makeqtl’ (Broman, 2003) assignments. Then the R squared value and estimated effect for each QTL were determined with ‘fitqtl’ (Broman, 2003) with the following adjustments: the Dt1 locus was included as a covariate and the method was set to Haley-Knott regression.

DNA Sequencing and Candidate Gene Search

DNA Essex and Williams 82 was isolated from ~40 mg of lyophilized seedling leaf tissue using a DNAeasy plant mini kit, according to manufacturer’s recommendations (Qiagen). Indexed Illumina libraries were prepared at Global Biologics, LLC (Columbia, MO) and libraries were sequenced using a HiSeq 2000 instrument at the DNAcore facility at the Univesrity of Missouri (2x100 bps).

All resequencing analyses used the Wm82.a2.v1 genome sequence and annotation build, which was downloaded from (http://phytozome.jgi.doe.gov/). Read mapping and variant calling were done using CLC Genomics Workbench software Version 8.0 (Qiagen/CLC Biotech, Cambridge, MA). Read mapping used the following settings: insertion cost of 3, deletion cost of 3, a similarity fraction setting of 0.8, automatic detection of paired-end distances, and non-specific matching handling was set to ignore. For variant calling, the fixed ploidy program was used with settings: minimum coverage of 9, variant probability cutoff of 90.0, require variant count of 2, and non-specific matches were ignored. Total variant calls were filtered to call genes with amino acid changes.

A candidate gene search for major genes influencing soybean fatty acids in close proximity with QTLs detected in this study was performed using selected genes listed in Gillman and Bilyeu (2012). Given relatively normal range of the fatty acids in this population (Table 4.1) (note: all tables and figures in this chapter are located in Appendix D) when compared with other studies (Pantalone et al., 2002; Pham et al., 2010; Bilyeu et al., 2011; Boersma et al., 2012; Gillman et al., 2014), it seems reasonable to assume that fatty acid QTLs from this study would be for small effect modifier genes rather than large effect genes. Thus, a candidate search for genes with amino acid changes between the parent lines was performed for these small effect fatty acid QTLs; as well as for yield, protein, and oil QTLs. Genes with parental amino acid differences located within one mega base-pair (mb) of QTLs in this study based on Glyma 2.0 positions were processed to determine which gene ontologies (GO) for biological processes were associated (www.SoyBase.org, "Gene Annotation Lookup", accessed 7/25/2015). As the QTL on chromosome 14 for stearic acid accounted for approximately 52% of the variation for that trait (Table 4.2), representing an unusually strong QTL, an expanded search (+/- 5 mb) for biological process GOs associated with that QTL was performed.

Results and Discussion

The approximate sequence coverage with respect to the Glyma 2.0 reference genome was 93% (~16.6 Giga base-pairs) for Essex and 94% (~15.6 Giga base-pairs) for Williams 82, with an average depth of 15.76-fold and 14.84-fold, respectively. Excluding genomic regions with no coverage, the average depth increases to 16.92-fold for Essex and 15.73-fold for Williams 82. In comparison with the Glyma 2.0 reference genome, Essex had approximately 831,000 total nucleotide differences and 14,000 amino acid changes. These numbers are greatly reduced for Williams 82, with approximately 69,000 total nucleotide differences and 1,800 amino acid changes with respect to the Glyma 2.0 reference genome. The nucleotide and amino acid differences between this

strain of Williams 82 and the one used as the Glyma 2.0 reference genome can likely be accounted for in the different strains used, particularly with the additional generations of selfing for increased homozygosity in the Williams 82 strain used in this study. The amino acid differences (> 12,000) between parents in this study are indicative of the gene pools for origination; with Essex in the southern pool and Williams 82 in the northern pool. Further, these amino acid differences provide plentiful opportunities for detecting influential genomic regions for yield, fatty acids, protein, and oil.

Each of the traits studied for QTLs (yield, fatty acids, protein, and oil) displayed a significant difference ($P < 0.05$) among RILs in the data analysis from the combined 2010, 2013, and 2014 field seasons (Table 4.1). Additionally, transgressive segregation was observed for each of these traits, with the range among RILs extending beyond the parental range (Table 4.1). This finding indicates that both parents carry different heritable alleles which influence yield, fatty acids, protein, and oil. Finding loci governing these heritable alleles could have merit for researchers seeking trait improvement. Thus, QTL detection was performed, with the findings displayed in Table 4.2. Additionally, genes with amino acid changes near (< 1 mb) the detected QTLs are listed for chromosomes 4, 6, 7, 9, 11, 13, 14, 17, and 19 (Tables 4.3-4.11, respectively) as possible candidates influencing trait variation.

Only one QTL, located on chromosome 14, was detected for yield (Figure 4.1). This QTL, designated as Yld14, was highly significant ($P < 0.01$), with a LOD score of 13, an R squared value of 0.23, and an effect of 144.43 kg ha^{-1} (Table 4.2). None of the previously reported markers associated with QTLs for seed yield listed in SoyBase were within five mb of Yld14 (www.SoyBase.org, “SoyBase browser”, accessed 7/25/2015) based on positions from both Glyma 1.01 and Glyma 2.0. There were 54 unique genes, representing 195 different biological process GOs within close proximity (<1 mb) of Yld14 (Table 4.9).

Given the density of genes in this range, it is difficult to single out one or two that may be influencing yield. Instead, it is possible that many of these genes interact to influence soybean yield. Further efforts using a fine-mapping approach may be useful for narrowing the list of candidate genes for Yld14, which could be quite useful for plant breeders as it is a major QTL ($R^2 > 0.10$) that has not been previously reported in SoyBase (www.SoyBase.org, "SoyBase browser", accessed 7/25/2015). A, since Essex (southern) and Williams 82 (northern) are representative of different genetic groups, it is possible that breeders seeking to improve yield using this QTL might benefit by making similar genetically diverse crosses.

For soybean seed fatty acids, a total of 19 QTLs were detected, with six for palmitic (Figure 4.2), one for stearic (Figure 4.3), four for oleic (Figure 4.4), three for linoleic (Figure 4.5), and five for linolenic (Figure 4.6). The palmitic acid QTLs were designated Pal4, Pal9, Pal13, Pal17.1, Pal17.2, and Pal19 (Table 4.2). While Pal4 is located on chromosome 4, no other QTLs for seed palmitic acid listed in SoyBase have previously been located on chromosome 4 (www.SoyBase.org, "SoyBase browser", accessed 7/25/2015). Notably, a previous study using nearly the same parent lines for RIL development detected palmitic acid QTLs on the same chromosomes as Pal9, Pal 17.1 and Pal17.2, and Pal19 (Hyten et al., 2004b). While Essex was used in both studies, Hyten et al. (2004) used 'Williams' (Bernard and Lindahl, 1972) as a parent rather than Williams 82: Williams is the recurrent parent of Williams 82 (Bernard and Cremeens, 1988). However, confirming the QTLs from Hyten et al. (2004b) is difficult, as the associated markers are not listed for the palmitic acid QTLs detected in that study. Yet the Hyten et al. (2004b) palmitic acid QTL on chromosome 19 is in the same interval as the Dt1 locus. Comparison of the IM procedure (Dt1 covariate included) and the CIM procedure (Dt1 covariate not included) provides strong evidence that Pal19 is tightly linked with the Dt1 locus (Figure 4.2). Thus, it is likely that Pal19 is the same QTL as the palmitic acid

QTL on chromosome 19 detected by Hyten et al. (2004b), even though Pal19 falls short of the 1% significance threshold for confirmed QTL status (http://www.soybase.org/resources/QTL.php, accessed 7/25/2015).

With regard to palmitic acid candidate gene search, a major gene (Glyma.17g047000) is located near (< 1.5 mb) Pal17.1. However, Glyma.17g047000 does not segregate in this population, requiring a different genetic explanation for Pal17.1. Of the 12 unique genes with amino acid changes within one mb of Pal17.1, Glyma.17g027600 (GO:0006631, GO:0006635, and GO:0009062) and Glyma.17g034100 (GO:0006635) are associated with biological processes involving fatty acids (Table 4.10). Thus, while not guaranteed to be causative genes for Pal17.1, Glyma.17g027600 and Glyma.17g034100 become primary candidates. None of the 10 genes with amino acid changes near (< 1 mb) Pal17.2 contribute to biological processes involving fatty acids, so determining a candidate gene or genes is more challenging than for Pal17.1 (Table 4.10). However, as the QTL peak for Pal17.2 is not very precise (Figure 4.2), it is possible that a candidate gene may be found outside the range of 1 mb. As with Pal17.2, none of the proximal (< 1 mb) genes with amino acid changes contributed toward fatty acid biological processes (Tables 4.3, 4.6, 4.8, and 4.11). However, as they are involved in lipid metabolic process (GO:00006629), Glyma.13g079700 for Pal13 and Glyma.19g166800 and Glyma.19g171000 for Pal19 emerge as possible gene candidates.

The stearic acid QTL, located on chromosome 14, was designated as Ste14 (Figure 4.3). In comparison with the other QTL detected in this study, Ste14 was the largest by a wide margin with respect to LOD score (41.62) and R squared value (0.52) (Table 4.2). In addition to this study, several previous groups have identified stearic acid QTL on chromosome 14 (Hyten et al., 2004b; Panthee et al., 2006b; Li et al., 2011). Similar to this research, Hyten et al. (2004b) found a major stearic acid QTL (R^2 = 0.47) near Satt070 using nearly the same parents

as this population. While the position of Satt070 is not listed in Glyma 2.0, it is located in the same vicinity (< 1.5 mb) as Ste14 based on Glyma 1.01 positioning (Table 4.2). Further, both the Hyten et al. QTL and Ste14 are highly significant ($P < 0.01$), and both have Essex as the favorable parent; thus ste14 is an excellent candidate for a confirmed QTL of the stearic acid QTL on chromosome 14 detected in Hyten et al. (2014) (http://www.soybase.org/resources/QTL.php, accessed 7/25/2015). Therefore, the confirmed QTL symbol cqSeed stearic-001 is proposed for Ste14. While the major stearic acid gene Glyma.14g121400 was closely associated (< 1.5 mb) with Ste14, it does not differ between Essex and Williams 82. Neither gene with amino acid changes near (< 1 mb) Ste14 (Glyma.14g124400 and Glyma.14g125000) are likely candidates for this QTL based on biological process GOs (Table 4.9). Thus, the search for candidate genes was expanded to five mb in either direction, yielding Glyma.14g120200 (GO:0006629) as a possible choice for Ste14, as it is associated with lipid metabolism (Table 4.12). QTLs such as Ste14 could be useful when combined with SACPD-C mutants in an effort to achieve elevated stearic acid levels (> 200 g kg^{-1}) that are competitive with high yielding cultivars (Gillman et al., 2014).

Four QTLs on chromosomes 9, 13, 17, and 19 (designated as Ole9, Ole13, Ole17, and Ole19, respectively) were detected for oleic acid (Figure 4.4). Previous QTLs for oleic acid have been detected on chromosomes 9, 13, 17, and 19 (www.SoyBase.org, “SoyBase browser”, accessed 7/26/2015). Of these, Hyten et al. (2004b) is the only study with parental overlap with this research. However, the Hyten et al. (2004b) oleic acid QTL on chromosome 19 was in the same vicinity as the Dt1 locus, which differs from Ole19 (Figure 4.4).

Major oleic acid genes Glyma.10g278000 and Glyma.20g111000 listed in Gillman and Bilyeu (2012) were not segregating in this population, and thus could not be associated with the oleic acid QTLs detected in this study. However, searching for causative genes in close proximity (< 1 mb) with Ole9, Ole13, and

Ole17 yielded several possible candidates. Of the 49 unique genes with amino acid changes near (< 1 mb) Ole9, five emerge as leading candidates due to their involvement in biological processes involving fatty acids and/or lipids (Table 4.6). These genes are Glyma.09g191400 (GO:0006629), Glyma.09g191700 (GO:0006631 and GO:0006635), Glyma.09g200500 (GO:0006636 and GO:0019216), Glyma.09g207900 (GO:0006636), and Glyma.09g209400 (GO:0006629 and GO:0016042) (Table 4.6). For Ole13, Glyma.13g043000, Glyma.13g043500, and Glyma.13g043600 are possible gene candidates as they are each involved in fatty acid transport (GO:0015908) (Table 4.8). Leading candidates for Ole17 included Glyma.17g027600 (GO:0006631, GO:0006635, and GO:0009062) and Glyma.17g034100 (GO:0006635) due to their involvement in biological processes involving fatty acids (Table 4.10). Candidate genes for Ole19 are not readily predictable, as no genes with amino acid changes within one mb are associated with fatty acid or lipid biological processes (Table 4.11). However, the QTLs detected for oleic acid (Ole9, Ole13, Ole17, and Ole19) in this study could be useful to plant breeders seeking to stabilize high concentrations (> 800 g kg^{-1}) for that trait (Lee et al., 2012; Fallen et al., 2012).

The QTLs for linoleic acid detected in this study, designated as Lin13.1, Lin13.2, and Lin19 were located on chromosomes 13, 13, and 19, respectively (Figure 4.5). Previous QTL on these chromosomes have been detected in Hyten et al. (2004b) using nearly the same parents. Yet, Lin19 is not near the Dt1 locus, whereas the linoleic acid QTL on chromosome 19 in Hyten et al. (2004b) is. Additionally, Both Lin13.1 and Lin13.2 fail to meet the 1% significance threshold needed for confirmed QTL status (http://www.soybase.org/resources/QTL.php, accessed 7/26/2015). While none of the genes with amino acid changes near (< 1 mb) Lin13.2 or Lin19 are involved in fatty acid or lipid biological processes, there are several candidates associated with Lin13.1 (Tables 4.8, 4.11). These candidates for Lin13.1 include Glyma.13g213500 (GO:0006636 and GO:0019216) and Glyma.13g215400 (GO:0006636) (Table 4.8). While few

efforts seeking to improve soybean fatty acids are focused on linoleic, it is possible that the QTLs listed in this study could be used to decrease linoleic acid. Such a decrease would allow other fatty acids, primarily oleic as the second most abundant fatty acid, to fill this void. By using such an approach, breeders may be able to maintain higher levels of stability in cultivars with high levels (> 800 g kg^{-1}) of oleic acid (Lee et al., 2012; Fallen et al., 2012).

Five QTLs on chromosomes 9, 13, 17, and 19 (designated as Len9.1, Len9.2, Len13, Len17, and Len19) were detected for linolenic acid (Figure 4.6). While Len17 is located on chromosome 17, no other QTLs for seed linolenic acid listed in SoyBase have previously been located on chromosome 17 (www.SoyBase.org, “SoyBase browser”, accessed 7/26/2015). This is not the case for any of the other linolenic acid QTLs in this study. Hyten et al. (2004b) previously detected a QTL on chromosome 19 for linolenic acid using nearly the same parents as in this study. Yet, Len19 is not near the Dt1 locus, whereas the linolenic acid QTL on chromosome 19 in Hyten et al. (2004b) is. However, previous studies have detected QTLs for linolenic acid on chromosome 19 linked to Satt238 (Kim et al., 2010) and Satt495 (Bachlava et al., 2009); both of which are very close (< 2 mb) to Len19 based on the Glyma 1.01 reference genome position (Table 4.2). While this study used different parents from Kim et al. (2010) and Bachlava et al. (2009), the significance ($P < 0.01$) and location of Len19 with respect to previous QTLs make it a strong candidate as a positional QTL confirmation (Smallwood et al., 2014). Additionally, Len9.2 is a candidate for positional QTL confirmation (Smallwood et al., 2014); as it is highly significant ($P < 0.01$) and closely associated with Satt499 (approximately 2 mb based on Glyma 1.01), which has previously been linked to a QTL for linolenic acid (Li et al., 2010).

As with oleic acid, none of the QTLs for linolenic acid detected in this study were located on chromosomes 2, 14, or 18, and so could not be associated with the

major linolenic acid genes (Glyma.02g227200, Glyma.14g194300, and Glyma.18g062000) listed in Gillman and Bilyeu (2012). However, searching for causative genes in close proximity (< 1 mb) with Len9.1, Len9.2, Len13, and Len17 yielded several possible candidates. Candidate genes with amino acid changes near Len9.1 with involvement in fatty acid or lipid biological processes include Glyma.09g041200 (GO:0006629 and GO:0016042), Glyma.09g043700 (GO:0000038, GO:0006633, and GO:0008610), and Glyma.09g043800 (GO:0019915) (Table 4.6). For Len9.2, the gene candidates were Glyma.09g191400 (GO:0006629), Glyma.09g191700 (GO:0006631 and GO:0006635), Glyma.09g200500 (GO:0006636 and GO:0019216), Glyma.09g207900 (GO:0006636), and Glyma.09g209400 (GO:0006629 and GO:0016042) (Table 4.6). For Len13, Glyma.13g043000, Glyma.13g043500, and Glyma.13g043600 are possible gene candidates as they are each involved in fatty acid transport (GO:0015908) (Table 4.8). The candidate gene involved in fatty acid biological processes for Len17 is Glyma.17g027600 (GO:0006631, GO:0006635, and GO:0009062) (Table 4.10). Candidate genes for Len19 are not readily predictable, as no genes with amino acid changes within one mb are associated with fatty acid or lipid biological processes (Table 4.11). However, the QTLs detected for linolenic acid (Len9.1, Len9.2, Len13, Len17, and Len19) in this study could be useful to plant breeders seeking to meet the dual goal of high oleic (> 800 g kg^{-1}), low linolenic (< 30 g kg^{-1}) soybeans.

Seven QTLs on chromosomes 6, 7, 9, and 13 (designated as Prot6.1, Prot6.2, Prot7, Prot9.1, Prot9.2, Prot13.1, and Prot13.2) were detected for protein in this study (Figure 4.7). Numerous protein QTLs have been previously reported on each of these chromosomes (www.SoyBase.org, “SoyBase browser”, accessed 7/26/2015). Of particular note are those detected by Hyten et al. (2004a), which used nearly the same parents for RIL development as this study. While Essex was used in both studies, Hyten et al. (2004a) had ‘Williams’ (Bernard and Lindahl, 1972) rather than Williams 82. Notably, the markers listed in the interval

range (Satt335 and Satt144) for the Hyten et al. (2004a) protein QTL on chromosome 13 are located approximately 2 mb from Prot 13.2 and Prot13.1, respectively, based on Glyma 1.01 positions (www.SoyBase.org, “SoyBase browser”, accessed 7/26/2015). As the additional RILs and markers used in this study afforded more recombination opportunities and higher resolution mapping, respectively, it is possible that Hyten et al. (2004a) was unable to detect the presence of two QTL on chromosome 13. Since both Prot13.1 and Prot13.2 are highly significant ($P < 0.01$) and closely associated with the QTL(s) on chromosome 13 from Hyten et al. (2004a), they are excellent candidates for confirmed QTL (http://www.soybase.org/resources/QTL.php, accessed 7/26/2015). Thus, confirmed QTL symbols of cqSeed protein-004 (Prot13.1) and cqSeed protein-005 (Prot13.2) are proposed. Also, Prot13.2 is close to Sat_090 (< 1 mb based on Glyma 1.01), which has previously been linked to a protein QTL (Reinprecht et al., 2006). Thus, Prot13.2 is also a candidate for positional confirmation of QTL (Smallwood et al., 2014). Further, Prot9.2 is closely associated (approximately 2 mb based on Glyma 1.01 and Glyma 2.0) with Satt475 (Lu et al., 2012) and Satt273 (Rossi et al., 2013). As Prot9.2 is highly significant ($P < 0.01$) it is also a strong candidate for positional confirmation of QTL (Smallwood et al., 2014). Numerous genes with a wide variety of associated biological processes are in close proximity (< 1 mb) from the protein QTLs detected in this study (Tables 4.4-4.6, 4.8). While further research is needed to fully understand which genes influence these QTLs, it is still possible for breeders to use this information in making protein improvements.

The two QTLs for oil detected in this population were located on chromosomes 6 and 11 (designated as Oil6 and Oil11, respectively) (Figure 4.8). Numerous studies have previously reported seed oil QTLs on chromosomes 6 and 11 (www.SoyBase.org, “SoyBase browser”, accessed 7/26/2015). Notably, a QTL for seed oil on chromosome 6 was detected by Hyten et al. (2004a), which used nearly the same parents for RIL development as this study. However, the Hyten

et al. (2004a) QTL on chromosome 6 did not display consistent effects across the environments used in that study, making it an unfavorable candidate for a confirmed QTL. Glyma.06g294800 (GO:0006629) is a possible candidate gene for Oil6, as it is located within one mb and is involved in lipid metabolism (Table 4.4). Of the 14 genes with amino acid changes near (< 1 mb) Oil11, none were involved in lipid specific biological processes (Table 4.7). Detection of these oil QTLs could be useful for breeders seeking to improve that trait.

As many of the QTLs detected in this study for different traits are proximally close (< 2 mb with overlapping candidate genes), it is worth considering the possibility of pleiotropy in these circumstances (Table 4.2). On chromosome 6, Prot6.2 and Oil6 are located within 2 mb of each other (Table 4.2). Given the historical evidence that oil and protein in soybean seed are negatively correlated (Yaklich et al., 2002), it seems reasonable to think that whatever causative gene or genes in this region is influencing seed accumulation for protein and oil in opposite directions. Further evidence is provided with the opposing direction of effects in this QTL region for protein and oil (Table 4.2).

Several of these traits had overlapping gene candidates on chromosome 9 (Table 4.6). Pal9 had overlapping candidate genes with both Prot9.2 and Ole9, suggesting possible pleiotropic effects in this region for these traits (Table 4.6). In addition to overlapping with Pal9, Ole9 had overlapping candidate genes with Len9.2 (Table 4.6). The possibility of pleiotropic effects for the fatty acid QTLs in this region (Pal9, Ole9, and Len9.2) is especially strong, as they are derived from the same biosynthetic pathway (Gillman and Bilyeu, 2012). Further possible pleiotropic regions involving fatty acids were found on chromosome 13 (Ole13 and Len13) (Table 4.8), chromosome 17 (Pal17.1, Ole17, and Len17) (Table 4.10), and chromosome 19 (Ole19, Lin19, and Len19) (Table 4.11). Understanding these pleiotropic relationships will be helpful for breeders seeking to improve these traits.

Conclusions

Due to the importance of yield, fatty acids, protein, and oil in soybean production, it is critical to develop an improved understanding of significant genomic regions governing these traits. In this study, 29 QTLs were detected for yield (1), palmitic acid (6), stearic acid (1), oleic acid (4), linoleic acid (3), linolenic acid (5), protein (7), and oil (2) (Table 4.2). Of these QTL, Ste14 stands out as the most influential QTL (R^2 = 0.52) (Table 4.2). Additionally, Ste14, along with Prot13.1 and Prot13.2 are excellent candidates for confirmed QTLs (http://www.soybase.org/resources/QTL.php, accessed 7/25/2015); while Len9.2, Len19, Prot9.2, and Prot13.2 are all strong candidates for positional QTL confirmations (Smallwood et al., 2014).

Since none of the major fatty acid genes listed in Gillman and Bilyeu (2012) segregate in this population, the QTLs detected for fatty acids are likely the result of small effect modifier genes (Hyten et al., 2004b). Many of the genes with amino acid changes in close proximity to the fatty acid QTLs are involved in biological processes for fatty acids and/or lipids (Tables 4.3, 4.6, 4.8-4.12). Such genes may be useful in breeding for goals of fatty acid improvement (Lee et al., 2012; Fallen et al., 2012; Gillman et al., 2014). Similarly, genes with amino acid changes in genomic regions near yield, protein, and oil were plentiful (Tables 4.4-4.9), and may contribute to some of the variation observed in these traits (Table 4.1).

All of the traits except yield were involved in probable pleiotropic relationships with other traits in this study. Of note is the pleiotropic effect between protein and oil on chromosome 6 (Table 4.4). Given the well-established negative relationship between protein and oil (Yaklich et al., 2002), it likely that whatever causative gene or genes in this region is influencing seed accumulation of both protein and oil, but in opposite directions. While the E1 maturity gene does segregate in this population, it is likely not the causative gene for this pleiotropic

effect between protein and oil, as it is not linked (> 25 mb) to either Prot6.2 or Oil6. Each of the fatty acids except stearic had at least two QTLs with overlapping candidate genes for other fatty acids (Tables 4.6, 4.8, 4.10-4.11), indicating probable pleiotropy for seed fatty acids at these loci. Accounting for such pleiotropic effects will be useful for breeders trying to adjust multiple traits at once.

Overall, the findings from this research contribute new information to the genetic understanding of soybean yield, fatty acids, protein and oil. This understanding will be useful in making trait improvements. Further research seeking to narrow the list of candidate genes for these QTLs would be beneficial.

References

Bachlava, E., R. Dewey, J. Burton, and A. Cardinal. 2009. Mapping and comparison of quantitative trait loci for oleic acid seed content in two segregating soybean populations. Crop Sci. 49:433-442.

Bernard, R.L., and C.R. Cremeens. 1988. Registration of 'Williams 82' soybean. Crop Sci. 28:1027–1028.

Bernard, R.L., and D.A. Lindahl. 1972. Registration of 'Williams' soybean. Crop Sci. 12:716.

Bilyeu, K., J.D. Gillman, and A.R. LeRoy. 2011. Novel FAD3 mutant allele combinations produce soybeans containing 1% linolenic acid in the seed oil. Crop Sci. 51:259–264.

Boersma, J.G., J.D. Gillman, K.D. Bilyeu, G.R. Ablett, C. Grainger, and I. Rajcan. 2012. New mutations in a delta-9-stearoyl-acyl carrier protein desaturase gene associated with enhanced stearic acid levels in soybean seed. Crop Sci. 52:1736–1742.

Bolon, Y., W.J. Haun, W.W. Xu, D. Grant, M.G. Stacey, R.T. Nelson, D.J. Gerhardt, J.A. Jeddeloh, G. Stacey, G.J. Muehlbauer, J.H. Orf, S.L. Naeve, R.M. Stupar, and C.P. Vance. 2011. Phenotypic and genomic analyses of fast neutron mutant population resource in soybean. Plant Physiol. 156:240-253.

Brim, C.A. 1966. A modified pedigree method of selection in soybeans. Crop Sci. 6:220.

Broman, K.W., H. Wu, Ś. Sen, and G.A. Churchill. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890.

Browning, B.L., and S.R. Browning. 2009. A unified approach to genotype imputation and haplotype phase inference for large data sets of trios and unrelated individuals. Am. J. Hum. Genet. 84:210-223.

Browning, S.R., and B.L. Browning. 2007. Rapid and accurate haplotype phasing and missing data inference for whole genome association studies using localized haplotype clustering. Am. J. Hum. Genet. 81:1084-1097.

Cardinal, A.J., R. Whetten, S. Wang, J. Auclair, D. Hyten, P. Cregan, E.

Bachlava, J. Gillman, M. Ramirez, R. Dewey, G. Upchurch, L. Miranda, and J.W. Burton. 2014. Mapping the low palmitate fap1 mutation and validation of its effects in soybean oil and agronomic traits in three soybean populations. Theor. Appl. Genet. 127:97-111.

Churchill, G.A., and R.W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963–971.

Fallen, B.D., K. Rainey, C.E. Sams, D.A.Kopsell, and V.R. Pantalone. 2012. Evaluation of agronomic and seed characteristics in elevated oleic acid soybean lines in the south-eastern US. J. Am. Oil Chem. Soc. 89:1333-1343.

Federal Register. 2015. Final determination regarding partially hydrogenated oils. https://www.federalregister.gov/articles/2015/06/17/2015-14883/final-determination-regarding-partially-hydrogenated-oils (accessed 24 July 2015).

Fehr W.R., and C.E. Caviness. 1977. Stages of soybean development. Special Report, Agriculture and Home Economics Experiment Station, Iowa State University, 1977, issue 80, p 11.

Gillman, J.D., and K.D. Bilyeu. 2012. Genes and alleles for quality traits on the soybean genetic/physical map. In: R.F. Wilson, editor, Designing soybean for 21st century markets. AOCS Press, Urbana, IL. p. 67-96.

Gillman, J.D., M.G. Stacy, Y. Cui, H.R. Berg, and G. Stacey. 2014. Deletions of the SACPD-C locus elevate seed stearic acid but also result in fatty acid and morphological alterations in nitrogen fixing nodules. BMC Plant Biol. 14:143.

Goodstein, D.M., S. Shu, R. Howson, R. Neupane, R.D. Hayes, J. Fazo, T. Mitros, W. Dirks, U. Hellsten, N. Putnam, and D.S. Rokhsar. 2012. Phytozome: a comparative platform for green plant genomics. Nucl. Acids Res. 40:D1178-D1186

Grant, D., R.T. Nelson, S.B. Cannon, and R.C. Shoemaker. 2010. SoyBase, the USDA-ARS soybean genetics and genomics database. Nucl. Acids Res.

38:D843-D846.

Hyten, D.L., V.R. Pantalone, C.E. Sams, A.M. Saxton, D. Landau-Ellis, T.R. Stefaniak, and M.E. Schmidt. 2004a. Seed quality QTL in a prominent soybean population. Theor. Appl. Genet. 109:552-561.

Hyten, D.L., V.R. Pantalone, A.M. Saxton, M.E. Schmidt, and C.E. Sams. 2004b. Molecular mapping and identification of soybean fatty acid modifier quantitative trait loci. J. Am. Oil Chem. Soc. 81:1115-1118.

Kim, H. Y. Kim, S. Kim, B. Son, Y. Choi, J. Kang, Y. Park, Y. Cho, and I. Cho. 2010. Analysis of quantitiative trait loci (QTLs) for seed size and fatty acid composition using recombinant inbred lines in soybean. J. Life Sci. 20:1186-1192.

Kinney, A.J. 1996. Development of genetically engineered soybean oils for food application. J. Food Lipids. 3:273-292.

Kinney, A.J., and T.E. Clemente. 2005. Modifying soybean oil for enhanced performance in biodiesel blends. Fuel Pro. Technol. 86:1137–1147.

Kris-Etherton, P.M., and S. Yu. 1997. Individual fatty acid effects on plasma lipids and lipoproteins: Human studies. Am. J. Clin. Nutr. 65:S1628–S1644.

Lee, J.D., K.D. Bilyeu, V.R. Pantalone, A.M. Gillen, Y.S. So, and J.G. Shannon. 2012. Environmental stability of oleic acid concentration in seed oil for soybean lines with FAD2-1A and FAD2-1B mutant genes. Crop Sci. 52:1290–1297.

Li, H., T. Zhao, Y. Wang, D. Yu, S. Chen, R. Zhou, and J. Gai. 2011. Genetic structure composed of additive QTL, epistatic QTL pairs and collective unmapped minor QTL conferring oil content and fatty acid components of soybeans. Euphytica 182:117-132.

Lu, W., Z. Wen, H. Li, D. Yuan, J. Li, H. Zhang, Z. Huang, S. Cui, and W. Du. 2013. Identification of the quantitative trait loci (QTL) underlying water soluble protein content in soybean. Theor. Appl. Genet. 126:425-433.

Martínez, O., and R.N. Curnow. 1992. Estimating the locations and sizes of the effects of quantitative trait loci using flanking markers. Theor. Appl. Genet.

85:480-488.

Nyquist, W.E. 1991. Estimation of heritability and prediction of selection response in plant populations. Crit. Rev. Plant Sci.10:235–322.

Pantalone, V.R., R.F. Wilson, W.P. Novitzky, and J.W. Burton. 2002. Genetic regulation of elevated stearic acid concentration in soybean oil. J. Am. Oil Chem. Soc. 79:543–553.

Panthee, D.R., V.R. Pantalone, C.E. Sams, A.M. Saxton, D.R. West, J.H. Orf, and A.S. Killam. 2006a. Quantitative trait loci controlling sulfur containing amino acids, methionine and cysteine, in soybean seeds. Theor. Appl. Genet. 112:546–553.

Panthee, D.R., V.R. Pantalone, and A.M. Saxton. 2006b. Modifier QTL for fatty acid composition in soybean oil. Euphytica 152:67-73.

Pham, A.T., J.D. Lee, J.G. Shannon, and K.D. Bilyeu. 2010. Mutant alleles of FAD2-1A and FAD2-1B combine to produce soybeans with the high oleic acid seed oil trait. BMC Plant Biol. 10:195.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Reinprecht, Y., V. Poysa, K. Yu, I. Rajcan, G. Abblett, and K. Pauls. 2006. Seed and agronomic QTL in low linolenic acid, lipoxygenase-free soybean (Glycine max (L.) Merrill) germplasm. Genome 49:1510-1527.

Rossi, M., J. Orf, L. Liu, Z. Dong, and I. Rajcan. 2013. Genetic basis of soybean adaptation to North American vs. Asian mega-environments in two independent populations from Canadian x Chinese crosses. Theor. Appl. Genet. 126:1809-1823.

SAS Institute Inc. 2002-2012. Cary, NC, USA. SAS 9.4)

Schmutz, J., S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D.L. Hyten, Q. Song, J.J. Thelen, J. Cheng, et al. 2010. Genome sequence of the palaeopolyploid soybean. Nature 463:178–183.

Smallwood, C.J., C.N. Nyinyi, D.A. Kopsell, C.E. Sams, D.R. West, P. Chen, S.K.

Kantartzi, P.B. Cregan, D.L. Hyten, and V.R. Pantalone. 2014. Detection and confirmation of quantitative trait loci for soybean seed isoflavones. Crop Sci. 54:1-12.

Smith, T.J., and H.M. Camper. 1973. Registration of Essex Soybean (Reg. No. 97). Crop Sci. 13:495.

Song, Q., D.L. Hyten, G. Jia, C.V. Quigley, E.W. Fickus, R.L. Nelson, and P.B. Cregan. 2013. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS ONE 8:e54985.

SoyBase and the Soybean Breeder's Toolbox. 2007. QTL nomenclature. http://www.soybase.org/resources/QTL.php (accessed 26 July 2015).

Spencer, M.M., D. Landau-Ellis, E.J. Meyer, and V.R. Pantalone. 2004. Molecular markers associated with linolenic acid content in soybean. J. Am. Oil Chem. Soc. 81:559–562.

USDA, ARS, National Genetic Resources Program. Germplasm Resources Information Network - (GRIN) [Online Database]. National Germplasm Resources Laboratory, Beltsville, Maryland. URL: http://www.ars-grin.gov.4/cgi-bin/npgs/html/index.pl?language=en (24 July 2015)

Wilson, R.F. 2004. Seed composition. In: H.R. Boerma and J.E. Specht, editors, Soybeans: Improvement, production, and uses. 3rd ed. ASA, CSSA, and SSSA, Madison, WI. p. 621–678.

Wimmer, V., T. Albrecht, H.J. Auinger, and C.C. Schön. 2012. Synbreed: A framework for the analysis of genomic prediction using R. Bioinformatics. 28:2086-2087.

Yaklich, R.W., B. Vinyard, M. Camp, and S. Douglass. 2002. Analysis of seed protein and oil from soybean northern and southern region uniform tests. Crop Sci. 42:1504–1515.

Appendix D-Chapter 4 Tables and Figures

Tables

Table 4.1 Simple statistics taken from combined analysis using data from 2010, 2013, and 2014 field seasons for soybean population E×W-50K subset consisting of 203 F5 derived RILs.

Trait	Genotype P value	GxE P value	Essex	Williams 82	min	mean	max	std. dev.†	LSD value	h^{2}‡
Maturity (Julian)	***	***	273.80	264.65	269.73	272.51	277.12	1.45	1.79	0.81
Height (cm)	***	***	75.25	85.72	57.42	90.31	125.07	16.87	7.93	0.97
Yield (kg ha^{-1})	***	***	3296.88	2463.91	2298.00	3048.27	3706.13	287.63	480.91	0.64
Palmitic (g kg^{-1} seed oil)	***	***	123.30	114.39	97.91	119.65	133.94	6.98	3.74	0.96
Stearic (g kg^{-1} seed oil)	***	***	43.92	40.43	35.80	41.59	49.63	2.96	1.99	0.94
Oleic (g kg^{-1} seed oil)	***	***	325.72	350.67	262.04	335.23	470.48	39.22	20.31	0.97
Linoleic (g kg^{-1} seed oil)	***	***	431.05	424.76	333.86	430.58	476.08	27.36	14.74	0.96
Linolenic (g kg^{-1} seed oil)	***	***	76.09	69.80	55.40	73.00	92.45	7.83	4.18	0.96
Protein (g kg^{-1} seed dry weight)	***	***	432.58	429.61	391.04	420.32	449.00	10.74	8.10	0.93
Oil (g kg^{-1} seed dry weight)	***	***	210.12	225.86	198.03	216.67	234.66	5.41	3.77	0.94

***Significant at 0.001 probability level

E×W-50K = soybean population with parental lines Essex and Williams 82

† std. deviation of LSMEANs

‡ heritability calculated using entry means basis (Nyquist, 1991)

Table 4.2 Quantitative trait loci (QTL) for yield, fatty acids, protein, and oil detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was estimated from the combined analysis using data from 2010, 2013, and 2014 field seasons. Significant QTL at the 1% threshold (based on 10,000 permutations) are displayed in bold text.

Trait	LOD Threshold	QTL	Chr	SNP	Position (Glyma 1.01)	Position (Glyma 2.0)	LOD	R2†	Effect‡
Yield	1% 4.13	**Yld14**	14	ss715617925	1879468	1874810	13.05	0.23	144.43
	5% 3.41								
Palmitic	1% 4.19	**Pal4**	4	ss715587892	38135736	41298992	4.82	0.04	1.46
	5% 3.43	**Pal9**	9	ss715603983	38013391	40608710	8.50	0.09	-2.12
		Pal13	13	ss715614223	2424574	19443133	3.60	0.02	0.96
		Pal17.1	17	ss715626470	2379823	2373531	5.34	0.03	-1.35
		Pal17.2	17	ss715628181	7792405	7523398	3.80	0.02	-1.03
		Pal19	19	ss715635231	42655124	42856209	3.79	0.04	-2.53
Stearic	1% 4.22	**Ste14**	14	ss715618483	35525630	18696303	41.62	0.52	-2.18
	5% 3.44								
Oleic	1% 4.1	Ole9	9	ss715604251	39726952	42404009	3.51	0.02	5.79
	5% 3.4	Ole13	13	ss715617125	7331729	14520516	3.81	0.02	-5.76
		Ole17	17	ss715626470	2379823	2373531	3.58	0.01	4.54
		Ole19	19	ss715633482	2391589	2437848	5.50	0.02	6.08
Linoleic	1% 4.17	Lin13.1	13	ss715615238	30836854	32049208	3.48	0.02	3.92
	5% 3.40	Lin13.2	13	ss715615962	36512850	37653753	3.82	0.02	3.82
		Lin19	19	ss715633482	2391589	2437848	4.47	0.04	-5.46
Linolenic	1% 4.24	**Len9.1**	9	ss715603590	3168642	3209966	4.43	0.04	-1.55
	5% 3.44	**Len9.2**	9	ss715604287	39933930	42611042	6.68	0.04	-1.66
		Len13	13	ss715617125	7331729	14520516	6.65	0.05	1.79
		Len17	17	ss715625991	1315668	1309414	3.54	0.01	-0.74
		Len19	19	ss715633482	2391589	2437848	5.43	0.02	-1.26
Protein	1% 4.18	Prot6.1	6	ss715595361	5723944	5729763	3.51	0.03	-2.13
	5% 3.41	Prot6.2	6	ss715594899	47847021	48468113	3.69	0.07	3.20
		Prot7	7	ss715598820	8377451	8417485	3.85	0.05	-2.44
		Prot9.1	9	ss715605452	911566	913310	3.81	0.07	-2.90
		Prot9.2	9	ss715603959	37781856	40376477	5.74	0.07	2.87
		Prot13.1	13	ss715615583	33597361	34809421	5.18	0.01	-1.52
		Prot13.2	13	ss715616094	38249824	39427301	5.12	0.06	-3.03
Oil	1% 4.11	**Oil6**	6	ss715594709	46914819	47461885	6.07	0.12	-2.10
	5% 3.39	Oil11	11	ss715610384	37219747	32767711	3.50	0.07	1.45

† estimated variance in trait captured by QTL, with range between 0 (none) and 1 (all)

‡ Estimated effect in kg ha-1 (yield) or g kg-1 (all other traits) with respect to the Williams 82 allele

E×W-50K = soybean population with parental lines Essex and Williams 82

Table 4.3 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 4 based on Glyma 2.0 position.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Pal4	Glyma.04g162700	40326865	40331483	GO:0006564	L-serine biosynthetic process
Pal4	Glyma.04g162700	40326865	40331483	GO:0009555	pollen development
Pal4	Glyma.04g162700	40326865	40331483	GO:0009790	embryo development
Pal4	Glyma.04g162700	40326865	40331483	GO:0048364	root development
Pal4	Glyma.04g165000	41246946	41250808	GO:0042744	hydrogen peroxide catabolic process
Pal4	Glyma.04g165000	41246946	41250808	GO:0055114	oxidation-reduction process
Pal4	Glyma.04g165700	41535757	41559071	GO:0016043	cellular component organization
Pal4	Glyma.04g165700	41535757	41559071	GO:0030036	actin cytoskeleton organization
Pal4	Glyma.04g167900	42124874	42127039	GO:0009765	photosynthesis light harvesting
Pal4	Glyma.04g167900	42124874	42127039	GO:0015979	photosynthesis
Pal4	Glyma.04g167900	42124874	42127039	GO:0019344	cysteine biosynthetic process
Pal4	Glyma.04g167900	42124874	42127039	GO:0030003	cellular cation homeostasis
Pal4	Glyma.04g167900	42124874	42127039	GO:0070838	divalent metal ion transport
Pal4	Glyma.04g167900	42124874	42127039	GO:0080167	response to karrikin

Table 4.4 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 6 based on Glyma 2.0 position.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0000902	cell morphogenesis
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0006096	glycolysis
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0006816	calcium ion transport
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0006833	water transport
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0006972	hyperosmotic response
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0007030	Golgi organization
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0009266	response to temperature stimulus
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0009651	response to salt stress
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0009664	plant-type cell wall organization
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0009750	response to fructose stimulus
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0009825	multidimensional cell growth
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0009832	plant-type cell wall biogenesis
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0009833	primary cell wall biogenesis
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0016049	cell growth
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0030243	cellulose metabolic process
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0030244	cellulose biosynthetic process
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0042538	hyperosmotic salinity response
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0046686	response to cadmium ion
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0048193	Golgi vesicle transport
Prot6.1	Glyma.06g069600	5338316	5344365	GO:0048767	root hair elongation
Prot6.1	Glyma.06g072800	5614187	5615402	GO:0005975	carbohydrate metabolic process
Prot6.1	Glyma.06g074000	5701847	5704926	GO:0006886	intracellular protein transport
Prot6.1	Glyma.06g074000	5701847	5704926	GO:0006944	cellular membrane fusion
Prot6.1	Glyma.06g074000	5701847	5704926	GO:0016192	vesicle-mediated transport
Prot6.1	Glyma.06g074200	5715485	5718978	GO:0008150	biological_process
Prot6.1	Glyma.06g075100	5807768	5812501	GO:0005975	carbohydrate metabolic process
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0006511	ubiquitin-dependent protein catabolic process
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0007062	sister chromatid cohesion
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0009640	photomorphogenesis
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0009790	embryo development
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0009793	embryo development ending in seed dormancy
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0009845	seed germination
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0009880	embryonic pattern specification
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0009909	regulation of flower development
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0009933	meristem structural organization
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0010072	primary shoot apical meristem specification
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0010162	seed dormancy
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0010182	sugar mediated signaling pathway
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0010228	vegetative to reproductive phase transition of meristem
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0010431	seed maturation
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0010564	regulation of cell cycle process
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0016567	protein ubiquitination
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0019915	lipid storage
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0045595	regulation of cell differentiation
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0048316	seed development
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0048366	leaf development
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0048825	cotyledon development
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0050826	response to freezing
Prot6.1	Glyma.06g075500	5835169	5847257	GO:0051301	cell division
Prot6.1	Glyma.06g075800	5865835	5869879	GO:0006635	fatty acid beta-oxidation
Prot6.1	Glyma.06g075800	5865835	5869879	GO:0008150	biological_process
Prot6.1	Glyma.06g075800	5865835	5869879	GO:0016558	protein import into peroxisome matrix
Prot6.1	Glyma.06g075800	5865835	5869879	GO:0048573	photoperiodism flowering
Prot6.1	Glyma.06g076000	5880948	5885101	GO:0000338	protein deneddylation
Prot6.1	Glyma.06g076000	5880948	5885101	GO:0009640	photomorphogenesis
Prot6.1	Glyma.06g076000	5880948	5885101	GO:0009733	response to auxin stimulus
Prot6.1	Glyma.06g076000	5880948	5885101	GO:0010100	negative regulation of photomorphogenesis
Prot6.1	Glyma.06g076000	5880948	5885101	GO:0010387	signalosome assembly
Prot6.1	Glyma.06g076000	5880948	5885101	GO:0010388	cullin deneddylation
Prot6.1	Glyma.06g076000	5880948	5885101	GO:0010971	positive regulation of G2/M transition of mitotic cell cycle
Prot6.1	Glyma.06g076000	5880948	5885101	GO:0016567	protein ubiquitination
Prot6.1	Glyma.06g076000	5880948	5885101	GO:0016571	histone methylation
Prot6.1	Glyma.06g076000	5880948	5885101	GO:0016579	protein deubiquitination
Prot6.1	Glyma.06g076000	5880948	5885101	GO:0045732	positive regulation of protein catabolic process
Prot6.1	Glyma.06g076000	5880948	5885101	GO:0045893	positive regulation of transcription DNA-dependent
Prot6.1	Glyma.06g076000	5880948	5885101	GO:2000082	regulation of L-ascorbic acid biosynthetic process
Prot6.1	Glyma.06g076200	5902265	5902594	GO:0000154	rRNA modification
Prot6.1	Glyma.06g076200	5902265	5902594	GO:0001708	cell fate specification
Prot6.1	Glyma.06g076200	5902265	5902594	GO:0051301	cell division
Prot6.1	Glyma.06g076900	5947978	5950087	GO:0009585	red far-red light phototransduction
Prot6.1	Glyma.06g076900	5947978	5950087	GO:0009744	response to sucrose stimulus
Prot6.1	Glyma.06g076900	5947978	5950087	GO:0009813	flavonoid biosynthetic process
Prot6.1	Glyma.06g076900	5947978	5950087	GO:0010017	red or far-red light signaling pathway
Prot6.1	Glyma.06g076900	5947978	5950087	GO:0010224	response to UV-B
Prot6.1	Glyma.06g076900	5947978	5950087	GO:0010264	myo-inositol hexakisphosphate biosynthetic process
Prot6.1	Glyma.06g080300	6170065	6172699	GO:0006400	tRNA modification
Prot6.1	Glyma.06g080300	6170065	6172699	GO:0006626	protein targeting to mitochondrion
Prot6.1	Glyma.06g080300	6170065	6172699	GO:0009165	nucleotide biosynthetic process
Prot6.1	Glyma.06g080300	6170065	6172699	GO:0015946	methanol oxidation
Prot6.1	Glyma.06g080300	6170065	6172699	GO:0019415	acetate biosynthetic process from carbon monoxide

Table 4.4 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Prot6.1	Glyma.06g081400	6260611	6264979	GO:0006412	translation
Prot6.1	Glyma.06g081400	6260611	6264979	GO:0009220	pyrimidine ribonucleotide biosynthetic process
Oil6	Glyma.06g273800	46473511	46474932	GO:0006508	proteolysis
Oil6	Glyma.06g273800	46473511	46474932	GO:0007568	aging
Oil6	Glyma.06g273800	46473511	46474932	GO:0009723	response to ethylene stimulus
Oil6	Glyma.06g273800	46473511	46474932	GO:0009817	defense response to fungus incompatible interaction
Oil6	Glyma.06g273800	46473511	46474932	GO:0010150	leaf senescence
Oil6	Glyma.06g273900	46477620	46484375	GO:0006508	proteolysis
Oil6	Glyma.06g273900	46477620	46484375	GO:0009651	response to salt stress
Oil6	Glyma.06g273900	46477620	46484375	GO:0052541	plant-type cell wall cellulose metabolic process
Oil6	Glyma.06g273900	46477620	46484375	GO:0052546	cell wall pectin metabolic process
Oil6	Glyma.06g275100	46608023	46610332	GO:0006508	proteolysis
Oil6	Glyma.06g275200	46618561	46619824	GO:0006508	proteolysis
Oil6	Glyma.06g275400	46631440	46637225	GO:0006457	protein folding
Oil6	Glyma.06g277100	46831714	46834771	GO:0005975	carbohydrate metabolic process
Oil6	Glyma.06g277200	46838315	46841977	GO:0005975	carbohydrate metabolic process
Oil6	Glyma.06g277200	46838315	46841977	GO:0009664	plant-type cell wall organization
Oil6	Glyma.06g277200	46838315	46841977	GO:0010583	response to cyclopentenone
Oil6	Glyma.06g277200	46838315	46841977	GO:0042545	cell wall modification
Oil6	Glyma.06g277200	46838315	46841977	GO:0042547	cell wall modification involved in multidimensional cell growth
Oil6	Glyma.06g280800	47066404	47067073	GO:0009409	response to cold
Oil6	Glyma.06g280800	47066404	47067073	GO:0009733	response to auxin stimulus
Oil6	Glyma.06g280800	47066404	47067073	GO:0042742	defense response to bacterium
Oil6	Glyma.06g281400	47078762	47079840	GO:0006508	proteolysis
Oil6	Glyma.06g281400	47078762	47079840	GO:0050790	regulation of catalytic activity
Oil6	Glyma.06g281500	47086024	47086296	GO:0009409	response to cold
Oil6	Glyma.06g281500	47086024	47086296	GO:0009733	response to auxin stimulus
Oil6	Glyma.06g281500	47086024	47086296	GO:0042742	defense response to bacterium
Oil6	Glyma.06g283200	47114654	47119627	GO:0006094	gluconeogenesis
Oil6	Glyma.06g283200	47114654	47119627	GO:0006096	glycolysis
Oil6	Glyma.06g283200	47114654	47119627	GO:0006164	purine nucleotide biosynthetic process
Oil6	Glyma.06g283200	47114654	47119627	GO:0008652	cellular amino acid biosynthetic process
Oil6	Glyma.06g283200	47114654	47119627	GO:0009082	branched chain family amino acid biosynthetic process
Oil6	Glyma.06g283200	47114654	47119627	GO:0009651	response to salt stress
Oil6	Glyma.06g283200	47114654	47119627	GO:0046686	response to cadmium ion
Oil6	Glyma.06g283200	47114654	47119627	GO:0055114	oxidation-reduction process
Oil6	Glyma.06g283800	47191044	47195889	GO:0008150	biological_process
Oil6	Glyma.06g284700	47313416	47315646	GO:0008150	biological_process
Oil6	Glyma.06g284900	47339957	47343650	GO:0006355	regulation of transcription DNA-dependent
Oil6	Glyma.06g284900	47339957	47343650	GO:0009410	response to xenobiotic stimulus
Oil6	Glyma.06g284900	47339957	47343650	GO:0030968	endoplasmic reticulum unfolded protein response
Oil6	Glyma.06g284900	47339957	47343650	GO:0045893	positive regulation of transcription DNA-dependent
Oil6	Glyma.06g285500	47406908	47412676	GO:0006952	defense response
Oil6	Glyma.06g285500	47406908	47412676	GO:0007165	signal transduction
Oil6	Glyma.06g285600	47428801	47429895	GO:0009407	toxin catabolic process
Oil6	Glyma.06g285600	47428801	47429895	GO:0010583	response to cyclopentenone
Oil6	Glyma.06g285700	47431347	47433567	GO:0008152	metabolic process
Oil6	Glyma.06g285800	47436638	47441549	GO:0009610	response to symbiotic fungus
Oil6	Glyma.06g285800	47436638	47441549	GO:0045893	positive regulation of transcription DNA-dependent
Oil6	Glyma.06g285800	47436638	47441549	GO:0048574	long-day photoperiodism flowering
Oil6	Glyma.06g285800	47436638	47441549	GO:0048578	positive regulation of long-day photoperiodism flowering
Oil6; Prot6.2	Glyma.06g286300	47468679	47468996	GO:0008150	biological_process
Oil6; Prot6.2	Glyma.06g286800	47558508	47565529	GO:0002679	respiratory burst involved in defense response
Oil6; Prot6.2	Glyma.06g286800	47558508	47565529	GO:0006857	oligopeptide transport
Oil6; Prot6.2	Glyma.06g286800	47558508	47565529	GO:0006865	amino acid transport
Oil6; Prot6.2	Glyma.06g286800	47558508	47565529	GO:0006952	defense response
Oil6; Prot6.2	Glyma.06g286800	47558508	47565529	GO:0008219	cell death
Oil6; Prot6.2	Glyma.06g286800	47558508	47565529	GO:0009407	toxin catabolic process
Oil6; Prot6.2	Glyma.06g286800	47558508	47565529	GO:0009817	defense response to fungus incompatible interaction
Oil6; Prot6.2	Glyma.06g286800	47558508	47565529	GO:0010150	leaf senescence
Oil6; Prot6.2	Glyma.06g286800	47558508	47565529	GO:0010200	response to chitin
Oil6; Prot6.2	Glyma.06g286800	47558508	47565529	GO:0010583	response to cyclopentenone
Oil6; Prot6.2	Glyma.06g286800	47558508	47565529	GO:0015824	proline transport
Oil6; Prot6.2	Glyma.06g287200	47607955	47611348	GO:0006355	regulation of transcription DNA-dependent
Oil6; Prot6.2	Glyma.06g287200	47607955	47611348	GO:0006826	iron ion transport
Oil6; Prot6.2	Glyma.06g287200	47607955	47611348	GO:0010106	cellular response to iron ion starvation
Oil6; Prot6.2	Glyma.06g287200	47607955	47611348	GO:0010167	response to nitrate
Oil6; Prot6.2	Glyma.06g287200	47607955	47611348	GO:0010260	organ senescence
Oil6; Prot6.2	Glyma.06g287200	47607955	47611348	GO:0015706	nitrate transport
Oil6; Prot6.2	Glyma.06g287200	47607955	47611348	GO:0045893	positive regulation of transcription DNA-dependent
Oil6; Prot6.2	Glyma.06g287200	47607955	47611348	GO:0048510	regulation of timing of transition from vegetative to reproductive phase
Oil6; Prot6.2	Glyma.06g287200	47607955	47611348	GO:0048573	photoperiodism flowering
Oil6; Prot6.2	Glyma.06g287500	47646262	47646552	GO:0008150	biological_process
Oil6; Prot6.2	Glyma.06g288500	47712486	47716034	GO:0006355	regulation of transcription DNA-dependent
Oil6; Prot6.2	Glyma.06g288500	47712486	47716034	GO:0007275	multicellular organismal development
Oil6; Prot6.2	Glyma.06g288500	47712486	47716034	GO:0010413	glucuronoxylan metabolic process
Oil6; Prot6.2	Glyma.06g288500	47712486	47716034	GO:0045492	xylan biosynthetic process
Oil6; Prot6.2	Glyma.06g288500	47712486	47716034	GO:0045893	positive regulation of transcription DNA-dependent
Oil6; Prot6.2	Glyma.06g288500	47712486	47716034	GO:2000652	regulation of secondary cell wall biogenesis
Oil6; Prot6.2	Glyma.06g288900	47747596	47762326	GO:0006355	regulation of transcription DNA-dependent
Oil6; Prot6.2	Glyma.06g288900	47747596	47762326	GO:0009630	gravitropism
Oil6; Prot6.2	Glyma.06g288900	47747596	47762326	GO:0010228	vegetative to reproductive phase transition of meristem
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0000165	MAPKKK cascade
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0006355	regulation of transcription DNA-dependent
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0006364	rRNA processing

Table 4.4 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0006612	protein targeting to membrane
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0007165	signal transduction
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0009617	response to bacterium
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0009658	chloroplast organization
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0009862	systemic acquired resistance salicylic acid mediated signaling pathway
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0009867	jasmonic acid mediated signaling pathway
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0009902	chloroplast relocation
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0009910	negative regulation of flower development
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0009965	leaf morphogenesis
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0010027	thylakoid membrane organization
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0010207	photosystem II assembly
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0010310	regulation of hydrogen peroxide metabolic process
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0010363	regulation of plant-type hypersensitive response
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0010380	regulation of chlorophyll biosynthetic process
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0030154	cell differentiation
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0031348	negative regulation of defense response
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0035304	regulation of protein dephosphorylation
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0042793	transcription from plastid promoter
Oil6; Prot6.2	Glyma.06g289300	47797509	47802933	GO:0045893	positive regulation of transcription DNA-dependent
Oil6; Prot6.2	Glyma.06g289600	47831855	47835534	GO:0006508	proteolysis
Oil6; Prot6.2	Glyma.06g290000	47878148	47879505	GO:0006355	regulation of transcription DNA-dependent
Oil6; Prot6.2	Glyma.06g290000	47878148	47879505	GO:0006857	oligopeptide transport
Oil6; Prot6.2	Glyma.06g290000	47878148	47879505	GO:0009873	ethylene mediated signaling pathway
Oil6; Prot6.2	Glyma.06g290000	47878148	47879505	GO:0045892	negative regulation of transcription DNA-dependent
Oil6; Prot6.2	Glyma.06g290200	47906561	47909375	GO:0009627	systemic acquired resistance
Oil6; Prot6.2	Glyma.06g290200	47906561	47909375	GO:0031347	regulation of defense response
Oil6; Prot6.2	Glyma.06g290200	47906561	47909375	GO:0031348	negative regulation of defense response
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0000226	microtubule cytoskeleton organization
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0000280	nuclear division
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0000911	cytokinesis by cell plate formation
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0006270	DNA-dependent DNA replication initiation
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0006275	regulation of DNA replication
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0006306	DNA methylation
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0006342	chromatin silencing
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0008283	cell proliferation
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0009909	regulation of flower development
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0010389	regulation of G2/M transition of mitotic cell cycle
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0016458	gene silencing
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0016572	histone phosphorylation
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0031047	gene silencing by RNA
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0034968	histone lysine methylation
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0042023	DNA endoreduplication
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0042127	regulation of cell proliferation
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0045010	actin nucleation
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0051225	spindle assembly
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0051258	protein polymerization
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0051567	histone H3-K9 methylation
Oil6; Prot6.2	Glyma.06g290300	47919790	47924970	GO:0051726	regulation of cell cycle
Oil6; Prot6.2	Glyma.06g291300	48016841	48019943	GO:0048446	petal morphogenesis
Oil6; Prot6.2	Glyma.06g292300	48121149	48122369	GO:0006661	phosphatidylinositol biosynthetic process
Oil6; Prot6.2	Glyma.06g292300	48121149	48122369	GO:0008150	biological_process
Oil6; Prot6.2	Glyma.06g292400	48130296	48134955	GO:0005975	carbohydrate metabolic process
Oil6; Prot6.2	Glyma.06g292900	48170133	48173010	GO:0008150	biological_process
Oil6; Prot6.2	Glyma.06g293000	48178520	48179509	GO:0008150	biological_process
Oil6; Prot6.2	Glyma.06g293600	48251255	48252791	GO:0008152	metabolic process
Oil6; Prot6.2	Glyma.06g294100	48301296	48305950	GO:0008150	biological_process
Oil6; Prot6.2	Glyma.06g294200	48309187	48316006	GO:0008150	biological_process
Oil6; Prot6.2	Glyma.06g294300	48326079	48329001	GO:0008150	biological_process
Oil6; Prot6.2	Glyma.06g294400	48336232	48340519	GO:0008150	biological_process
Oil6; Prot6.2	Glyma.06g294800	48376832	48379482	GO:0006629	lipid metabolic process
Oil6; Prot6.2	Glyma.06g294800	48376832	48379482	GO:0019761	glucosinolate biosynthetic process
Oil6; Prot6.2	Glyma.06g294800	48376832	48379482	GO:0019953	sexual reproduction
Prot6.2	Glyma.06g295700	48460473	48468376	GO:0006598	polyamine catabolic process
Prot6.2	Glyma.06g295700	48460473	48468376	GO:0009611	response to wounding
Prot6.2	Glyma.06g295700	48460473	48468376	GO:0009698	phenylpropanoid metabolic process
Prot6.2	Glyma.06g295700	48460473	48468376	GO:0009805	coumarin biosynthetic process
Prot6.2	Glyma.06g295700	48460473	48468376	GO:0009809	lignin biosynthetic process
Prot6.2	Glyma.06g295700	48460473	48468376	GO:0009963	positive regulation of flavonoid biosynthetic process
Prot6.2	Glyma.06g295700	48460473	48468376	GO:0016126	sterol biosynthetic process
Prot6.2	Glyma.06g295700	48460473	48468376	GO:0042398	cellular modified amino acid biosynthetic process
Prot6.2	Glyma.06g295700	48460473	48468376	GO:0051555	flavonol biosynthetic process
Prot6.2	Glyma.06g297000	48622864	48625532	GO:0006606	protein import into nucleus
Prot6.2	Glyma.06g297000	48622864	48625532	GO:0031120	snRNA pseudouridine synthesis
Prot6.2	Glyma.06g297000	48622864	48625532	GO:0042254	ribosome biogenesis
Prot6.2	Glyma.06g298300	48720151	48724708	GO:0055114	oxidation-reduction process
Prot6.2	Glyma.06g298600	48741426	48742170	GO:0009611	response to wounding
Prot6.2	Glyma.06g298700	48745551	48746297	GO:0009611	response to wounding
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0006364	rRNA processing
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0006417	regulation of translation
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0006655	phosphatidylglycerol biosynthetic process
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0007186	G-protein coupled receptor protein signaling pathway
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0009657	plastid organization
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0009773	photosynthetic electron transport in photosystem I
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0009902	chloroplast relocation
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0010027	thylakoid membrane organization

Table 4.4 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0010182	sugar mediated signaling pathway
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0010207	photosystem II assembly
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0015979	photosynthesis
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0019288	isopentenyl diphosphate biosynthetic process mevalonate-independent pathway
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0034660	ncRNA metabolic process
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0035304	regulation of protein dephosphorylation
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0042742	defense response to bacterium
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0042793	transcription from plastid promoter
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0045037	protein import into chloroplast stroma
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0045038	protein import into chloroplast thylakoid membrane
Prot6.2	Glyma.06g299100	48765626	48769670	GO:0045893	positive regulation of transcription DNA-dependent
Prot6.2	Glyma.06g299200	48770091	48776577	GO:0007205	activation of protein kinase C activity by G-protein coupled receptor protein signaling pathway
Prot6.2	Glyma.06g299200	48770091	48776577	GO:0009723	response to ethylene stimulus
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0006355	regulation of transcription DNA-dependent
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0006970	response to osmotic stress
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0007165	signal transduction
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0009414	response to water deprivation
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0009611	response to wounding
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0009651	response to salt stress
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0009723	response to ethylene stimulus
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0009733	response to auxin stimulus
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0009737	response to abscisic acid stimulus
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0009738	abscisic acid mediated signaling pathway
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0009753	response to jasmonic acid stimulus
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0015824	proline transport
Prot6.2	Glyma.06g299300	48784796	48786545	GO:0042538	hyperosmotic salinity response
Prot6.2	Glyma.06g299600	48825808	48829252	GO:0006810	transport
Prot6.2	Glyma.06g299600	48825808	48829252	GO:0015858	nucleoside transport
Prot6.2	Glyma.06g299600	48825808	48829252	GO:0015864	pyrimidine nucleoside transport
Prot6.2	Glyma.06g299700	48844926	48847912	GO:0000278	mitotic cell cycle
Prot6.2	Glyma.06g299700	48844926	48847912	GO:0010440	stomatal lineage progression
Prot6.2	Glyma.06g299700	48844926	48847912	GO:0042023	DNA endoreduplication
Prot6.2	Glyma.06g299700	48844926	48847912	GO:0045736	negative regulation of cyclin-dependent protein kinase activity
Prot6.2	Glyma.06g299700	48844926	48847912	GO:0051726	regulation of cell cycle
Prot6.2	Glyma.06g299900	48860250	48861922	GO:0006355	regulation of transcription DNA-dependent
Prot6.2	Glyma.06g299900	48860250	48861922	GO:0009651	response to salt stress
Prot6.2	Glyma.06g299900	48860250	48861922	GO:0009723	response to ethylene stimulus
Prot6.2	Glyma.06g299900	48860250	48861922	GO:0009733	response to auxin stimulus
Prot6.2	Glyma.06g299900	48860250	48861922	GO:0009751	response to salicylic acid stimulus
Prot6.2	Glyma.06g299900	48860250	48861922	GO:0009753	response to jasmonic acid stimulus
Prot6.2	Glyma.06g300000	48880589	48882272	GO:0006355	regulation of transcription DNA-dependent
Prot6.2	Glyma.06g300000	48880589	48882272	GO:0009809	lignin biosynthetic process
Prot6.2	Glyma.06g300000	48880589	48882272	GO:0045893	positive regulation of transcription DNA-dependent
Prot6.2	Glyma.06g300000	48880589	48882272	GO:2000652	regulation of secondary cell wall biogenesis
Prot6.2	Glyma.06g300600	48955701	48958858	GO:0006468	protein phosphorylation
Prot6.2	Glyma.06g300600	48955701	48958858	GO:0048544	recognition of pollen
Prot6.2	Glyma.06g300700	48960266	48962892	GO:0008150	biological_process
Prot6.2	Glyma.06g300800	48965413	48969623	GO:0046777	protein autophosphorylation
Prot6.2	Glyma.06g301000	49001317	49005605	GO:0009733	response to auxin stimulus
Prot6.2	Glyma.06g301000	49001317	49005605	GO:0009734	auxin mediated signaling pathway
Prot6.2	Glyma.06g301000	49001317	49005605	GO:0009826	unidimensional cell growth
Prot6.2	Glyma.06g301000	49001317	49005605	GO:0010252	auxin homeostasis
Prot6.2	Glyma.06g301000	49001317	49005605	GO:0010583	response to cyclopentenone
Prot6.2	Glyma.06g301200	49019337	49020137	GO:0008150	biological_process
Prot6.2	Glyma.06g301700	49058716	49065196	GO:0009860	pollen tube growth
Prot6.2	Glyma.06g301700	49058716	49065196	GO:0016043	cellular component organization
Prot6.2	Glyma.06g301700	49058716	49065196	GO:0016482	cytoplasmic transport
Prot6.2	Glyma.06g301700	49058716	49065196	GO:0030036	actin cytoskeleton organization
Prot6.2	Glyma.06g301700	49058716	49065196	GO:0045010	actin nucleation
Prot6.2	Glyma.06g301700	49058716	49065196	GO:0051017	actin filament bundle assembly
Prot6.2	Glyma.06g301800	49066568	49068108	GO:0008150	biological_process
Prot6.2	Glyma.06g301900	49073240	49092983	GO:0009791	post-embryonic development
Prot6.2	Glyma.06g301900	49073240	49092983	GO:0010228	vegetative to reproductive phase transition of meristem
Prot6.2	Glyma.06g301900	49073240	49092983	GO:0048440	carpel development
Prot6.2	Glyma.06g301900	49073240	49092983	GO:0048443	stamen development
Prot6.2	Glyma.06g302100	49110296	49114225	GO:0009416	response to light stimulus
Prot6.2	Glyma.06g302300	49159813	49162435	GO:0005975	carbohydrate metabolic process
Prot6.2	Glyma.06g302300	49159813	49162435	GO:0006073	cellular glucan metabolic process
Prot6.2	Glyma.06g302300	49159813	49162435	GO:0010075	regulation of meristem growth
Prot6.2	Glyma.06g303900	49338422	49340244	GO:0006865	amino acid transport
Prot6.2	Glyma.06g303900	49338422	49340244	GO:0008150	biological_process
Prot6.2	Glyma.06g304100	49348843	49350580	GO:0006865	amino acid transport
Prot6.2	Glyma.06g304100	49348843	49350580	GO:0008150	biological_process
Prot6.2	Glyma.06g304400	49364042	49364550	GO:0008150	biological_process
Prot6.2	Glyma.06g304600	49371143	49372980	GO:0008150	biological_process

Table 4.5 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 7 based on Glyma 2.0 position.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Prot7	Glyma.07g082000	7534011	7536705	GO:0008152	metabolic process
Prot7	Glyma.07g082000	7534011	7536705	GO:0055114	oxidation-reduction process
Prot7	Glyma.07g082200	7549036	7554463	GO:0000956	nuclear-transcribed mRNA catabolic process
Prot7	Glyma.07g082200	7549036	7554463	GO:0006487	protein N-linked glycosylation
Prot7	Glyma.07g082200	7549036	7554463	GO:0048573	photoperiodism flowering
Prot7	Glyma.07g082500	7568946	7570260	GO:0008152	metabolic process
Prot7	Glyma.07g082800	7609774	7613605	GO:0015937	coenzyme A biosynthetic process
Prot7	Glyma.07g082800	7609774	7613605	GO:0042538	hyperosmotic salinity response
Prot7	Glyma.07g082900	7616974	7620509	GO:0000398	nuclear mRNA splicing via spliceosome
Prot7	Glyma.07g082900	7616974	7620509	GO:0008150	biological_process
Prot7	Glyma.07g082900	7616974	7620509	GO:0051604	protein maturation
Prot7	Glyma.07g083100	7649197	7652550	GO:0008150	biological_process
Prot7	Glyma.07g083200	7657310	7668139	GO:0055114	oxidation-reduction process
Prot7	Glyma.07g083400	7685748	7694834	GO:0055114	oxidation-reduction process
Prot7	Glyma.07g083500	7716572	7720978	GO:0006355	regulation of transcription DNA-dependent
Prot7	Glyma.07g084200	7759819	7763556	GO:0006184	GTP catabolic process
Prot7	Glyma.07g084200	7759819	7763556	GO:0007015	actin filament organization
Prot7	Glyma.07g084200	7759819	7763556	GO:0007165	signal transduction
Prot7	Glyma.07g084200	7759819	7763556	GO:0007264	small GTPase mediated signal transduction
Prot7	Glyma.07g084200	7759819	7763556	GO:0015031	protein transport
Prot7	Glyma.07g084300	7770933	7773893	GO:0000079	regulation of cyclin-dependent protein kinase activity
Prot7	Glyma.07g084300	7770933	7773893	GO:0000278	mitotic cell cycle
Prot7	Glyma.07g084300	7770933	7773893	GO:0006468	protein phosphorylation
Prot7	Glyma.07g084300	7770933	7773893	GO:0008152	metabolic process
Prot7	Glyma.07g084300	7770933	7773893	GO:0010078	maintenance of root meristem identity
Prot7	Glyma.07g084300	7770933	7773893	GO:0010440	stomatal lineage progression
Prot7	Glyma.07g084300	7770933	7773893	GO:0042023	DNA endoreduplication
Prot7	Glyma.07g084300	7770933	7773893	GO:0045736	negative regulation of cyclin-dependent protein kinase activity
Prot7	Glyma.07g084500	7788320	7793377	GO:0008150	biological_process
Prot7	Glyma.07g084900	7812324	7816700	GO:0008150	biological_process
Prot7	Glyma.07g085200	7845887	7852685	GO:0006631	fatty acid metabolic process
Prot7	Glyma.07g085200	7845887	7852685	GO:0006635	fatty acid beta-oxidation
Prot7	Glyma.07g085200	7845887	7852685	GO:0007275	multicellular organismal development
Prot7	Glyma.07g085200	7845887	7852685	GO:0008152	metabolic process
Prot7	Glyma.07g085200	7845887	7852685	GO:0009695	jasmonic acid biosynthetic process
Prot7	Glyma.07g085200	7845887	7852685	GO:0009845	seed germination
Prot7	Glyma.07g085200	7845887	7852685	GO:0009908	flower development
Prot7	Glyma.07g085200	7845887	7852685	GO:0055114	oxidation-reduction process
Prot7	Glyma.07g085400	7870356	7876333	GO:0008033	tRNA processing
Prot7	Glyma.07g085400	7870356	7876333	GO:0009658	chloroplast organization
Prot7	Glyma.07g085400	7870356	7876333	GO:0009793	embryo development ending in seed dormancy
Prot7	Glyma.07g085400	7870356	7876333	GO:0010098	suspensor development
Prot7	Glyma.07g086600	8024187	8029402	GO:0008150	biological_process
Prot7	Glyma.07g086700	8031061	8031833	GO:0008150	biological_process
Prot7	Glyma.07g086800	8032435	8034515	GO:0008150	biological_process
Prot7	Glyma.07g087000	8053652	8055911	GO:0000902	cell morphogenesis
Prot7	Glyma.07g087000	8053652	8055911	GO:0008150	biological_process
Prot7	Glyma.07g087000	8053652	8055911	GO:0016049	cell growth
Prot7	Glyma.07g087000	8053652	8055911	GO:0048193	Golgi vesicle transport
Prot7	Glyma.07g087300	8067710	8068385	GO:0002237	response to molecule of bacterial origin
Prot7	Glyma.07g087300	8067710	8068385	GO:0002679	respiratory burst involved in defense response
Prot7	Glyma.07g087300	8067710	8068385	GO:0006979	response to oxidative stress
Prot7	Glyma.07g087300	8067710	8068385	GO:0008150	biological_process
Prot7	Glyma.07g087300	8067710	8068385	GO:0009611	response to wounding
Prot7	Glyma.07g087300	8067710	8068385	GO:0010200	response to chitin
Prot7	Glyma.07g087300	8067710	8068385	GO:0030968	endoplasmic reticulum unfolded protein response
Prot7	Glyma.07g087300	8067710	8068385	GO:0035556	intracellular signal transduction
Prot7	Glyma.07g087300	8067710	8068385	GO:0050832	defense response to fungus
Prot7	Glyma.07g087600	8103388	8104969	GO:0008152	metabolic process
Prot7	Glyma.07g087600	8103388	8104969	GO:0009627	systemic acquired resistance
Prot7	Glyma.07g087600	8103388	8104969	GO:0009699	phenylpropanoid biosynthetic process
Prot7	Glyma.07g087600	8103388	8104969	GO:0010167	response to nitrate
Prot7	Glyma.07g087600	8103388	8104969	GO:0015706	nitrate transport
Prot7	Glyma.07g087600	8103388	8104969	GO:0034976	response to endoplasmic reticulum stress
Prot7	Glyma.07g087600	8103388	8104969	GO:0055114	oxidation-reduction process
Prot7	Glyma.07g087800	8125996	8135514	GO:0009827	plant-type cell wall modification
Prot7	Glyma.07g087800	8125996	8135514	GO:0009846	pollen germination
Prot7	Glyma.07g087800	8125996	8135514	GO:0009860	pollen tube growth
Prot7	Glyma.07g087800	8125996	8135514	GO:0032012	regulation of ARF protein signal transduction
Prot7	Glyma.07g087800	8125996	8135514	GO:0050790	regulation of catalytic activity
Prot7	Glyma.07g088100	8155777	8157594	GO:0006412	translation
Prot7	Glyma.07g089000	8296453	8305007	GO:0006306	DNA methylation
Prot7	Glyma.07g089000	8296453	8305007	GO:0006342	chromatin silencing
Prot7	Glyma.07g089000	8296453	8305007	GO:0009409	response to cold
Prot7	Glyma.07g089000	8296453	8305007	GO:0009640	photomorphogenesis

Table 4.5 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Prot7	Glyma.07g089000	8296453	8305007	GO:0009793	embryo development ending in seed dormancy
Prot7	Glyma.07g089000	8296453	8305007	GO:0009845	seed germination
Prot7	Glyma.07g089000	8296453	8305007	GO:0009909	regulation of flower development
Prot7	Glyma.07g089000	8296453	8305007	GO:0009933	meristem structural organization
Prot7	Glyma.07g089000	8296453	8305007	GO:0010048	vernalization response
Prot7	Glyma.07g089000	8296453	8305007	GO:0010162	seed dormancy
Prot7	Glyma.07g089000	8296453	8305007	GO:0010182	sugar mediated signaling pathway
Prot7	Glyma.07g089000	8296453	8305007	GO:0010228	vegetative to reproductive phase transition of meristem
Prot7	Glyma.07g089000	8296453	8305007	GO:0010388	cullin deneddylation
Prot7	Glyma.07g089000	8296453	8305007	GO:0016567	protein ubiquitination
Prot7	Glyma.07g089000	8296453	8305007	GO:0016568	chromatin modification
Prot7	Glyma.07g089000	8296453	8305007	GO:0016571	histone methylation
Prot7	Glyma.07g089000	8296453	8305007	GO:0016572	histone phosphorylation
Prot7	Glyma.07g089000	8296453	8305007	GO:0016579	protein deubiquitination
Prot7	Glyma.07g089000	8296453	8305007	GO:0019915	lipid storage
Prot7	Glyma.07g089000	8296453	8305007	GO:0031047	gene silencing by RNA
Prot7	Glyma.07g089000	8296453	8305007	GO:0045814	negative regulation of gene expression epigenetic
Prot7	Glyma.07g089000	8296453	8305007	GO:0045893	positive regulation of transcription DNA-dependent
Prot7	Glyma.07g089000	8296453	8305007	GO:0048572	short-day photoperiodism
Prot7	Glyma.07g089000	8296453	8305007	GO:0048575	short-day photoperiodism flowering
Prot7	Glyma.07g089000	8296453	8305007	GO:0050826	response to freezing
Prot7	Glyma.07g089000	8296453	8305007	GO:0051567	histone H3-K9 methylation
Prot7	Glyma.07g089000	8296453	8305007	GO:0051571	positive regulation of histone H3-K4 methylation
Prot7	Glyma.07g089000	8296453	8305007	GO:0061087	positive regulation of histone H3-K27 methylation
Prot7	Glyma.07g089100	8306577	8310277	GO:0000162	tryptophan biosynthetic process
Prot7	Glyma.07g089200	8312557	8324908	GO:0000023	maltose metabolic process
Prot7	Glyma.07g089200	8312557	8324908	GO:0006098	pentose-phosphate shunt
Prot7	Glyma.07g089200	8312557	8324908	GO:0006364	rRNA processing
Prot7	Glyma.07g089200	8312557	8324908	GO:0009902	chloroplast relocation
Prot7	Glyma.07g089200	8312557	8324908	GO:0010027	thylakoid membrane organization
Prot7	Glyma.07g089200	8312557	8324908	GO:0015979	photosynthesis
Prot7	Glyma.07g089200	8312557	8324908	GO:0015995	chlorophyll biosynthetic process
Prot7	Glyma.07g089200	8312557	8324908	GO:0016117	carotenoid biosynthetic process
Prot7	Glyma.07g089200	8312557	8324908	GO:0016226	iron-sulfur cluster assembly
Prot7	Glyma.07g089200	8312557	8324908	GO:0019252	starch biosynthetic process
Prot7	Glyma.07g089200	8312557	8324908	GO:0019288	isopentenyl diphosphate biosynthetic process mevalonate-independent pathway
Prot7	Glyma.07g089200	8312557	8324908	GO:0034660	ncRNA metabolic process
Prot7	Glyma.07g089200	8312557	8324908	GO:0043085	positive regulation of catalytic activity
Prot7	Glyma.07g090100	8434337	8436847	GO:0016575	histone deacetylation
Prot7	Glyma.07g090100	8434337	8436847	GO:0048364	root development
Prot7	Glyma.07g090100	8434337	8436847	GO:0051568	histone H3-K4 methylation
Prot7	Glyma.07g090100	8434337	8436847	GO:0055114	oxidation-reduction process
Prot7	Glyma.07g091100	8526332	8527857	GO:0006355	regulation of transcription DNA-dependent
Prot7	Glyma.07g091700	8559654	8561002	GO:0007018	microtubule-based movement
Prot7	Glyma.07g091800	8570088	8572245	GO:0008150	biological_process
Prot7	Glyma.07g091900	8587444	8591560	GO:0006412	translation
Prot7	Glyma.07g092400	8617793	8625516	GO:0008150	biological_process
Prot7	Glyma.07g092500	8632559	8633302	GO:0008150	biological_process
Prot7	Glyma.07g092800	8673359	8674772	GO:0006355	regulation of transcription DNA-dependent
Prot7	Glyma.07g093000	8681254	8683612	GO:0006468	protein phosphorylation
Prot7	Glyma.07g093200	8694869	8700465	GO:0000956	nuclear-transcribed mRNA catabolic process
Prot7	Glyma.07g093200	8694869	8700465	GO:0009651	response to salt stress
Prot7	Glyma.07g093200	8694869	8700465	GO:0009737	response to abscisic acid stimulus
Prot7	Glyma.07g093200	8694869	8700465	GO:0010187	negative regulation of seed germination
Prot7	Glyma.07g093700	8723740	8734056	GO:0009793	embryo development ending in seed dormancy
Prot7	Glyma.07g093700	8723740	8734056	GO:0010106	cellular response to iron ion starvation
Prot7	Glyma.07g094100	8787245	8791003	GO:0006468	protein phosphorylation
Prot7	Glyma.07g094200	8792240	8797837	GO:0006468	protein phosphorylation
Prot7	Glyma.07g094500	8810207	8815766	GO:0006468	protein phosphorylation
Prot7	Glyma.07g095000	8843665	8855430	GO:0006468	protein phosphorylation
Prot7	Glyma.07g095200	8889231	8902172	GO:0006468	protein phosphorylation
Prot7	Glyma.07g095400	8914725	8920132	GO:0006468	protein phosphorylation
Prot7	Glyma.07g095500	8934807	8940626	GO:0006468	protein phosphorylation
Prot7	Glyma.07g095500	8934807	8940626	GO:0009620	response to fungus
Prot7	Glyma.07g095500	8934807	8940626	GO:0009691	cytokinin biosynthetic process
Prot7	Glyma.07g095800	8969502	8976133	GO:0006468	protein phosphorylation
Prot7	Glyma.07g095800	8969502	8976133	GO:0009620	response to fungus
Prot7	Glyma.07g095800	8969502	8976133	GO:0009691	cytokinin biosynthetic process
Prot7	Glyma.07g096000	8990337	8991434	GO:0006468	protein phosphorylation
Prot7	Glyma.07g096300	9025738	9030949	GO:0002679	respiratory burst involved in defense response
Prot7	Glyma.07g096300	9025738	9030949	GO:0006468	protein phosphorylation
Prot7	Glyma.07g096300	9025738	9030949	GO:0006984	ER-nucleus signaling pathway
Prot7	Glyma.07g096300	9025738	9030949	GO:0009407	toxin catabolic process
Prot7	Glyma.07g096300	9025738	9030949	GO:0010200	response to chitin
Prot7	Glyma.07g096300	9025738	9030949	GO:0010583	response to cyclopentenone
Prot7	Glyma.07g096300	9025738	9030949	GO:0034976	response to endoplasmic reticulum stress
Prot7	Glyma.07g096600	9057087	9064826	GO:0000719	photoreactive repair
Prot7	Glyma.07g096600	9057087	9064826	GO:0005975	carbohydrate metabolic process
Prot7	Glyma.07g096600	9057087	9064826	GO:0006048	UDP-N-acetylglucosamine biosynthetic process
Prot7	Glyma.07g096600	9057087	9064826	GO:0009411	response to UV
Prot7	Glyma.07g096600	9057087	9064826	GO:0019255	glucose 1-phosphate metabolic process

Table 4.6 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 9 based on Glyma 2.0 position.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Prot9.1	Glyma.09g004200	331820	337460	GO:0008152	metabolic process
Prot9.1	Glyma.09g004200	331820	337460	GO:0009094	L-phenylalanine biosynthetic process
Prot9.1	Glyma.09g004200	331820	337460	GO:0010048	vernalization response
Prot9.1	Glyma.09g004200	331820	337460	GO:0043481	anthocyanin accumulation in tissues in response to UV light
Prot9.1	Glyma.09g004200	331820	337460	GO:0048440	carpel development
Prot9.1	Glyma.09g005900	459265	461797	GO:0008150	biological_process
Prot9.1	Glyma.09g006000	463598	465470	GO:0008150	biological_process
Prot9.1	Glyma.09g006400	476050	478738	GO:0008150	biological_process
Prot9.1	Glyma.09g006800	492341	494466	GO:0006355	regulation of transcription DNA-dependent
Prot9.1	Glyma.09g011700	894858	897607	GO:0008150	biological_process
Prot9.1	Glyma.09g011700	894858	897607	GO:0048446	petal morphogenesis
Prot9.1	Glyma.09g012200	937048	940244	GO:0008150	biological_process
Prot9.1	Glyma.09g012200	937048	940244	GO:0009741	response to brassinosteroid stimulus
Prot9.1	Glyma.09g012200	937048	940244	GO:0015996	chlorophyll catabolic process
Prot9.1	Glyma.09g014600	1107273	1124993	GO:0007067	mitosis
Prot9.1	Glyma.09g014600	1107273	1124993	GO:0008150	biological_process
Prot9.1	Glyma.09g014900	1135061	1144285	GO:0006468	protein phosphorylation
Prot9.1	Glyma.09g014900	1135061	1144285	GO:0009620	response to fungus
Prot9.1	Glyma.09g015400	1174890	1179112	GO:0000226	microtubule cytoskeleton organization
Prot9.1	Glyma.09g015400	1174890	1179112	GO:0000911	cytokinesis by cell plate formation
Prot9.1	Glyma.09g015400	1174890	1179112	GO:0006468	protein phosphorylation
Prot9.1	Glyma.09g015900	1220701	1228294	GO:0009416	response to light stimulus
Prot9.1	Glyma.09g016300	1256220	1258810	GO:0005975	carbohydrate metabolic process
Prot9.1	Glyma.09g016600	1285133	1290884	GO:0006979	response to oxidative stress
Prot9.1	Glyma.09g016600	1285133	1290884	GO:0034484	raffinose catabolic process
Prot9.1	Glyma.09g016600	1285133	1290884	GO:0080167	response to karrikin
Prot9.1	Glyma.09g016800	1300913	1311894	GO:0006402	mRNA catabolic process
Prot9.1	Glyma.09g016800	1300913	1311894	GO:0009791	post-embryonic development
Prot9.1	Glyma.09g016800	1300913	1311894	GO:0009965	leaf morphogenesis
Prot9.1	Glyma.09g016800	1300913	1311894	GO:0010071	root meristem specification
Prot9.1	Glyma.09g016800	1300913	1311894	GO:0010072	primary shoot apical meristem specification
Prot9.1	Glyma.09g016800	1300913	1311894	GO:0031087	deadenylation-independent decapping of nuclear-transcribed mRNA
Prot9.1	Glyma.09g016800	1300913	1311894	GO:0071365	cellular response to auxin stimulus
Prot9.1	Glyma.09g017900	1407056	1409287	GO:0006508	proteolysis
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0000226	microtubule cytoskeleton organization
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0000280	nuclear division
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0000911	cytokinesis by cell plate formation
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0006342	chromatin silencing
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0007000	nucleolus organization
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0009960	endosperm development
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0009965	leaf morphogenesis
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0010342	endosperm cellularization
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0016572	histone phosphorylation
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0048316	seed development
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0048451	petal formation
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0048453	sepal formation
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0051225	spindle assembly
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0051301	cell division
Prot9.1	Glyma.09g018100	1416403	1420543	GO:0051567	histone H3-K9 methylation
Prot9.1	Glyma.09g018600	1460075	1465805	GO:0009416	response to light stimulus
Prot9.1	Glyma.09g018600	1460075	1465805	GO:0055114	oxidation-reduction process
Prot9.1	Glyma.09g018800	1479784	1487258	GO:0006468	protein phosphorylation
Prot9.1	Glyma.09g018800	1479784	1487258	GO:0006569	tryptophan catabolic process
Prot9.1	Glyma.09g018800	1479784	1487258	GO:0009684	indoleacetic acid biosynthetic process
Prot9.1	Glyma.09g018800	1479784	1487258	GO:0030003	cellular cation homeostasis
Prot9.1	Glyma.09g018800	1479784	1487258	GO:0044242	cellular lipid catabolic process
Prot9.1	Glyma.09g018800	1479784	1487258	GO:0070838	divalent metal ion transport
Prot9.1	Glyma.09g018900	1493815	1500130	GO:0006468	protein phosphorylation
Prot9.1	Glyma.09g018900	1493815	1500130	GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway
Prot9.1	Glyma.09g018900	1493815	1500130	GO:0010413	glucuronoxylan metabolic process
Prot9.1	Glyma.09g018900	1493815	1500130	GO:0042546	cell wall biogenesis
Prot9.1	Glyma.09g018900	1493815	1500130	GO:0044036	cell wall macromolecule metabolic process
Prot9.1	Glyma.09g018900	1493815	1500130	GO:0045492	xylan biosynthetic process
Prot9.1	Glyma.09g019100	1505442	1509230	GO:0030244	cellulose biosynthetic process
Prot9.1	Glyma.09g019100	1505442	1509230	GO:0048193	Golgi vesicle transport
Prot9.1	Glyma.09g019300	1514278	1517148	GO:0009639	response to red or far red light
Prot9.1	Glyma.09g019500	1521389	1524233	GO:0006810	transport
Prot9.1	Glyma.09g019500	1521389	1524233	GO:0006886	intracellular protein transport
Prot9.1	Glyma.09g019600	1525323	1528348	GO:0030001	metal ion transport
Prot9.1	Glyma.09g019600	1525323	1528348	GO:0055085	transmembrane transport
Prot9.1	Glyma.09g019700	1530313	1536505	GO:0008150	biological_process
Prot9.1	Glyma.09g019900	1562146	1563912	GO:0043067	regulation of programmed cell death

Table 4.6 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Prot9.1	Glyma.09g020200	1591523	1592970	GO:0042742	defense response to bacterium
Prot9.1	Glyma.09g020200	1591523	1592970	GO:0051607	defense response to virus
Prot9.1	Glyma.09g020500	1614751	1617943	GO:0006952	defense response
Prot9.1	Glyma.09g020500	1614751	1617943	GO:0009627	systemic acquired resistance
Prot9.1	Glyma.09g020500	1614751	1617943	GO:0009697	salicylic acid biosynthetic process
Prot9.1	Glyma.09g020700	1629727	1635164	GO:0006952	defense response
Prot9.1	Glyma.09g020700	1629727	1635164	GO:0009627	systemic acquired resistance
Prot9.1	Glyma.09g020700	1629727	1635164	GO:0009697	salicylic acid biosynthetic process
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0006612	protein targeting to membrane
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0006944	cellular membrane fusion
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0009410	response to xenobiotic stimulus
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0009617	response to bacterium
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0009627	systemic acquired resistance
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0009697	salicylic acid biosynthetic process
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0009816	defense response to bacterium incompatible interaction
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0009817	defense response to fungus incompatible interaction
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0009862	systemic acquired resistance salicylic acid mediated signaling pathway
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0009863	salicylic acid mediated signaling pathway
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0009867	jasmonic acid mediated signaling pathway
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0010200	response to chitin
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0010363	regulation of plant-type hypersensitive response
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0031348	negative regulation of defense response
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0043069	negative regulation of programmed cell death
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0045087	innate immune response
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0050832	defense response to fungus
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0052542	defense response by callose deposition
Prot9.1	Glyma.09g020800	1637560	1643183	GO:0080185	
Prot9.1	Glyma.09g021200	1676351	1689033	GO:0000910	cytokinesis
Prot9.1	Glyma.09g021200	1676351	1689033	GO:0007021	tubulin complex assembly
Prot9.1	Glyma.09g021200	1676351	1689033	GO:0009793	embryo development ending in seed dormancy
Prot9.1	Glyma.09g021400	1694747	1705508	GO:0000956	nuclear-transcribed mRNA catabolic process
Prot9.1	Glyma.09g021400	1694747	1705508	GO:0006635	fatty acid beta-oxidation
Prot9.1	Glyma.09g021400	1694747	1705508	GO:0010048	vernalization response
Prot9.1	Glyma.09g021400	1694747	1705508	GO:0016558	protein import into peroxisome matrix
Prot9.1	Glyma.09g021500	1706246	1712104	GO:0008150	biological_process
Prot9.1	Glyma.09g021500	1706246	1712104	GO:0015979	photosynthesis
Prot9.1	Glyma.09g021600	1720708	1725048	GO:0008150	biological_process
Prot9.1	Glyma.09g021600	1720708	1725048	GO:0009855	determination of bilateral symmetry
Prot9.1	Glyma.09g021600	1720708	1725048	GO:0009944	polarity specification of adaxial/abaxial axis
Prot9.1	Glyma.09g021600	1720708	1725048	GO:0010014	meristem initiation
Prot9.1	Glyma.09g021600	1720708	1725048	GO:0010075	regulation of meristem growth
Prot9.1	Glyma.09g021700	1731533	1733308	GO:0006520	cellular amino acid metabolic process
Prot9.1	Glyma.09g021700	1731533	1733308	GO:0006559	L-phenylalanine catabolic process
Prot9.1	Glyma.09g021700	1731533	1733308	GO:0009611	response to wounding
Prot9.1	Glyma.09g022100	1762850	1765582	GO:0008219	cell death
Prot9.1	Glyma.09g022100	1762850	1765582	GO:0009751	response to salicylic acid stimulus
Prot9.1	Glyma.09g022100	1762850	1765582	GO:0009816	defense response to bacterium incompatible interaction
Prot9.1	Glyma.09g022100	1762850	1765582	GO:0016192	vesicle-mediated transport
Prot9.1	Glyma.09g022100	1762850	1765582	GO:0046836	glycolipid transport
Prot9.1	Glyma.09g022200	1768807	1778364	GO:0000278	mitotic cell cycle
Prot9.1	Glyma.09g022200	1768807	1778364	GO:0006378	mRNA polyadenylation
Prot9.1	Glyma.09g022200	1768807	1778364	GO:0006396	RNA processing
Prot9.1	Glyma.09g022200	1768807	1778364	GO:0006397	mRNA processing
Prot9.1	Glyma.09g022200	1768807	1778364	GO:0006979	response to oxidative stress
Prot9.1	Glyma.09g022200	1768807	1778364	GO:1900363	
Prot9.1	Glyma.09g022500	1805052	1807394	GO:0006979	response to oxidative stress
Prot9.1	Glyma.09g022500	1805052	1807394	GO:0055114	oxidation-reduction process
Prot9.1	Glyma.09g022600	1808637	1810238	GO:0055114	oxidation-reduction process
Prot9.1	Glyma.09g023300	1879980	1884416	GO:0006278	RNA-dependent DNA replication
Prot9.1	Glyma.09g023300	1879980	1884416	GO:0006397	mRNA processing
Prot9.1	Glyma.09g023300	1879980	1884416	GO:0008380	RNA splicing
Len9.1	Glyma.09g027500	2235834	2239074	GO:0006468	protein phosphorylation
Len9.1	Glyma.09g027600	2248230	2255144	GO:0006468	protein phosphorylation
Len9.1	Glyma.09g027700	2258161	2263661	GO:0006468	protein phosphorylation
Len9.1	Glyma.09g027700	2258161	2263661	GO:0009620	response to fungus
Len9.1	Glyma.09g027800	2278104	2289291	GO:0006468	protein phosphorylation
Len9.1	Glyma.09g028100	2314639	2316995	GO:0006865	amino acid transport
Len9.1	Glyma.09g028100	2314639	2316995	GO:0009407	toxin catabolic process
Len9.1	Glyma.09g028100	2314639	2316995	GO:0010583	response to cyclopentenone
Len9.1	Glyma.09g028100	2314639	2316995	GO:0055114	oxidation-reduction process
Len9.1	Glyma.09g028200	2320663	2322174	GO:0006865	amino acid transport
Len9.1	Glyma.09g028200	2320663	2322174	GO:0009407	toxin catabolic process
Len9.1	Glyma.09g028200	2320663	2322174	GO:0010583	response to cyclopentenone
Len9.1	Glyma.09g028200	2320663	2322174	GO:0055114	oxidation-reduction process
Len9.1	Glyma.09g028300	2323505	2325118	GO:0006865	amino acid transport
Len9.1	Glyma.09g028300	2323505	2325118	GO:0009407	toxin catabolic process
Len9.1	Glyma.09g028300	2323505	2325118	GO:0010583	response to cyclopentenone
Len9.1	Glyma.09g028300	2323505	2325118	GO:0055114	oxidation-reduction process
Len9.1	Glyma.09g028600	2341364	2354040	GO:0006260	DNA replication
Len9.1	Glyma.09g028800	2360628	2368329	GO:0006260	DNA replication

Table 4.6 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Len9.1	Glyma.09g028900	2373572	2374252	GO:0006511	ubiquitin-dependent protein catabolic process
Len9.1	Glyma.09g028900	2373572	2374252	GO:0007275	multicellular organismal development
Len9.1	Glyma.09g028900	2373572	2374252	GO:0016567	protein ubiquitination
Len9.1	Glyma.09g031600	2592440	2595781	GO:0005975	carbohydrate metabolic process
Len9.1	Glyma.09g031700	2609839	2612385	GO:0006470	protein dephosphorylation
Len9.1	Glyma.09g032000	2638311	2644265	GO:0006464	protein modification process
Len9.1	Glyma.09g032000	2638311	2644265	GO:0016567	protein ubiquitination
Len9.1	Glyma.09g032000	2638311	2644265	GO:0042787	protein ubiquitination involved in ubiquitin-dependent protein catabolic process
Len9.1	Glyma.09g033900	2834638	2835687	GO:0006139	nucleobase nucleoside nucleotide and nucleic acid metabolic process
Len9.1	Glyma.09g034000	2839448	2840293	GO:0008150	biological_process
Len9.1	Glyma.09g035500	2960395	2967229	GO:0006325	chromatin organization
Len9.1	Glyma.09g035500	2960395	2967229	GO:0006355	regulation of transcription DNA-dependent
Len9.1	Glyma.09g035500	2960395	2967229	GO:0007623	circadian rhythm
Len9.1	Glyma.09g035500	2960395	2967229	GO:0009409	response to cold
Len9.1	Glyma.09g035500	2960395	2967229	GO:0009630	gravitropism
Len9.1	Glyma.09g035500	2960395	2967229	GO:0009638	phototropism
Len9.1	Glyma.09g035500	2960395	2967229	GO:0009640	photomorphogenesis
Len9.1	Glyma.09g035500	2960395	2967229	GO:0009649	entrainment of circadian clock
Len9.1	Glyma.09g035500	2960395	2967229	GO:0009687	abscisic acid metabolic process
Len9.1	Glyma.09g035500	2960395	2967229	GO:0009867	jasmonic acid mediated signaling pathway
Len9.1	Glyma.09g035500	2960395	2967229	GO:0010017	red or far-red light signaling pathway
Len9.1	Glyma.09g035500	2960395	2967229	GO:0010029	regulation of seed germination
Len9.1	Glyma.09g035500	2960395	2967229	GO:0010148	transpiration
Len9.1	Glyma.09g035500	2960395	2967229	GO:0010155	regulation of proton transport
Len9.1	Glyma.09g035500	2960395	2967229	GO:0010161	red light signaling pathway
Len9.1	Glyma.09g035500	2960395	2967229	GO:0010202	response to low fluence red light stimulus
Len9.1	Glyma.09g035500	2960395	2967229	GO:0010218	response to far red light
Len9.1	Glyma.09g035500	2960395	2967229	GO:0010374	stomatal complex development
Len9.1	Glyma.09g035500	2960395	2967229	GO:0010617	circadian regulation of calcium ion oscillation
Len9.1	Glyma.09g035500	2960395	2967229	GO:0015979	photosynthesis
Len9.1	Glyma.09g035500	2960395	2967229	GO:0031347	regulation of defense response
Len9.1	Glyma.09g035500	2960395	2967229	GO:0046685	response to arsenic-containing substance
Len9.1	Glyma.09g035500	2960395	2967229	GO:0046777	protein autophosphorylation
Len9.1	Glyma.09g037900	3163020	3178239	GO:0008150	biological_process
Len9.1	Glyma.09g037900	3163020	3178239	GO:0009630	gravitropism
Len9.1	Glyma.09g037900	3163020	3178239	GO:0009639	response to red or far red light
Len9.1	Glyma.09g039000	3270971	3276530	GO:0008150	biological_process
Len9.1	Glyma.09g039300	3296096	3300443	GO:0009639	response to red or far red light
Len9.1	Glyma.09g041100	3430561	3432082	GO:0006457	protein folding
Len9.1	Glyma.09g041200	3440285	3445911	GO:0006629	lipid metabolic process
Len9.1	Glyma.09g041200	3440285	3445911	GO:0016042	lipid catabolic process
Len9.1	Glyma.09g041300	3456325	3459245	GO:0006952	defense response
Len9.1	Glyma.09g041300	3456325	3459245	GO:0007165	signal transduction
Len9.1	Glyma.09g041300	3456325	3459245	GO:0009416	response to light stimulus
Len9.1	Glyma.09g041300	3456325	3459245	GO:0010114	response to red light
Len9.1	Glyma.09g041300	3456325	3459245	GO:0042742	defense response to bacterium
Len9.1	Glyma.09g043300	3706717	3708193	GO:0008150	biological_process
Len9.1	Glyma.09g043400	3712337	3713601	GO:0009414	response to water deprivation
Len9.1	Glyma.09g043400	3712337	3713601	GO:0009737	response to abscisic acid stimulus
Len9.1	Glyma.09g043400	3712337	3713601	GO:0009790	embryo development
Len9.1	Glyma.09g043700	3722797	3723765	GO:0000038	very long-chain fatty acid metabolic process
Len9.1	Glyma.09g043700	3722797	3723765	GO:0006633	fatty acid biosynthetic process
Len9.1	Glyma.09g043700	3722797	3723765	GO:0008152	metabolic process
Len9.1	Glyma.09g043700	3722797	3723765	GO:0008610	lipid biosynthetic process
Len9.1	Glyma.09g043700	3722797	3723765	GO:0009409	response to cold
Len9.1	Glyma.09g043700	3722797	3723765	GO:0042335	cuticle development
Len9.1	Glyma.09g043800	3724972	3731181	GO:0000278	mitotic cell cycle
Len9.1	Glyma.09g043800	3724972	3731181	GO:0006355	regulation of transcription DNA-dependent
Len9.1	Glyma.09g043800	3724972	3731181	GO:0006396	RNA processing
Len9.1	Glyma.09g043800	3724972	3731181	GO:0006508	proteolysis
Len9.1	Glyma.09g043800	3724972	3731181	GO:0007062	sister chromatid cohesion
Len9.1	Glyma.09g043800	3724972	3731181	GO:0009640	photomorphogenesis
Len9.1	Glyma.09g043800	3724972	3731181	GO:0009749	response to glucose stimulus
Len9.1	Glyma.09g043800	3724972	3731181	GO:0009755	hormone-mediated signaling pathway
Len9.1	Glyma.09g043800	3724972	3731181	GO:0009793	embryo development ending in seed dormancy
Len9.1	Glyma.09g043800	3724972	3731181	GO:0009845	seed germination
Len9.1	Glyma.09g043800	3724972	3731181	GO:0009870	defense response signaling pathway resistance gene-dependent
Len9.1	Glyma.09g043800	3724972	3731181	GO:0009880	embryonic pattern specification
Len9.1	Glyma.09g043800	3724972	3731181	GO:0009909	regulation of flower development
Len9.1	Glyma.09g043800	3724972	3731181	GO:0009933	meristem structural organization
Len9.1	Glyma.09g043800	3724972	3731181	GO:0010072	primary shoot apical meristem specification
Len9.1	Glyma.09g043800	3724972	3731181	GO:0010154	fruit development
Len9.1	Glyma.09g043800	3724972	3731181	GO:0010162	seed dormancy
Len9.1	Glyma.09g043800	3724972	3731181	GO:0010182	sugar mediated signaling pathway
Len9.1	Glyma.09g043800	3724972	3731181	GO:0010204	defense response signaling pathway resistance gene-independent
Len9.1	Glyma.09g043800	3724972	3731181	GO:0010228	vegetative to reproductive phase transition of meristem
Len9.1	Glyma.09g043800	3724972	3731181	GO:0010431	seed maturation
Len9.1	Glyma.09g043800	3724972	3731181	GO:0010564	regulation of cell cycle process
Len9.1	Glyma.09g043800	3724972	3731181	GO:0016567	protein ubiquitination
Len9.1	Glyma.09g043800	3724972	3731181	GO:0019915	lipid storage

Table 4.6 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Len9.1	Glyma.09g043800	3724972	3731181	GO:0042742	defense response to bacterium
Len9.1	Glyma.09g043800	3724972	3731181	GO:0043687	post-translational protein modification
Len9.1	Glyma.09g043800	3724972	3731181	GO:0045595	regulation of cell differentiation
Len9.1	Glyma.09g043800	3724972	3731181	GO:0045892	negative regulation of transcription DNA-dependent
Len9.1	Glyma.09g043800	3724972	3731181	GO:0045893	positive regulation of transcription DNA-dependent
Len9.1	Glyma.09g043800	3724972	3731181	GO:0048364	root development
Len9.1	Glyma.09g043800	3724972	3731181	GO:0048366	leaf development
Len9.1	Glyma.09g043800	3724972	3731181	GO:0048825	cotyledon development
Len9.1	Glyma.09g043800	3724972	3731181	GO:0050826	response to freezing
Len9.1	Glyma.09g043800	3724972	3731181	GO:0050832	defense response to fungus
Len9.1	Glyma.09g043800	3724972	3731181	GO:0051301	cell division
Len9.1	Glyma.09g044000	3755498	3761683	GO:0006457	protein folding
Len9.1	Glyma.09g044100	3764690	3772608	GO:0007067	mitosis
Len9.1	Glyma.09g044100	3764690	3772608	GO:0030261	chromosome condensation
Len9.1	Glyma.09g044300	3782343	3791516	GO:0000278	mitotic cell cycle
Len9.1	Glyma.09g044300	3782343	3791516	GO:0003002	regionalization
Len9.1	Glyma.09g044300	3782343	3791516	GO:0007155	cell adhesion
Len9.1	Glyma.09g044300	3782343	3791516	GO:0009887	organ morphogenesis
Len9.1	Glyma.09g044300	3782343	3791516	GO:0009888	tissue development
Len9.1	Glyma.09g044300	3782343	3791516	GO:0010014	meristem initiation
Len9.1	Glyma.09g044300	3782343	3791516	GO:0010090	trichome morphogenesis
Len9.1	Glyma.09g044300	3782343	3791516	GO:0010228	vegetative to reproductive phase transition of meristem
Len9.1	Glyma.09g044300	3782343	3791516	GO:0010638	positive regulation of organelle organization
Len9.1	Glyma.09g044300	3782343	3791516	GO:0016049	cell growth
Len9.1	Glyma.09g044300	3782343	3791516	GO:0016926	protein desumoylation
Len9.1	Glyma.09g044300	3782343	3791516	GO:0033043	regulation of organelle organization
Len9.1	Glyma.09g044300	3782343	3791516	GO:0033044	regulation of chromosome organization
Len9.1	Glyma.09g044300	3782343	3791516	GO:0045010	actin nucleation
Len9.1	Glyma.09g044300	3782343	3791516	GO:0045595	regulation of cell differentiation
Len9.1	Glyma.09g044300	3782343	3791516	GO:0048449	floral organ formation
Len9.1	Glyma.09g044300	3782343	3791516	GO:0048589	developmental growth
Len9.1	Glyma.09g044300	3782343	3791516	GO:0048765	root hair cell differentiation
Len9.1	Glyma.09g044300	3782343	3791516	GO:0050665	hydrogen peroxide biosynthetic process
Len9.1	Glyma.09g044300	3782343	3791516	GO:0071555	cell wall organization
Len9.1	Glyma.09g044400	3793019	3798367	GO:0006406	mRNA export from nucleus
Len9.1	Glyma.09g044400	3793019	3798367	GO:0006508	proteolysis
Len9.1	Glyma.09g044400	3793019	3798367	GO:0009909	regulation of flower development
Len9.1	Glyma.09g044400	3793019	3798367	GO:0009911	positive regulation of flower development
Len9.1	Glyma.09g044400	3793019	3798367	GO:0010074	maintenance of meristem identity
Len9.1	Glyma.09g044400	3793019	3798367	GO:0010413	glucuronoxylan metabolic process
Len9.1	Glyma.09g044400	3793019	3798367	GO:0016926	protein desumoylation
Len9.1	Glyma.09g044400	3793019	3798367	GO:0045492	xylan biosynthetic process
Len9.1	Glyma.09g044500	3801066	3811353	GO:0006810	transport
Len9.1	Glyma.09g044500	3801066	3811353	GO:0006855	drug transmembrane transport
Len9.1	Glyma.09g044500	3801066	3811353	GO:0009414	response to water deprivation
Len9.1	Glyma.09g044500	3801066	3811353	GO:0009611	response to wounding
Len9.1	Glyma.09g044500	3801066	3811353	GO:0009624	response to nematode
Len9.1	Glyma.09g044500	3801066	3811353	GO:0009627	systemic acquired resistance
Len9.1	Glyma.09g044500	3801066	3811353	GO:0010118	stomatal movement
Len9.1	Glyma.09g044500	3801066	3811353	GO:0034976	response to endoplasmic reticulum stress
Len9.1	Glyma.09g044500	3801066	3811353	GO:0055085	transmembrane transport
Len9.1	Glyma.09g045800	3928530	3933105	GO:0006810	transport
Len9.1	Glyma.09g045800	3928530	3933105	GO:0006839	mitochondrial transport
Len9.1	Glyma.09g045800	3928530	3933105	GO:0030974	thiamine pyrophosphate transport
Len9.1	Glyma.09g046300	3968064	3977175	GO:0008150	biological_process
Len9.1	Glyma.09g046400	3991078	3993254	GO:0008152	metabolic process
Len9.1	Glyma.09g046400	3991078	3993254	GO:0055114	oxidation-reduction process
Len9.1	Glyma.09g046500	4001013	4011424	GO:0000278	mitotic cell cycle
Len9.1	Glyma.09g046500	4001013	4011424	GO:0006396	RNA processing
Len9.1	Glyma.09g046500	4001013	4011424	GO:0006486	protein glycosylation
Len9.1	Glyma.09g046500	4001013	4011424	GO:0006499	N-terminal protein myristoylation
Len9.1	Glyma.09g046500	4001013	4011424	GO:0006511	ubiquitin-dependent protein catabolic process
Len9.1	Glyma.09g046500	4001013	4011424	GO:0006888	ER to Golgi vesicle-mediated transport
Len9.1	Glyma.09g046500	4001013	4011424	GO:0007062	sister chromatid cohesion
Len9.1	Glyma.09g046500	4001013	4011424	GO:0009733	response to auxin stimulus
Len9.1	Glyma.09g046500	4001013	4011424	GO:0009753	response to jasmonic acid stimulus
Len9.1	Glyma.09g046500	4001013	4011424	GO:0009790	embryo development
Len9.1	Glyma.09g046500	4001013	4011424	GO:0009793	embryo development ending in seed dormancy
Len9.1	Glyma.09g046500	4001013	4011424	GO:0009867	jasmonic acid mediated signaling pathway
Len9.1	Glyma.09g046500	4001013	4011424	GO:0009880	embryonic pattern specification
Len9.1	Glyma.09g046500	4001013	4011424	GO:0010072	primary shoot apical meristem specification
Len9.1	Glyma.09g046500	4001013	4011424	GO:0010087	phloem or xylem histogenesis
Len9.1	Glyma.09g046500	4001013	4011424	GO:0010162	seed dormancy
Len9.1	Glyma.09g046500	4001013	4011424	GO:0010265	SCF complex assembly
Len9.1	Glyma.09g046500	4001013	4011424	GO:0010431	seed maturation
Len9.1	Glyma.09g046500	4001013	4011424	GO:0010564	regulation of cell cycle process
Len9.1	Glyma.09g046500	4001013	4011424	GO:0042752	regulation of circadian rhythm
Len9.1	Glyma.09g046500	4001013	4011424	GO:0043090	amino acid import
Len9.1	Glyma.09g046500	4001013	4011424	GO:0045595	regulation of cell differentiation
Len9.1	Glyma.09g046500	4001013	4011424	GO:0048316	seed development

Table 4.6 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Len9.1	Glyma.09g046500	4001013	4011424	GO:0048366	leaf development
Len9.1	Glyma.09g046500	4001013	4011424	GO:0048825	cotyledon development
Len9.1	Glyma.09g046500	4001013	4011424	GO:0051301	cell division
Len9.1	Glyma.09g047400	4085492	4086957	GO:0006863	purine base transport
Len9.1	Glyma.09g047400	4085492	4086957	GO:0015931	nucleobase nucleoside nucleotide and nucleic acid transport
Prot9.2	Glyma.09g169300	39386161	39389901	GO:0006355	regulation of transcription DNA-dependent
Prot9.2	Glyma.09g169300	39386161	39389901	GO:0009737	response to abscisic acid stimulus
Prot9.2	Glyma.09g169300	39386161	39389901	GO:2000652	regulation of secondary cell wall biogenesis
Prot9.2	Glyma.09g169500	39409999	39418798	GO:0006468	protein phosphorylation
Prot9.2	Glyma.09g169600	39420294	39430115	GO:0000956	nuclear-transcribed mRNA catabolic process
Prot9.2	Glyma.09g169600	39420294	39430115	GO:0007346	regulation of mitotic cell cycle
Prot9.2	Glyma.09g169600	39420294	39430115	GO:0008150	biological_process
Prot9.2	Glyma.09g169600	39420294	39430115	GO:0010048	vernalization response
Prot9.2	Glyma.09g169800	39439452	39441906	GO:0009873	ethylene mediated signaling pathway
Prot9.2	Glyma.09g169800	39439452	39441906	GO:0055114	oxidation-reduction process
Prot9.2	Glyma.09g170700	39562228	39566261	GO:0008152	metabolic process
Prot9.2; Pal9	Glyma.09g171700	39651499	39653374	GO:0006970	response to osmotic stress
Prot9.2; Pal9	Glyma.09g171700	39651499	39653374	GO:0009651	response to salt stress
Prot9.2; Pal9	Glyma.09g171700	39651499	39653374	GO:0009658	chloroplast organization
Prot9.2; Pal9	Glyma.09g171700	39651499	39653374	GO:0009737	response to abscisic acid stimulus
Prot9.2; Pal9	Glyma.09g171800	39653896	39655287	GO:0009409	response to cold
Prot9.2; Pal9	Glyma.09g171800	39653896	39655287	GO:0009414	response to water deprivation
Prot9.2; Pal9	Glyma.09g171800	39653896	39655287	GO:0009737	response to abscisic acid stimulus
Prot9.2; Pal9	Glyma.09g171800	39653896	39655287	GO:0042538	hyperosmotic salinity response
Prot9.2; Pal9	Glyma.09g171900	39656849	39660683	GO:0008150	biological_process
Prot9.2; Pal9	Glyma.09g172100	39685047	39686773	GO:0002679	respiratory burst involved in defense response
Prot9.2; Pal9	Glyma.09g172100	39685047	39686773	GO:0010200	response to chitin
Prot9.2; Pal9	Glyma.09g172100	39685047	39686773	GO:0015824	proline transport
Prot9.2; Pal9	Glyma.09g172100	39685047	39686773	GO:0016567	protein ubiquitination
Prot9.2; Pal9	Glyma.09g172100	39685047	39686773	GO:0035556	intracellular signal transduction
Prot9.2; Pal9	Glyma.09g172400	39707036	39709045	GO:0006412	translation
Prot9.2; Pal9	Glyma.09g172400	39707036	39709045	GO:0006783	heme biosynthetic process
Prot9.2; Pal9	Glyma.09g172400	39707036	39709045	GO:0009965	leaf morphogenesis
Prot9.2; Pal9	Glyma.09g172400	39707036	39709045	GO:0016556	mRNA modification
Prot9.2; Pal9	Glyma.09g172400	39707036	39709045	GO:0030154	cell differentiation
Prot9.2; Pal9	Glyma.09g173800	39855229	39857735	GO:0006468	protein phosphorylation
Prot9.2; Pal9	Glyma.09g173800	39855229	39857735	GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway
Prot9.2; Pal9	Glyma.09g173800	39855229	39857735	GO:0010075	regulation of meristem growth
Prot9.2; Pal9	Glyma.09g174200	39883774	39890294	GO:0009611	response to wounding
Prot9.2; Pal9	Glyma.09g174200	39883774	39890294	GO:0009695	jasmonic acid biosynthetic process
Prot9.2; Pal9	Glyma.09g174200	39883774	39890294	GO:0009867	jasmonic acid mediated signaling pathway
Prot9.2; Pal9	Glyma.09g174200	39883774	39890294	GO:0031347	regulation of defense response
Prot9.2; Pal9	Glyma.09g174300	39894789	39896284	GO:0006869	lipid transport
Prot9.2; Pal9	Glyma.09g174400	39897429	39901815	GO:0008535	respiratory chain complex IV assembly
Prot9.2; Pal9	Glyma.09g174900	39929601	39934071	GO:0005975	carbohydrate metabolic process
Prot9.2; Pal9	Glyma.09g174900	39929601	39934071	GO:0006012	galactose metabolic process
Prot9.2; Pal9	Glyma.09g174900	39929601	39934071	GO:0019318	hexose metabolic process
Prot9.2; Pal9	Glyma.09g174900	39929601	39934071	GO:0042732	D-xylose metabolic process
Prot9.2; Pal9	Glyma.09g175100	39945561	39952432	GO:0080156	mitochondrial mRNA modification
Prot9.2; Pal9	Glyma.09g175900	40040733	40043715	GO:0008150	biological_process
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0008361	regulation of cell size
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009165	nucleotide biosynthetic process
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009624	response to nematode
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009630	gravitropism
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009637	response to blue light
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009640	photomorphogenesis
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009733	response to auxin stimulus
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009790	embryo development
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009855	determination of bilateral symmetry
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009887	organ morphogenesis
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009908	flower development
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009926	auxin polar transport
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009944	polarity specification of adaxial/abaxial axis
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0009965	leaf morphogenesis
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0010014	meristem initiation
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0010051	xylem and phloem pattern formation
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0010073	meristem maintenance
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0010089	xylem development
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0010229	inflorescence development
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0010338	leaf formation
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0010358	leaf shaping
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0010583	response to cyclopentenone
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0016310	phosphorylation
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0043481	anthocyanin accumulation in tissues in response to UV light
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0044036	cell wall macromolecule metabolic process
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0048364	root development
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0048367	shoot development
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0048439	flower morphogenesis
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0048443	stamen development
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0048519	negative regulation of biological process

Table 4.6 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0048825	cotyledon development
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0048826	cotyledon morphogenesis
Prot9.2; Pal9	Glyma.09g176300	40077953	40079215	GO:0055085	transmembrane transport
Prot9.2; Pal9	Glyma.09g176800	40118262	40121782	GO:0008150	biological_process
Prot9.2; Pal9	Glyma.09g177600	40214417	40227685	GO:0006468	protein phosphorylation
Prot9.2; Pal9	Glyma.09g177600	40214417	40227685	GO:0009651	response to salt stress
Prot9.2; Pal9	Glyma.09g177600	40214417	40227685	GO:0048573	photoperiodism flowering
Prot9.2; Pal9	Glyma.09g177700	40233549	40238012	GO:0007264	small GTPase mediated signal transduction
Prot9.2; Pal9	Glyma.09g177700	40233549	40238012	GO:0010089	xylem development
Prot9.2; Pal9	Glyma.09g177700	40233549	40238012	GO:0010413	glucuronoxylan metabolic process
Prot9.2; Pal9	Glyma.09g177700	40233549	40238012	GO:0015031	protein transport
Prot9.2; Pal9	Glyma.09g177700	40233549	40238012	GO:0044036	cell wall macromolecule metabolic process
Prot9.2; Pal9	Glyma.09g177700	40233549	40238012	GO:0045492	xylan biosynthetic process
Prot9.2; Pal9	Glyma.09g177800	40242274	40245392	GO:0006355	regulation of transcription DNA-dependent
Prot9.2; Pal9	Glyma.09g177800	40242274	40245392	GO:0048513	organ development
Prot9.2; Pal9	Glyma.09g178000	40279869	40286232	GO:0006355	regulation of transcription DNA-dependent
Prot9.2; Pal9	Glyma.09g178000	40279869	40286232	GO:0009688	abscisic acid biosynthetic process
Prot9.2; Pal9	Glyma.09g178100	40288436	40293420	GO:0005975	carbohydrate metabolic process
Prot9.2; Pal9	Glyma.09g178400	40330140	40331372	GO:0006979	response to oxidative stress
Prot9.2; Pal9	Glyma.09g178600	40342421	40342771	GO:0006091	generation of precursor metabolites and energy
Prot9.2; Pal9	Glyma.09g178600	40342421	40342771	GO:0006354	transcription elongation DNA-dependent
Prot9.2; Pal9	Glyma.09g178600	40342421	40342771	GO:0015979	photosynthesis
Prot9.2; Pal9	Glyma.09g178600	40342421	40342771	GO:0048564	photosystem I assembly
Prot9.2; Pal9	Glyma.09g178700	40344306	40347537	GO:0006855	drug transmembrane transport
Prot9.2; Pal9	Glyma.09g178700	40344306	40347537	GO:0009624	response to nematode
Prot9.2; Pal9	Glyma.09g178700	40344306	40347537	GO:0055085	transmembrane transport
Prot9.2; Pal9	Glyma.09g178800	40354319	40358609	GO:0006855	drug transmembrane transport
Prot9.2; Pal9	Glyma.09g178800	40354319	40358609	GO:0009624	response to nematode
Prot9.2; Pal9	Glyma.09g178800	40354319	40358609	GO:0055085	transmembrane transport
Prot9.2; Pal9	Glyma.09g179000	40373197	40377187	GO:0006855	drug transmembrane transport
Prot9.2; Pal9	Glyma.09g179000	40373197	40377187	GO:0009624	response to nematode
Prot9.2; Pal9	Glyma.09g179000	40373197	40377187	GO:0055085	transmembrane transport
Prot9.2; Pal9	Glyma.09g179600	40442367	40444892	GO:0006094	gluconeogenesis
Prot9.2; Pal9	Glyma.09g179600	40442367	40444892	GO:0006508	proteolysis
Prot9.2; Pal9	Glyma.09g179600	40442367	40444892	GO:0007010	cytoskeleton organization
Prot9.2; Pal9	Glyma.09g179600	40442367	40444892	GO:0009926	auxin polar transport
Prot9.2; Pal9	Glyma.09g179600	40442367	40444892	GO:0010359	regulation of anion channel activity
Prot9.2; Pal9	Glyma.09g179600	40442367	40444892	GO:0010498	proteasomal protein catabolic process
Prot9.2; Pal9	Glyma.09g181300	40620634	40626543	GO:0008150	biological_process
Prot9.2; Pal9	Glyma.09g184400	40955846	40963711	GO:0007018	microtubule-based movement
Prot9.2; Pal9	Glyma.09g184900	41002131	41005811	GO:0000226	microtubule cytoskeleton organization
Prot9.2; Pal9	Glyma.09g184900	41002131	41005811	GO:0016192	vesicle-mediated transport
Prot9.2; Pal9	Glyma.09g185800	41085271	41088053	GO:0016575	histone deacetylation
Prot9.2; Pal9	Glyma.09g185800	41085271	41088053	GO:0048364	root development
Prot9.2; Pal9	Glyma.09g185800	41085271	41088053	GO:0051568	histone H3-K4 methylation
Prot9.2; Pal9	Glyma.09g185800	41085271	41088053	GO:0055114	oxidation-reduction process
Prot9.2; Pal9	Glyma.09g186000	41095580	41099227	GO:0055114	oxidation-reduction process
Prot9.2; Pal9	Glyma.09g186100	41100358	41101236	GO:0009411	response to UV
Prot9.2; Pal9	Glyma.09g186100	41100358	41101236	GO:0009718	anthocyanin biosynthetic process
Prot9.2; Pal9	Glyma.09g186100	41100358	41101236	GO:0009733	response to auxin stimulus
Prot9.2; Pal9	Glyma.09g186100	41100358	41101236	GO:0009744	response to sucrose stimulus
Prot9.2; Pal9	Glyma.09g186100	41100358	41101236	GO:0009813	flavonoid biosynthetic process
Prot9.2; Pal9	Glyma.09g186100	41100358	41101236	GO:0010224	response to UV-B
Prot9.2; Pal9	Glyma.09g186100	41100358	41101236	GO:0055114	oxidation-reduction process
Prot9.2; Pal9	Glyma.09g186400	41114501	41117379	GO:0055114	oxidation-reduction process
Prot9.2; Pal9	Glyma.09g186600	41131553	41141747	GO:0008150	biological_process
Prot9.2; Pal9	Glyma.09g186900	41156622	41158854	GO:0006084	acetyl-CoA metabolic process
Prot9.2; Pal9	Glyma.09g186900	41156622	41158854	GO:0016125	sterol metabolic process
Prot9.2; Pal9	Glyma.09g186900	41156622	41158854	GO:0016126	sterol biosynthetic process
Prot9.2; Pal9	Glyma.09g186900	41156622	41158854	GO:0016132	brassinosteroid biosynthetic process
Prot9.2; Pal9	Glyma.09g186900	41156622	41158854	GO:0019932	second-messenger-mediated signaling
Prot9.2; Pal9	Glyma.09g186900	41156622	41158854	GO:0060964	regulation of gene silencing by miRNA
Prot9.2; Pal9	Glyma.09g187000	41168622	41169926	GO:0008150	biological_process
Prot9.2; Pal9	Glyma.09g187100	41171964	41188123	GO:0006891	intra-Golgi vesicle-mediated transport
Prot9.2; Pal9	Glyma.09g187100	41171964	41188123	GO:0030244	cellulose biosynthetic process
Prot9.2; Pal9	Glyma.09g187100	41171964	41188123	GO:0048193	Golgi vesicle transport
Prot9.2; Pal9	Glyma.09g187300	41195049	41199507	GO:0030163	protein catabolic process
Prot9.2; Pal9	Glyma.09g188700	41339819	41345857	GO:0006810	transport
Prot9.2; Pal9	Glyma.09g188700	41339819	41345857	GO:0008272	sulfate transport
Prot9.2; Pal9	Glyma.09g188700	41339819	41345857	GO:0055085	transmembrane transport
Prot9.2; Pal9	Glyma.09g188900	41373565	41377066	GO:0006355	regulation of transcription DNA-dependent
Prot9.2; Pal9	Glyma.09g188900	41373565	41377066	GO:0043687	post-translational protein modification
Prot9.2; Pal9	Glyma.09g188900	41373565	41377066	GO:0045893	positive regulation of transcription DNA-dependent
Pal9	Glyma.09g189000	41379183	41385121	GO:0009827	plant-type cell wall modification
Pal9	Glyma.09g189000	41379183	41385121	GO:0009846	pollen germination
Pal9	Glyma.09g189000	41379183	41385121	GO:0009860	pollen tube growth
Pal9	Glyma.09g189000	41379183	41385121	GO:0032012	regulation of ARF protein signal transduction
Pal9	Glyma.09g189000	41379183	41385121	GO:0050790	regulation of catalytic activity
Pal9	Glyma.09g189100	41386307	41394858	GO:0000710	meiotic mismatch repair
Pal9	Glyma.09g189100	41386307	41394858	GO:0000911	cytokinesis by cell plate formation

Table 4.6 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Pal9	Glyma.09g189100	41386307	41394858	GO:0006200	ATP catabolic process
Pal9	Glyma.09g189100	41386307	41394858	GO:0006260	DNA replication
Pal9	Glyma.09g189100	41386307	41394858	GO:0006270	DNA-dependent DNA replication initiation
Pal9	Glyma.09g189100	41386307	41394858	GO:0006275	regulation of DNA replication
Pal9	Glyma.09g189100	41386307	41394858	GO:0006298	mismatch repair
Pal9	Glyma.09g189100	41386307	41394858	GO:0006306	DNA methylation
Pal9	Glyma.09g189100	41386307	41394858	GO:0007067	mitosis
Pal9	Glyma.09g189100	41386307	41394858	GO:0007126	meiosis
Pal9	Glyma.09g189100	41386307	41394858	GO:0007129	synapsis
Pal9	Glyma.09g189100	41386307	41394858	GO:0007131	reciprocal meiotic recombination
Pal9	Glyma.09g189100	41386307	41394858	GO:0008283	cell proliferation
Pal9	Glyma.09g189100	41386307	41394858	GO:0009411	response to UV
Pal9	Glyma.09g189100	41386307	41394858	GO:0009909	regulation of flower development
Pal9	Glyma.09g189100	41386307	41394858	GO:0010389	regulation of G2/M transition of mitotic cell cycle
Pal9	Glyma.09g189100	41386307	41394858	GO:0010564	regulation of cell cycle process
Pal9	Glyma.09g189100	41386307	41394858	GO:0016458	gene silencing
Pal9	Glyma.09g189100	41386307	41394858	GO:0034968	histone lysine methylation
Pal9	Glyma.09g189100	41386307	41394858	GO:0043570	maintenance of DNA repeat elements
Pal9	Glyma.09g189100	41386307	41394858	GO:0045910	negative regulation of DNA recombination
Pal9	Glyma.09g189100	41386307	41394858	GO:0051567	histone H3-K9 methylation
Pal9	Glyma.09g189100	41386307	41394858	GO:0051726	regulation of cell cycle
Pal9; Ole9	Glyma.09g190100	41466759	41469357	GO:0008150	biological_process
Pal9; Ole9	Glyma.09g190200	41470908	41473026	GO:0008150	biological_process
Pal9; Ole9	Glyma.09g190400	41482745	41489524	GO:0008150	biological_process
Pal9; Ole9	Glyma.09g190500	41492391	41500159	GO:0007018	microtubule-based movement
Pal9; Ole9	Glyma.09g190900	41563721	41565975	GO:0006810	transport
Pal9; Ole9	Glyma.09g190900	41563721	41565975	GO:0031348	negative regulation of defense response
Pal9; Ole9	Glyma.09g190900	41563721	41565975	GO:0043090	amino acid import
Pal9; Ole9	Glyma.09g190900	41563721	41565975	GO:0055085	transmembrane transport
Ole9; Len9.2	Glyma.09g191400	41612990	41618958	GO:0006486	protein glycosylation
Ole9; Len9.2	Glyma.09g191400	41612990	41618958	GO:0006508	proteolysis
Ole9; Len9.2	Glyma.09g191400	41612990	41618958	GO:0006629	lipid metabolic process
Ole9; Len9.2	Glyma.09g191400	41612990	41618958	GO:0006888	ER to Golgi vesicle-mediated transport
Ole9; Len9.2	Glyma.09g191400	41612990	41618958	GO:0006972	hyperosmotic response
Ole9; Len9.2	Glyma.09g191400	41612990	41618958	GO:0042538	hyperosmotic salinity response
Ole9; Len9.2	Glyma.09g191500	41620935	41627598	GO:0008033	tRNA processing
Ole9; Len9.2	Glyma.09g191500	41620935	41627598	GO:0009658	chloroplast organization
Ole9; Len9.2	Glyma.09g191500	41620935	41627598	GO:0009793	embryo development ending in seed dormancy
Ole9; Len9.2	Glyma.09g191500	41620935	41627598	GO:0010098	suspensor development
Ole9; Len9.2	Glyma.09g191700	41642568	41650351	GO:0006631	fatty acid metabolic process
Ole9; Len9.2	Glyma.09g191700	41642568	41650351	GO:0006635	fatty acid beta-oxidation
Ole9; Len9.2	Glyma.09g191700	41642568	41650351	GO:0007275	multicellular organismal development
Ole9; Len9.2	Glyma.09g191700	41642568	41650351	GO:0008152	metabolic process
Ole9; Len9.2	Glyma.09g191700	41642568	41650351	GO:0009695	jasmonic acid biosynthetic process
Ole9; Len9.2	Glyma.09g191700	41642568	41650351	GO:0009845	seed germination
Ole9; Len9.2	Glyma.09g191700	41642568	41650351	GO:0009908	flower development
Ole9; Len9.2	Glyma.09g191700	41642568	41650351	GO:0055114	oxidation-reduction process
Ole9; Len9.2	Glyma.09g191800	41654391	41657674	GO:0010090	trichome morphogenesis
Ole9; Len9.2	Glyma.09g191800	41654391	41657674	GO:0045010	actin nucleation
Ole9; Len9.2	Glyma.09g192000	41674766	41679084	GO:0008150	biological_process
Ole9; Len9.2	Glyma.09g193200	41775058	41777851	GO:0032259	methylation
Ole9; Len9.2	Glyma.09g193300	41778393	41785291	GO:0006414	translational elongation
Ole9; Len9.2	Glyma.09g193300	41778393	41785291	GO:0009793	embryo development ending in seed dormancy
Ole9; Len9.2	Glyma.09g193300	41778393	41785291	GO:0009902	chloroplast relocation
Ole9; Len9.2	Glyma.09g193300	41778393	41785291	GO:0010027	thylakoid membrane organization
Ole9; Len9.2	Glyma.09g193300	41778393	41785291	GO:0016117	carotenoid biosynthetic process
Ole9; Len9.2	Glyma.09g193300	41778393	41785291	GO:0019288	isopentenyl diphosphate biosynthetic process mevalonate-independent pathway
Ole9; Len9.2	Glyma.09g193300	41778393	41785291	GO:0034660	ncRNA metabolic process
Ole9; Len9.2	Glyma.09g193300	41778393	41785291	GO:0042744	hydrogen peroxide catabolic process
Ole9; Len9.2	Glyma.09g193300	41778393	41785291	GO:0046686	response to cadmium ion
Ole9; Len9.2	Glyma.09g195100	41972424	41978298	GO:0006508	proteolysis
Ole9; Len9.2	Glyma.09g195100	41972424	41978298	GO:0008152	metabolic process
Ole9; Len9.2	Glyma.09g195100	41972424	41978298	GO:0016036	cellular response to phosphate starvation
Ole9; Len9.2	Glyma.09g195100	41972424	41978298	GO:0019375	galactolipid biosynthetic process
Ole9; Len9.2	Glyma.09g195100	41972424	41978298	GO:0042631	cellular response to water deprivation
Ole9; Len9.2	Glyma.09g195100	41972424	41978298	GO:0043086	negative regulation of catalytic activity
Ole9; Len9.2	Glyma.09g195500	42019562	42021201	GO:0008150	biological_process
Ole9; Len9.2	Glyma.09g195700	42036056	42043855	GO:0006346	methylation-dependent chromatin silencing
Ole9; Len9.2	Glyma.09g195700	42036056	42043855	GO:0006355	regulation of transcription DNA-dependent
Ole9; Len9.2	Glyma.09g195700	42036056	42043855	GO:0009630	gravitropism
Ole9; Len9.2	Glyma.09g195700	42036056	42043855	GO:0016246	RNA interference
Ole9; Len9.2	Glyma.09g196100	42073133	42079762	GO:0006486	protein glycosylation
Ole9; Len9.2	Glyma.09g196100	42073133	42079762	GO:0008150	biological_process
Ole9; Len9.2	Glyma.09g196100	42073133	42079762	GO:0030244	cellulose biosynthetic process
Ole9; Len9.2	Glyma.09g196100	42073133	42079762	GO:0048193	Golgi vesicle transport
Ole9; Len9.2	Glyma.09g196200	42081338	42089388	GO:0006360	transcription from RNA polymerase I promoter
Ole9; Len9.2	Glyma.09g196200	42081338	42089388	GO:0008150	biological_process
Ole9; Len9.2	Glyma.09g196400	42100890	42105875	GO:0048446	petal morphogenesis
Ole9; Len9.2	Glyma.09g197000	42169740	42181543	GO:0006487	protein N-linked glycosylation
Ole9; Len9.2	Glyma.09g197000	42169740	42181543	GO:0006987	activation of signaling protein activity involved in unfolded protein response

Table 4.6 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Ole9; Len9.2	Glyma.09g197000	42169740	42181543	GO:0008380	RNA splicing
Ole9; Len9.2	Glyma.09g197000	42169740	42181543	GO:0009751	response to salicylic acid stimulus
Ole9; Len9.2	Glyma.09g197000	42169740	42181543	GO:0009816	defense response to bacterium incompatible interaction
Ole9; Len9.2	Glyma.09g197000	42169740	42181543	GO:0030968	endoplasmic reticulum unfolded protein response
Ole9; Len9.2	Glyma.09g197000	42169740	42181543	GO:0046777	protein autophosphorylation
Ole9; Len9.2	Glyma.09g197100	42181980	42187122	GO:0006810	transport
Ole9; Len9.2	Glyma.09g197100	42181980	42187122	GO:0006874	cellular calcium ion homeostasis
Ole9; Len9.2	Glyma.09g197100	42181980	42187122	GO:0007186	G-protein coupled receptor protein signaling pathway
Ole9; Len9.2	Glyma.09g197100	42181980	42187122	GO:0009416	response to light stimulus
Ole9; Len9.2	Glyma.09g197100	42181980	42187122	GO:0030003	cellular cation homeostasis
Ole9; Len9.2	Glyma.09g197300	42208344	42210282	GO:0006816	calcium ion transport
Ole9; Len9.2	Glyma.09g197300	42208344	42210282	GO:0006874	cellular calcium ion homeostasis
Ole9; Len9.2	Glyma.09g197300	42208344	42210282	GO:0009416	response to light stimulus
Ole9; Len9.2	Glyma.09g197300	42208344	42210282	GO:0009630	gravitropism
Ole9; Len9.2	Glyma.09g197300	42208344	42210282	GO:0030003	cellular cation homeostasis
Ole9; Len9.2	Glyma.09g197300	42208344	42210282	GO:0045087	innate immune response
Ole9; Len9.2	Glyma.09g197300	42208344	42210282	GO:0071230	cellular response to amino acid stimulus
Ole9; Len9.2	Glyma.09g197400	42211661	42216967	GO:0006816	calcium ion transport
Ole9; Len9.2	Glyma.09g197400	42211661	42216967	GO:0006874	cellular calcium ion homeostasis
Ole9; Len9.2	Glyma.09g197400	42211661	42216967	GO:0009416	response to light stimulus
Ole9; Len9.2	Glyma.09g197400	42211661	42216967	GO:0009630	gravitropism
Ole9; Len9.2	Glyma.09g197400	42211661	42216967	GO:0030003	cellular cation homeostasis
Ole9; Len9.2	Glyma.09g197400	42211661	42216967	GO:0045087	innate immune response
Ole9; Len9.2	Glyma.09g197400	42211661	42216967	GO:0071230	cellular response to amino acid stimulus
Ole9; Len9.2	Glyma.09g198000	42254013	42259322	GO:0008150	biological_process
Ole9; Len9.2	Glyma.09g198200	42267017	42278973	GO:0048235	pollen sperm cell differentiation
Ole9; Len9.2	Glyma.09g199100	42355999	42363720	GO:0009657	plastid organization
Ole9; Len9.2	Glyma.09g199100	42355999	42363720	GO:0010020	chloroplast fission
Ole9; Len9.2	Glyma.09g199100	42355999	42363720	GO:0055085	transmembrane transport
Ole9; Len9.2	Glyma.09g199400	42386853	42391877	GO:0008150	biological_process
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0006355	regulation of transcription DNA-dependent
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0009855	determination of bilateral symmetry
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0009887	organ morphogenesis
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0009908	flower development
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0009944	polarity specification of adaxial/abaxial axis
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0010014	meristem initiation
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0010075	regulation of meristem growth
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0010080	regulation of floral meristem growth
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0010089	xylem development
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0010492	maintenance of shoot apical meristem identity
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0035265	organ growth
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0044036	cell wall macromolecule metabolic process
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0048364	root development
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0060771	phyllotactic patterning
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0060772	leaf phyllotactic patterning
Ole9; Len9.2	Glyma.09g199800	42427982	42433197	GO:0060774	auxin mediated signaling pathway involved in phyllotactic patterning
Ole9; Len9.2	Glyma.09g200000	42450808	42454178	GO:0008150	biological_process
Ole9; Len9.2	Glyma.09g200200	42456964	42459986	GO:0008150	biological_process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0000096	sulfur amino acid metabolic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0006546	glycine catabolic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0006636	unsaturated fatty acid biosynthetic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0006655	phosphatidylglycerol biosynthetic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0006733	oxidoreduction coenzyme metabolic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0006766	vitamin metabolic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0008152	metabolic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0008652	cellular amino acid biosynthetic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0009072	aromatic amino acid family metabolic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0009073	aromatic amino acid family biosynthetic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0009106	lipoate metabolic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0009108	coenzyme biosynthetic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0009117	nucleotide metabolic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0009416	response to light stimulus
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0009695	jasmonic acid biosynthetic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0009965	leaf morphogenesis
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0010155	regulation of proton transport
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0015994	chlorophyll metabolic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0015995	chlorophyll biosynthetic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0016114	terpenoid biosynthetic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0016117	carotenoid biosynthetic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0019216	regulation of lipid metabolic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0019288	isopentenyl diphosphate biosynthetic process mevalonate-independent pathway
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0019344	cysteine biosynthetic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0019748	secondary metabolic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0030154	cell differentiation
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0031408	oxylipin biosynthetic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0032880	regulation of protein localization
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0044272	sulfur compound biosynthetic process
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0045893	positive regulation of transcription DNA-dependent
Ole9; Len9.2	Glyma.09g200500	42478937	42487179	GO:0046777	protein autophosphorylation
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0000226	microtubule cytoskeleton organization

Table 4.6 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0000278	mitotic cell cycle
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0000911	cytokinesis by cell plate formation
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0006306	DNA methylation
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0006342	chromatin silencing
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0009909	regulation of flower development
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0016458	gene silencing
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0016570	histone modification
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0016572	histone phosphorylation
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0031047	gene silencing by RNA
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0034968	histone lysine methylation
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0048449	floral organ formation
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0048451	petal formation
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0048453	sepal formation
Ole9; Len9.2	Glyma.09g200800	42501441	42507940	GO:0051567	histone H3-K9 methylation
Ole9; Len9.2	Glyma.09g201000	42517743	42519443	GO:0006355	regulation of transcription DNA-dependent
Ole9; Len9.2	Glyma.09g201000	42517743	42519443	GO:0010052	guard cell differentiation
Ole9; Len9.2	Glyma.09g201000	42517743	42519443	GO:0010118	stomatal movement
Ole9; Len9.2	Glyma.09g201000	42517743	42519443	GO:0045893	positive regulation of transcription DNA-dependent
Ole9; Len9.2	Glyma.09g201000	42517743	42519443	GO:1902066	
Ole9; Len9.2	Glyma.09g201100	42530078	42535148	GO:0006865	amino acid transport
Ole9; Len9.2	Glyma.09g201100	42530078	42535148	GO:0009407	toxin catabolic process
Ole9; Len9.2	Glyma.09g201100	42530078	42535148	GO:0009617	response to bacterium
Ole9; Len9.2	Glyma.09g201100	42530078	42535148	GO:0009626	plant-type hypersensitive response
Ole9; Len9.2	Glyma.09g201100	42530078	42535148	GO:0009809	lignin biosynthetic process
Ole9; Len9.2	Glyma.09g201100	42530078	42535148	GO:0010583	response to cyclopentenone
Ole9; Len9.2	Glyma.09g201100	42530078	42535148	GO:0055114	oxidation-reduction process
Ole9; Len9.2	Glyma.09g201200	42539281	42541604	GO:0009809	lignin biosynthetic process
Ole9; Len9.2	Glyma.09g201200	42539281	42541604	GO:0055114	oxidation-reduction process
Ole9; Len9.2	Glyma.09g203100	42717097	42719742	GO:0008150	biological_process
Ole9; Len9.2	Glyma.09g203300	42740672	42745026	GO:0007020	microtubule nucleation
Ole9; Len9.2	Glyma.09g203300	42740672	42745026	GO:0009733	response to auxin stimulus
Ole9; Len9.2	Glyma.09g203300	42740672	42745026	GO:0009736	cytokinin mediated signaling pathway
Ole9; Len9.2	Glyma.09g203300	42740672	42745026	GO:0009740	gibberellic acid mediated signaling pathway
Ole9; Len9.2	Glyma.09g203300	42740672	42745026	GO:0010162	seed dormancy
Ole9; Len9.2	Glyma.09g203300	42740672	42745026	GO:0010311	lateral root formation
Ole9; Len9.2	Glyma.09g203300	42740672	42745026	GO:0045892	negative regulation of transcription DNA-dependent
Ole9; Len9.2	Glyma.09g207800	43209945	43212158	GO:0043069	negative regulation of programmed cell death
Ole9; Len9.2	Glyma.09g207900	43215179	43225630	GO:0006636	unsaturated fatty acid biosynthetic process
Ole9; Len9.2	Glyma.09g207900	43215179	43225630	GO:0006655	phosphatidylglycerol biosynthetic process
Ole9; Len9.2	Glyma.09g207900	43215179	43225630	GO:0008152	metabolic process
Ole9; Len9.2	Glyma.09g207900	43215179	43225630	GO:0015995	chlorophyll biosynthetic process
Ole9; Len9.2	Glyma.09g207900	43215179	43225630	GO:0016117	carotenoid biosynthetic process
Ole9; Len9.2	Glyma.09g207900	43215179	43225630	GO:0019288	isopentenyl diphosphate biosynthetic process mevalonate-independent pathway
Ole9; Len9.2	Glyma.09g208000	43237559	43243598	GO:0006499	N-terminal protein myristoylation
Ole9; Len9.2	Glyma.09g208000	43237559	43243598	GO:0009855	determination of bilateral symmetry
Ole9; Len9.2	Glyma.09g208000	43237559	43243598	GO:0009944	polarity specification of adaxial/abaxial axis
Ole9; Len9.2	Glyma.09g208000	43237559	43243598	GO:0010014	meristem initiation
Ole9; Len9.2	Glyma.09g208000	43237559	43243598	GO:0010075	regulation of meristem growth
Ole9; Len9.2	Glyma.09g208200	43253360	43258475	GO:0000271	polysaccharide biosynthetic process
Ole9; Len9.2	Glyma.09g208200	43253360	43258475	GO:0009409	response to cold
Ole9; Len9.2	Glyma.09g208200	43253360	43258475	GO:0009832	plant-type cell wall biogenesis
Ole9; Len9.2	Glyma.09g208200	43253360	43258475	GO:0010583	response to cyclopentenone
Ole9; Len9.2	Glyma.09g208200	43253360	43258475	GO:0030244	cellulose biosynthetic process
Ole9; Len9.2	Glyma.09g208400	43279596	43282064	GO:0006457	protein folding
Ole9; Len9.2	Glyma.09g208500	43279609	43291904	GO:0006351	transcription DNA-dependent
Ole9; Len9.2	Glyma.09g208500	43279609	43291904	GO:0006355	regulation of transcription DNA-dependent
Ole9; Len9.2	Glyma.09g208500	43279609	43291904	GO:0006473	protein acetylation
Ole9; Len9.2	Glyma.09g208500	43279609	43291904	GO:0045893	positive regulation of transcription DNA-dependent
Ole9; Len9.2	Glyma.09g208900	43333697	43335901	GO:0006952	defense response
Ole9; Len9.2	Glyma.09g209100	43355642	43360197	GO:0006351	transcription DNA-dependent
Ole9; Len9.2	Glyma.09g209100	43355642	43360197	GO:0006487	protein N-linked glycosylation
Ole9; Len9.2	Glyma.09g209100	43355642	43360197	GO:0006623	protein targeting to vacuole
Ole9; Len9.2	Glyma.09g209100	43355642	43360197	GO:0006944	cellular membrane fusion
Ole9; Len9.2	Glyma.09g209100	43355642	43360197	GO:0009627	systemic acquired resistance
Ole9; Len9.2	Glyma.09g209100	43355642	43360197	GO:0009697	salicylic acid biosynthetic process
Ole9; Len9.2	Glyma.09g209100	43355642	43360197	GO:0009742	brassinosteroid mediated signaling pathway
Ole9; Len9.2	Glyma.09g209100	43355642	43360197	GO:0016192	vesicle-mediated transport
Ole9; Len9.2	Glyma.09g209100	43355642	43360197	GO:0032784	regulation of transcription elongation DNA-dependent
Ole9; Len9.2	Glyma.09g209100	43355642	43360197	GO:0048193	Golgi vesicle transport
Ole9; Len9.2	Glyma.09g209200	43361634	43365478	GO:0008150	biological_process
Ole9; Len9.2	Glyma.09g209200	43361634	43365478	GO:0009965	leaf morphogenesis
Ole9; Len9.2	Glyma.09g209200	43361634	43365478	GO:0010155	regulation of proton transport
Ole9; Len9.2	Glyma.09g209200	43361634	43365478	GO:0030154	cell differentiation
Ole9; Len9.2	Glyma.09g209200	43361634	43365478	GO:0046777	protein autophosphorylation
Ole9; Len9.2	Glyma.09g209300	43370740	43372816	GO:0006259	DNA metabolic process
Ole9; Len9.2	Glyma.09g209300	43370740	43372816	GO:0009944	polarity specification of adaxial/abaxial axis
Ole9; Len9.2	Glyma.09g209300	43370740	43372816	GO:0016226	iron-sulfur cluster assembly
Ole9; Len9.2	Glyma.09g209300	43370740	43372816	GO:0042127	regulation of cell proliferation
Ole9; Len9.2	Glyma.09g209300	43370740	43372816	GO:0051726	regulation of cell cycle
Ole9; Len9.2	Glyma.09g209300	43370740	43372816	GO:1990067	

Table 4.6 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Ole9; Len9.2	Glyma.09g209300	43370740	43372816	GO:2001022	
Ole9; Len9.2	Glyma.09g209400	43377152	43380635	GO:0006629	lipid metabolic process
Ole9; Len9.2	Glyma.09g209400	43377152	43380635	GO:0016042	lipid catabolic process
Ole9; Len9.2	Glyma.09g209700	43396577	43398265	GO:0008150	biological_process
Ole9; Len9.2	Glyma.09g209700	43396577	43398265	GO:0010075	regulation of meristem growth
Ole9; Len9.2	Glyma.09g209700	43396577	43398265	GO:0016556	mRNA modification
Len9.2	Glyma.09g209800	43411309	43411773	GO:0000165	MAPKKK cascade
Len9.2	Glyma.09g209800	43411309	43411773	GO:0000289	nuclear-transcribed mRNA poly(A) tail shortening
Len9.2	Glyma.09g209800	43411309	43411773	GO:0002679	respiratory burst involved in defense response
Len9.2	Glyma.09g209800	43411309	43411773	GO:0009451	RNA modification
Len9.2	Glyma.09g209800	43411309	43411773	GO:0009693	ethylene biosynthetic process
Len9.2	Glyma.09g209800	43411309	43411773	GO:0009738	abscisic acid mediated signaling pathway
Len9.2	Glyma.09g209800	43411309	43411773	GO:0009814	defense response incompatible interaction
Len9.2	Glyma.09g209800	43411309	43411773	GO:0009873	ethylene mediated signaling pathway
Len9.2	Glyma.09g209800	43411309	43411773	GO:0010200	response to chitin
Len9.2	Glyma.09g209800	43411309	43411773	GO:0010228	vegetative to reproductive phase transition of meristem
Len9.2	Glyma.09g209800	43411309	43411773	GO:0035556	intracellular signal transduction
Len9.2	Glyma.09g209800	43411309	43411773	GO:0042742	defense response to bacterium
Len9.2	Glyma.09g209900	43412726	43419532	GO:0006184	GTP catabolic process
Len9.2	Glyma.09g209900	43412726	43419532	GO:0007165	signal transduction
Len9.2	Glyma.09g209900	43412726	43419532	GO:0007186	G-protein coupled receptor protein signaling pathway
Len9.2	Glyma.09g209900	43412726	43419532	GO:0007188	G-protein signaling coupled to cAMP nucleotide second messenger
Len9.2	Glyma.09g209900	43412726	43419532	GO:0009630	gravitropism
Len9.2	Glyma.09g209900	43412726	43419532	GO:0009652	thigmotropism
Len9.2	Glyma.09g209900	43412726	43419532	GO:0009723	response to ethylene stimulus
Len9.2	Glyma.09g209900	43412726	43419532	GO:0009737	response to abscisic acid stimulus
Len9.2	Glyma.09g209900	43412726	43419532	GO:0009744	response to sucrose stimulus
Len9.2	Glyma.09g209900	43412726	43419532	GO:0009749	response to glucose stimulus
Len9.2	Glyma.09g209900	43412726	43419532	GO:0009750	response to fructose stimulus
Len9.2	Glyma.09g209900	43412726	43419532	GO:0010555	response to mannitol stimulus
Len9.2	Glyma.09g209900	43412726	43419532	GO:0048364	root development
Len9.2	Glyma.09g209900	43412726	43419532	GO:2000067	regulation of root morphogenesis
Len9.2	Glyma.09g210300	43440140	43447326	GO:0006974	response to DNA damage stimulus
Len9.2	Glyma.09g210300	43440140	43447326	GO:0007050	cell cycle arrest
Len9.2	Glyma.09g210700	43466944	43471858	GO:0001510	RNA methylation
Len9.2	Glyma.09g210800	43476243	43482337	GO:0006944	cellular membrane fusion
Len9.2	Glyma.09g210800	43476243	43482337	GO:0015031	protein transport
Len9.2	Glyma.09g210800	43476243	43482337	GO:0016192	vesicle-mediated transport
Len9.2	Glyma.09g210900	43486670	43490921	GO:0000165	MAPKKK cascade
Len9.2	Glyma.09g210900	43486670	43490921	GO:0005975	carbohydrate metabolic process
Len9.2	Glyma.09g210900	43486670	43490921	GO:0006098	pentose-phosphate shunt
Len9.2	Glyma.09g210900	43486670	43490921	GO:0006364	rRNA processing
Len9.2	Glyma.09g210900	43486670	43490921	GO:0006612	protein targeting to membrane
Len9.2	Glyma.09g210900	43486670	43490921	GO:0008152	metabolic process
Len9.2	Glyma.09g210900	43486670	43490921	GO:0009058	biosynthetic process
Len9.2	Glyma.09g210900	43486670	43490921	GO:0009409	response to cold
Len9.2	Glyma.09g210900	43486670	43490921	GO:0009595	detection of biotic stimulus
Len9.2	Glyma.09g210900	43486670	43490921	GO:0009637	response to blue light
Len9.2	Glyma.09g210900	43486670	43490921	GO:0009657	plastid organization
Len9.2	Glyma.09g210900	43486670	43490921	GO:0009697	salicylic acid biosynthetic process
Len9.2	Glyma.09g210900	43486670	43490921	GO:0009773	photosynthetic electron transport in photosystem I
Len9.2	Glyma.09g210900	43486670	43490921	GO:0009814	defense response incompatible interaction
Len9.2	Glyma.09g210900	43486670	43490921	GO:0009862	systemic acquired resistance salicylic acid mediated signaling pathway
Len9.2	Glyma.09g210900	43486670	43490921	GO:0009867	jasmonic acid mediated signaling pathway
Len9.2	Glyma.09g210900	43486670	43490921	GO:0009902	chloroplast relocation
Len9.2	Glyma.09g210900	43486670	43490921	GO:0010103	stomatal complex morphogenesis
Len9.2	Glyma.09g210900	43486670	43490921	GO:0010114	response to red light
Len9.2	Glyma.09g210900	43486670	43490921	GO:0010200	response to chitin
Len9.2	Glyma.09g210900	43486670	43490921	GO:0010207	photosystem II assembly
Len9.2	Glyma.09g210900	43486670	43490921	GO:0010218	response to far red light
Len9.2	Glyma.09g210900	43486670	43490921	GO:0010310	regulation of hydrogen peroxide metabolic process
Len9.2	Glyma.09g210900	43486670	43490921	GO:0010363	regulation of plant-type hypersensitive response
Len9.2	Glyma.09g210900	43486670	43490921	GO:0015979	photosynthesis
Len9.2	Glyma.09g210900	43486670	43490921	GO:0018119	peptidyl-cysteine S-nitrosylation
Len9.2	Glyma.09g210900	43486670	43490921	GO:0019344	cysteine biosynthetic process
Len9.2	Glyma.09g210900	43486670	43490921	GO:0019684	photosynthesis light reaction
Len9.2	Glyma.09g210900	43486670	43490921	GO:0031348	negative regulation of defense response
Len9.2	Glyma.09g210900	43486670	43490921	GO:0035304	regulation of protein dephosphorylation
Len9.2	Glyma.09g210900	43486670	43490921	GO:0042742	defense response to bacterium
Len9.2	Glyma.09g210900	43486670	43490921	GO:0043900	regulation of multi-organism process
Len9.2	Glyma.09g210900	43486670	43490921	GO:0050832	defense response to fungus
Len9.2	Glyma.09g211000	43495032	43501760	GO:0008150	biological_process
Len9.2	Glyma.09g211100	43504097	43508529	GO:0006914	autophagy
Len9.2	Glyma.09g211100	43504097	43508529	GO:0008152	metabolic process
Len9.2	Glyma.09g211100	43504097	43508529	GO:0009610	response to symbiotic fungus
Len9.2	Glyma.09g211200	43510772	43516423	GO:0009658	chloroplast organization
Len9.2	Glyma.09g211200	43510772	43516423	GO:0010027	thylakoid membrane organization
Len9.2	Glyma.09g211200	43510772	43516423	GO:0030003	cellular cation homeostasis
Len9.2	Glyma.09g211200	43510772	43516423	GO:0070838	divalent metal ion transport
Len9.2	Glyma.09g211400	43534918	43538177	GO:0006355	regulation of transcription DNA-dependent

Table 4.7 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 11 based on Glyma 2.0 position.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Oil11	Glyma.11g223100	31828844	31831620	GO:0008150	biological_process
Oil11	Glyma.11g225500	32065768	32068093	GO:0008152	metabolic process
Oil11	Glyma.11g227100	32199169	32200630	GO:0000023	maltose metabolic process
Oil11	Glyma.11g227100	32199169	32200630	GO:0000165	MAPKKK cascade
Oil11	Glyma.11g227100	32199169	32200630	GO:0005975	carbohydrate metabolic process
Oil11	Glyma.11g227100	32199169	32200630	GO:0005986	sucrose biosynthetic process
Oil11	Glyma.11g227100	32199169	32200630	GO:0006098	pentose-phosphate shunt
Oil11	Glyma.11g227100	32199169	32200630	GO:0006364	rRNA processing
Oil11	Glyma.11g227100	32199169	32200630	GO:0006612	protein targeting to membrane
Oil11	Glyma.11g227100	32199169	32200630	GO:0009409	response to cold
Oil11	Glyma.11g227100	32199169	32200630	GO:0009595	detection of biotic stimulus
Oil11	Glyma.11g227100	32199169	32200630	GO:0009657	plastid organization
Oil11	Glyma.11g227100	32199169	32200630	GO:0009697	salicylic acid biosynthetic process
Oil11	Glyma.11g227100	32199169	32200630	GO:0009773	photosynthetic electron transport in photosystem I
Oil11	Glyma.11g227100	32199169	32200630	GO:0009814	defense response incompatible interaction
Oil11	Glyma.11g227100	32199169	32200630	GO:0009862	systemic acquired resistance salicylic acid mediated signaling pathway
Oil11	Glyma.11g227100	32199169	32200630	GO:0009867	jasmonic acid mediated signaling pathway
Oil11	Glyma.11g227100	32199169	32200630	GO:0010103	stomatal complex morphogenesis
Oil11	Glyma.11g227100	32199169	32200630	GO:0010200	response to chitin
Oil11	Glyma.11g227100	32199169	32200630	GO:0010207	photosystem II assembly
Oil11	Glyma.11g227100	32199169	32200630	GO:0010310	regulation of hydrogen peroxide metabolic process
Oil11	Glyma.11g227100	32199169	32200630	GO:0010363	regulation of plant-type hypersensitive response
Oil11	Glyma.11g227100	32199169	32200630	GO:0016051	carbohydrate biosynthetic process
Oil11	Glyma.11g227100	32199169	32200630	GO:0019252	starch biosynthetic process
Oil11	Glyma.11g227100	32199169	32200630	GO:0019253	reductive pentose-phosphate cycle
Oil11	Glyma.11g227100	32199169	32200630	GO:0019288	isopentenyl diphosphate biosynthetic process mevalonate-independent pathway
Oil11	Glyma.11g227100	32199169	32200630	GO:0019344	cysteine biosynthetic process
Oil11	Glyma.11g227100	32199169	32200630	GO:0019684	photosynthesis light reaction
Oil11	Glyma.11g227100	32199169	32200630	GO:0019760	glucosinolate metabolic process
Oil11	Glyma.11g227100	32199169	32200630	GO:0019761	glucosinolate biosynthetic process
Oil11	Glyma.11g227100	32199169	32200630	GO:0031348	negative regulation of defense response
Oil11	Glyma.11g227100	32199169	32200630	GO:0035304	regulation of protein dephosphorylation
Oil11	Glyma.11g227100	32199169	32200630	GO:0042742	defense response to bacterium
Oil11	Glyma.11g227100	32199169	32200630	GO:0043085	positive regulation of catalytic activity
Oil11	Glyma.11g227100	32199169	32200630	GO:0043900	regulation of multi-organism process
Oil11	Glyma.11g227100	32199169	32200630	GO:0050832	defense response to fungus
Oil11	Glyma.11g230400	32560337	32563722	GO:0006468	protein phosphorylation
Oil11	Glyma.11g230400	32560337	32563722	GO:0006612	protein targeting to membrane
Oil11	Glyma.11g230400	32560337	32563722	GO:0006944	cellular membrane fusion
Oil11	Glyma.11g230400	32560337	32563722	GO:0009723	response to ethylene stimulus
Oil11	Glyma.11g230400	32560337	32563722	GO:0009738	abscisic acid mediated signaling pathway
Oil11	Glyma.11g230400	32560337	32563722	GO:0010363	regulation of plant-type hypersensitive response
Oil11	Glyma.11g230400	32560337	32563722	GO:0035556	intracellular signal transduction
Oil11	Glyma.11g230400	32560337	32563722	GO:0043069	negative regulation of programmed cell death
Oil11	Glyma.11g233100	32840573	32844286	GO:0008150	biological_process
Oil11	Glyma.11g237000	33169449	33171449	GO:0008152	metabolic process
Oil11	Glyma.11g239600	33399330	33401730	GO:0008150	biological_process
Oil11	Glyma.11g239800	33424880	33439567	GO:0006635	fatty acid beta-oxidation
Oil11	Glyma.11g239800	33424880	33439567	GO:0006810	transport
Oil11	Glyma.11g239800	33424880	33439567	GO:0007031	peroxisome organization
Oil11	Glyma.11g239800	33424880	33439567	GO:0009407	toxin catabolic process
Oil11	Glyma.11g239800	33424880	33439567	GO:0010030	positive regulation of seed germination
Oil11	Glyma.11g239800	33424880	33439567	GO:0015916	fatty-acyl-CoA transport
Oil11	Glyma.11g239800	33424880	33439567	GO:0043161	proteasomal ubiquitin-dependent protein catabolic process
Oil11	Glyma.11g239800	33424880	33439567	GO:0051788	response to misfolded protein
Oil11	Glyma.11g239800	33424880	33439567	GO:0055085	transmembrane transport
Oil11	Glyma.11g239800	33424880	33439567	GO:0080129	proteasome core complex assembly
Oil11	Glyma.11g240600	33491895	33493311	GO:0006355	regulation of transcription DNA-dependent
Oil11	Glyma.11g240600	33491895	33493311	GO:0010044	response to aluminum ion
Oil11	Glyma.11g240600	33491895	33493311	GO:0010447	response to acidity
Oil11	Glyma.11g241100	33533212	33535772	GO:0008150	biological_process
Oil11	Glyma.11g241700	33581449	33584216	GO:0006334	nucleosome assembly
Oil11	Glyma.11g243600	33743871	33749429	GO:0005975	carbohydrate metabolic process
Oil11	Glyma.11g243600	33743871	33749429	GO:0005990	lactose catabolic process
Oil11	Glyma.11g243600	33743871	33749429	GO:0006499	N-terminal protein myristoylation
Oil11	Glyma.11g243600	33743871	33749429	GO:0009311	oligosaccharide metabolic process
Oil11	Glyma.11g243600	33743871	33749429	GO:0016139	glycoside catabolic process
Oil11	Glyma.11g243600	33743871	33749429	GO:0046477	glycosylceramide catabolic process
Oil11	Glyma.11g243700	33750461	33756998	GO:0006084	acetyl-CoA metabolic process
Oil11	Glyma.11g243700	33750461	33756998	GO:0016126	sterol biosynthetic process
Oil11	Glyma.11g243700	33750461	33756998	GO:0016132	brassinosteroid biosynthetic process
Oil11	Glyma.11g243800	33759603	33766398	GO:0006333	chromatin assembly or disassembly
Oil11	Glyma.11g243800	33759603	33766398	GO:0006346	methylation-dependent chromatin silencing
Oil11	Glyma.11g243800	33759603	33766398	GO:0008150	biological_process
Oil11	Glyma.11g243800	33759603	33766398	GO:0009220	pyrimidine ribonucleotide biosynthetic process
Oil11	Glyma.11g243800	33759603	33766398	GO:0009909	regulation of flower development
Oil11	Glyma.11g243800	33759603	33766398	GO:0016570	histone modification
Oil11	Glyma.11g243800	33759603	33766398	GO:0031048	chromatin silencing by small RNA
Oil11	Glyma.11g243800	33759603	33766398	GO:0048449	floral organ formation
Oil11	Glyma.11g243800	33759603	33766398	GO:0051567	histone H3-K9 methylation

Table 4.8 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 13 based on Glyma 2.0 position.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Ole13; Len13	Glyma.13g042200	13542395	13551520	GO:0009685	gibberellin metabolic process
Ole13; Len13	Glyma.13g042200	13542395	13551520	GO:0055114	oxidation-reduction process
Ole13; Len13	Glyma.13g042300	13554013	13555877	GO:0010413	glucuronoxylan metabolic process
Ole13; Len13	Glyma.13g042300	13554013	13555877	GO:0045492	xylan biosynthetic process
Ole13; Len13	Glyma.13g042700	13579883	13588272	GO:0008150	biological_process
Ole13; Len13	Glyma.13g042700	13579883	13588272	GO:0008284	positive regulation of cell proliferation
Ole13; Len13	Glyma.13g043000	13631954	13636186	GO:0009611	response to wounding
Ole13; Len13	Glyma.13g043000	13631954	13636186	GO:0009651	response to salt stress
Ole13; Len13	Glyma.13g043000	13631954	13636186	GO:0009737	response to abscisic acid stimulus
Ole13; Len13	Glyma.13g043000	13631954	13636186	GO:0015908	fatty acid transport
Ole13; Len13	Glyma.13g043000	13631954	13636186	GO:0080051	cutin transport
Ole13; Len13	Glyma.13g043300	13657741	13660361	GO:0006810	transport
Ole13; Len13	Glyma.13g043300	13657741	13660361	GO:0008272	sulfate transport
Ole13; Len13	Glyma.13g043300	13657741	13660361	GO:0055085	transmembrane transport
Ole13; Len13	Glyma.13g043400	13660578	13661855	GO:0006096	glycolysis
Ole13; Len13	Glyma.13g043400	13660578	13661855	GO:0009060	aerobic respiration
Ole13; Len13	Glyma.13g043400	13660578	13661855	GO:0009809	lignin biosynthetic process
Ole13; Len13	Glyma.13g043400	13660578	13661855	GO:0033587	shikimate biosynthetic process
Ole13; Len13	Glyma.13g043400	13660578	13661855	GO:0046686	response to cadmium ion
Ole13; Len13	Glyma.13g043500	13674949	13686040	GO:0009611	response to wounding
Ole13; Len13	Glyma.13g043500	13674949	13686040	GO:0009651	response to salt stress
Ole13; Len13	Glyma.13g043500	13674949	13686040	GO:0009737	response to abscisic acid stimulus
Ole13; Len13	Glyma.13g043500	13674949	13686040	GO:0015908	fatty acid transport
Ole13; Len13	Glyma.13g043500	13674949	13686040	GO:0080051	cutin transport
Ole13; Len13	Glyma.13g043600	13695151	13698734	GO:0009611	response to wounding
Ole13; Len13	Glyma.13g043600	13695151	13698734	GO:0009651	response to salt stress
Ole13; Len13	Glyma.13g043600	13695151	13698734	GO:0009737	response to abscisic acid stimulus
Ole13; Len13	Glyma.13g043600	13695151	13698734	GO:0015908	fatty acid transport
Ole13; Len13	Glyma.13g043600	13695151	13698734	GO:0080051	cutin transport
Ole13; Len13	Glyma.13g045300	13911395	13912205	GO:0008150	biological_process
Ole13; Len13	Glyma.13g045300	13911395	13912205	GO:0009741	response to brassinosteroid stimulus
Ole13; Len13	Glyma.13g048500	14448130	14454096	GO:0055114	oxidation-reduction process
Ole13; Len13	Glyma.13g048600	14456516	14458768	GO:0006499	N-terminal protein myristoylation
Ole13; Len13	Glyma.13g048600	14456516	14458768	GO:0006888	ER to Golgi vesicle-mediated transport
Ole13; Len13	Glyma.13g048600	14456516	14458768	GO:0016998	cell wall macromolecule catabolic process
Ole13; Len13	Glyma.13g048600	14456516	14458768	GO:0030968	endoplasmic reticulum unfolded protein response
Ole13; Len13	Glyma.13g048600	14456516	14458768	GO:0043090	amino acid import
Ole13; Len13	Glyma.13g049200	14515169	14519326	GO:0008150	biological_process
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0006312	mitotic recombination
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0007276	gamete generation
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0007346	regulation of mitotic cell cycle
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0009560	embryo sac egg cell differentiation
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0009733	response to auxin stimulus
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0009965	leaf morphogenesis
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0010048	vernalization response
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0010071	root meristem specification
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0030154	cell differentiation
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0032875	regulation of DNA endoreduplication
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0042023	DNA endoreduplication
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0043161	proteasomal ubiquitin-dependent protein catabolic process
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0043248	proteasome assembly
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0048364	root development
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0048829	root cap development
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0051301	cell division
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0051302	regulation of cell division
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0051510	regulation of unidimensional cell growth
Ole13; Len13	Glyma.13g049400	14530308	14541307	GO:0051788	response to misfolded protein
Ole13; Len13	Glyma.13g050300	14630394	14632571	GO:0009658	chloroplast organization
Ole13; Len13	Glyma.13g050300	14630394	14632571	GO:0009909	regulation of flower development
Ole13; Len13	Glyma.13g050400	14652395	14656183	GO:0006355	regulation of transcription DNA-dependent
Ole13; Len13	Glyma.13g050400	14652395	14656183	GO:0045893	positive regulation of transcription DNA-dependent
Ole13; Len13	Glyma.13g050400	14652395	14656183	GO:2000652	regulation of secondary cell wall biogenesis
Ole13; Len13	Glyma.13g050500	14674243	14683099	GO:0006952	defense response
Ole13; Len13	Glyma.13g050500	14674243	14683099	GO:0007165	signal transduction
Ole13; Len13	Glyma.13g050500	14674243	14683099	GO:0050832	defense response to fungus
Ole13; Len13	Glyma.13g050900	14730756	14736198	GO:0006913	nucleocytoplasmic transport
Ole13; Len13	Glyma.13g050900	14730756	14736198	GO:0008150	biological_process
Ole13; Len13	Glyma.13g050900	14730756	14736198	GO:0048527	lateral root development
Ole13; Len13	Glyma.13g052300	14939773	14940715	GO:0008150	biological_process
Ole13; Len13	Glyma.13g052300	14939773	14940715	GO:0009954	proximal/distal pattern formation
Ole13; Len13	Glyma.13g052300	14939773	14940715	GO:0010227	floral organ abscission
Ole13; Len13	Glyma.13g052300	14939773	14940715	GO:0048439	flower morphogenesis
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0000165	MAPKKK cascade

Table 4.8 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0001666	response to hypoxia
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0006612	protein targeting to membrane
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0009595	detection of biotic stimulus
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0009617	response to bacterium
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0009620	response to fungus
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0009627	systemic acquired resistance
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0009697	salicylic acid biosynthetic process
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0009813	flavonoid biosynthetic process
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0009862	systemic acquired resistance salicylic acid mediated signaling pathway
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0009867	jasmonic acid mediated signaling pathway
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0010200	response to chitin
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0010310	regulation of hydrogen peroxide metabolic process
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0010363	regulation of plant-type hypersensitive response
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0031347	regulation of defense response
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0031348	negative regulation of defense response
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0034976	response to endoplasmic reticulum stress
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0042742	defense response to bacterium
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0043900	regulation of multi-organism process
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0045087	innate immune response
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0050832	defense response to fungus
Ole13; Len13	Glyma.13g052800	15005141	15009623	GO:0055114	oxidation-reduction process
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0000271	polysaccharide biosynthetic process
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0000272	polysaccharide catabolic process
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0005982	starch metabolic process
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0006096	glycolysis
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0006816	calcium ion transport
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0006833	water transport
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0006972	hyperosmotic response
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0007030	Golgi organization
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0007389	pattern specification process
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0008361	regulation of cell size
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0009266	response to temperature stimulus
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0009651	response to salt stress
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0009664	plant-type cell wall organization
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0009750	response to fructose stimulus
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0009825	multidimensional cell growth
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0009832	plant-type cell wall biogenesis
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0009926	auxin polar transport
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0009932	cell tip growth
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0010015	root morphogenesis
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0010215	cellulose microfibril organization
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0010817	regulation of hormone levels
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0016049	cell growth
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0030243	cellulose metabolic process
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0040007	growth
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0043481	anthocyanin accumulation in tissues in response to UV light
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0046686	response to cadmium ion
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0048767	root hair elongation
Ole13; Len13	Glyma.13g053300	15064460	15069264	GO:0071555	cell wall organization
Ole13; Len13	Glyma.13g053400	15070294	15076224	GO:0006810	transport
Ole13; Len13	Glyma.13g053400	15070294	15076224	GO:0006839	mitochondrial transport
Ole13; Len13	Glyma.13g053400	15070294	15076224	GO:0055085	transmembrane transport
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0000902	cell morphogenesis
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0006007	glucose catabolic process
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0006096	glycolysis
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0006468	protein phosphorylation
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0006833	water transport
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0006972	hyperosmotic response
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0007030	Golgi organization
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0009266	response to temperature stimulus
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0009651	response to salt stress
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0009791	post-embryonic development
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0010193	response to ozone
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0010483	pollen tube reception
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0016049	cell growth
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0030244	cellulose biosynthetic process
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0046686	response to cadmium ion
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0046777	protein autophosphorylation
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0048193	Golgi vesicle transport
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0048767	root hair elongation
Ole13; Len13	Glyma.13g053700	15096852	15099898	GO:0050832	defense response to fungus
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0000902	cell morphogenesis
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0006007	glucose catabolic process
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0006096	glycolysis
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0006468	protein phosphorylation

Table 4.8 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0006833	water transport
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0006972	hyperosmotic response
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0007030	Golgi organization
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0009266	response to temperature stimulus
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0009651	response to salt stress
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0009791	post-embryonic development
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0010193	response to ozone
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0010483	pollen tube reception
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0016049	cell growth
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0030244	cellulose biosynthetic process
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0046686	response to cadmium ion
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0046777	protein autophosphorylation
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0048193	Golgi vesicle transport
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0048767	root hair elongation
Ole13; Len13	Glyma.13g053800	15110071	15112808	GO:0050832	defense response to fungus
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0000902	cell morphogenesis
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0006007	glucose catabolic process
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0006096	glycolysis
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0006468	protein phosphorylation
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0006833	water transport
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0006972	hyperosmotic response
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0007030	Golgi organization
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0009266	response to temperature stimulus
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0009651	response to salt stress
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0009791	post-embryonic development
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0010193	response to ozone
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0010483	pollen tube reception
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0016049	cell growth
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0030244	cellulose biosynthetic process
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0046686	response to cadmium ion
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0046777	protein autophosphorylation
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0048193	Golgi vesicle transport
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0048767	root hair elongation
Ole13; Len13	Glyma.13g054300	15153615	15156149	GO:0050832	defense response to fungus
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0000902	cell morphogenesis
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0006007	glucose catabolic process
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0006096	glycolysis
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0006468	protein phosphorylation
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0006833	water transport
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0006972	hyperosmotic response
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0007030	Golgi organization
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0009266	response to temperature stimulus
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0009651	response to salt stress
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0009791	post-embryonic development
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0010193	response to ozone
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0010483	pollen tube reception
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0016049	cell growth
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0030244	cellulose biosynthetic process
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0046686	response to cadmium ion
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0046777	protein autophosphorylation
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0048193	Golgi vesicle transport
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0048767	root hair elongation
Ole13; Len13	Glyma.13g054400	15164832	15168321	GO:0050832	defense response to fungus
Ole13; Len13	Glyma.13g054500	15173682	15182007	GO:0009793	embryo development ending in seed dormancy
Ole13; Len13	Glyma.13g054500	15173682	15182007	GO:0010413	glucuronoxylan metabolic process
Ole13; Len13	Glyma.13g054500	15173682	15182007	GO:0045492	xylan biosynthetic process
Ole13; Len13	Glyma.13g054600	15191769	15194310	GO:0009556	microsporogenesis
Ole13; Len13	Glyma.13g054600	15191769	15194310	GO:0009827	plant-type cell wall modification
Ole13; Len13	Glyma.13g054600	15191769	15194310	GO:0009860	pollen tube growth
Ole13; Len13	Glyma.13g054600	15191769	15194310	GO:0010584	pollen exine formation
Ole13; Len13	Glyma.13g055300	15237358	15240608	GO:0008150	biological_process
Pal13	Glyma.13g078600	18467981	18471050	GO:0006200	ATP catabolic process
Pal13	Glyma.13g078600	18467981	18471050	GO:0006754	ATP biosynthetic process
Pal13	Glyma.13g078600	18467981	18471050	GO:0015986	ATP synthesis coupled proton transport
Pal13	Glyma.13g078600	18467981	18471050	GO:0015991	ATP hydrolysis coupled proton transport
Pal13	Glyma.13g078600	18467981	18471050	GO:0015992	proton transport
Pal13	Glyma.13g078600	18467981	18471050	GO:0046034	ATP metabolic process
Pal13	Glyma.13g078600	18467981	18471050	GO:0046686	response to cadmium ion
Pal13	Glyma.13g078900	18542943	18553529	GO:0006084	acetyl-CoA metabolic process
Pal13	Glyma.13g078900	18542943	18553529	GO:0016125	sterol metabolic process
Pal13	Glyma.13g078900	18542943	18553529	GO:0016126	sterol biosynthetic process
Pal13	Glyma.13g078900	18542943	18553529	GO:0016132	brassinosteroid biosynthetic process
Pal13	Glyma.13g078900	18542943	18553529	GO:0019932	second-messenger-mediated signaling
Pal13	Glyma.13g078900	18542943	18553529	GO:0060964	regulation of gene silencing by miRNA
Pal13	Glyma.13g079100	18567152	18568147	GO:0006468	protein phosphorylation

Table 4.8 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Pal13	Glyma.13g079100	18567152	18568147	GO:0009620	response to fungus
Pal13	Glyma.13g079100	18567152	18568147	GO:0009691	cytokinin biosynthetic process
Pal13	Glyma.13g079700	18613446	18616026	GO:0006629	lipid metabolic process
Pal13	Glyma.13g088000	20100722	20108180	GO:0005975	carbohydrate metabolic process
Pal13	Glyma.13g088000	20100722	20108180	GO:0010075	regulation of meristem growth
Pal13	Glyma.13g088400	20228085	20231887	GO:0008150	biological_process
Pal13	Glyma.13g088500	20280557	20280960	GO:0006091	generation of precursor metabolites and energy
Pal13	Glyma.13g088500	20280557	20280960	GO:0006354	transcription elongation DNA-dependent
Pal13	Glyma.13g088500	20280557	20280960	GO:0009773	photosynthetic electron transport in photosystem I
Pal13	Glyma.13g088500	20280557	20280960	GO:0015979	photosynthesis
Pal13	Glyma.13g088500	20280557	20280960	GO:0019684	photosynthesis light reaction
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0000278	mitotic cell cycle
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0000724	double-strand break repair via homologous recombination
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0006261	DNA-dependent DNA replication
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0006275	regulation of DNA replication
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0006306	DNA methylation
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0006342	chromatin silencing
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0006351	transcription DNA-dependent
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0006354	transcription elongation DNA-dependent
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0009555	pollen development
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0016444	somatic cell DNA recombination
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0016568	chromatin modification
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0030422	production of siRNA involved in RNA interference
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0031047	gene silencing by RNA
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0035194	posttranscriptional gene silencing by RNA
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0048451	petal formation
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0048453	sepal formation
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0051567	histone H3-K9 methylation
Lin13.1	Glyma.13g196900	31055891	31072513	GO:0051726	regulation of cell cycle
Lin13.1	Glyma.13g197800	31181424	31185590	GO:0006302	double-strand break repair
Lin13.1	Glyma.13g197800	31181424	31185590	GO:0009751	response to salicylic acid stimulus
Lin13.1	Glyma.13g197800	31181424	31185590	GO:0010332	response to gamma radiation
Lin13.1	Glyma.13g199400	31318973	31320276	GO:0006508	proteolysis
Lin13.1	Glyma.13g199900	31359993	31364555	GO:0009408	response to heat
Lin13.1	Glyma.13g199900	31359993	31364555	GO:0009409	response to cold
Lin13.1	Glyma.13g199900	31359993	31364555	GO:0009414	response to water deprivation
Lin13.1	Glyma.13g199900	31359993	31364555	GO:0009651	response to salt stress
Lin13.1	Glyma.13g200100	31375134	31379683	GO:0009408	response to heat
Lin13.1	Glyma.13g200100	31375134	31379683	GO:0009409	response to cold
Lin13.1	Glyma.13g200100	31375134	31379683	GO:0009414	response to water deprivation
Lin13.1	Glyma.13g200100	31375134	31379683	GO:0009651	response to salt stress
Lin13.1	Glyma.13g203000	31683629	31689851	GO:0006826	iron ion transport
Lin13.1	Glyma.13g203000	31683629	31689851	GO:0006855	drug transmembrane transport
Lin13.1	Glyma.13g203000	31683629	31689851	GO:0006879	cellular iron ion homeostasis
Lin13.1	Glyma.13g203000	31683629	31689851	GO:0010106	cellular response to iron ion starvation
Lin13.1	Glyma.13g203000	31683629	31689851	GO:0010167	response to nitrate
Lin13.1	Glyma.13g203000	31683629	31689851	GO:0010413	glucuronoxylan metabolic process
Lin13.1	Glyma.13g203000	31683629	31689851	GO:0015706	nitrate transport
Lin13.1	Glyma.13g203000	31683629	31689851	GO:0045492	xylan biosynthetic process
Lin13.1	Glyma.13g203000	31683629	31689851	GO:0055085	transmembrane transport
Lin13.1	Glyma.13g203000	31683629	31689851	GO:0071281	cellular response to iron ion
Lin13.1	Glyma.13g203000	31683629	31689851	GO:0071369	cellular response to ethylene stimulus
Lin13.1	Glyma.13g203000	31683629	31689851	GO:0071732	cellular response to nitric oxide
Lin13.1	Glyma.13g204700	31872482	31883107	GO:0009058	biosynthetic process
Lin13.1	Glyma.13g204700	31872482	31883107	GO:0009250	glucan biosynthetic process
Lin13.1	Glyma.13g204700	31872482	31883107	GO:0019252	starch biosynthetic process
Lin13.1	Glyma.13g206000	32002491	32003259	GO:0006333	chromatin assembly or disassembly
Lin13.1	Glyma.13g206200	32005629	32006396	GO:0010628	positive regulation of gene expression
Lin13.1	Glyma.13g206900	32065701	32068163	GO:0006096	glycolysis
Lin13.1	Glyma.13g206900	32065701	32068163	GO:0006979	response to oxidative stress
Lin13.1	Glyma.13g206900	32065701	32068163	GO:0009060	aerobic respiration
Lin13.1	Glyma.13g206900	32065701	32068163	GO:0046686	response to cadmium ion
Lin13.1	Glyma.13g207200	32082001	32087695	GO:0008152	metabolic process
Lin13.1	Glyma.13g207200	32082001	32087695	GO:0030148	sphingolipid biosynthetic process
Lin13.1	Glyma.13g208400	32231119	32234068	GO:0030001	metal ion transport
Lin13.1	Glyma.13g209200	32324301	32328052	GO:0008150	biological_process
Lin13.1	Glyma.13g210000	32388338	32390888	GO:0006096	glycolysis
Lin13.1	Glyma.13g210000	32388338	32390888	GO:0006833	water transport
Lin13.1	Glyma.13g210000	32388338	32390888	GO:0006972	hyperosmotic response
Lin13.1	Glyma.13g210000	32388338	32390888	GO:0007030	Golgi organization
Lin13.1	Glyma.13g210000	32388338	32390888	GO:0009266	response to temperature stimulus
Lin13.1	Glyma.13g210000	32388338	32390888	GO:0009651	response to salt stress
Lin13.1	Glyma.13g210000	32388338	32390888	GO:0046686	response to cadmium ion
Lin13.1	Glyma.13g210400	32420426	32423292	GO:0006163	purine nucleotide metabolic process
Lin13.1	Glyma.13g211900	32559102	32575545	GO:0008150	biological_process

Table 4.8 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0000096	sulfur amino acid metabolic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0006546	glycine catabolic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0006636	unsaturated fatty acid biosynthetic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0006655	phosphatidylglycerol biosynthetic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0006733	oxidoreduction coenzyme metabolic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0006766	vitamin metabolic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0008152	metabolic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0008652	cellular amino acid biosynthetic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0009072	aromatic amino acid family metabolic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0009073	aromatic amino acid family biosynthetic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0009106	lipoate metabolic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0009108	coenzyme biosynthetic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0009117	nucleotide metabolic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0009416	response to light stimulus
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0009695	jasmonic acid biosynthetic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0009965	leaf morphogenesis
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0010155	regulation of proton transport
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0015994	chlorophyll metabolic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0015995	chlorophyll biosynthetic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0016114	terpenoid biosynthetic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0016117	carotenoid biosynthetic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0019216	regulation of lipid metabolic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0019288	isopentenyl diphosphate biosynthetic process mevalonate-independent pathway
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0019344	cysteine biosynthetic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0019748	secondary metabolic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0030154	cell differentiation
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0031408	oxylipin biosynthetic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0032880	regulation of protein localization
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0044272	sulfur compound biosynthetic process
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0045893	positive regulation of transcription DNA-dependent
Lin13.1	Glyma.13g213500	32711673	32716430	GO:0046777	protein autophosphorylation
Lin13.1	Glyma.13g214000	32748724	32749782	GO:0008150	biological_process
Lin13.1	Glyma.13g215400	32875517	32877386	GO:0006098	pentose-phosphate shunt
Lin13.1	Glyma.13g215400	32875517	32877386	GO:0006364	rRNA processing
Lin13.1	Glyma.13g215400	32875517	32877386	GO:0006636	unsaturated fatty acid biosynthetic process
Lin13.1	Glyma.13g215400	32875517	32877386	GO:0009073	aromatic amino acid family biosynthetic process
Lin13.1	Glyma.13g215400	32875517	32877386	GO:0009965	leaf morphogenesis
Lin13.1	Glyma.13g215400	32875517	32877386	GO:0015995	chlorophyll biosynthetic process
Lin13.1	Glyma.13g215400	32875517	32877386	GO:0016117	carotenoid biosynthetic process
Lin13.1	Glyma.13g215400	32875517	32877386	GO:0019288	isopentenyl diphosphate biosynthetic process mevalonate-independent pathway
Lin13.1	Glyma.13g215400	32875517	32877386	GO:0019344	cysteine biosynthetic process
Lin13.1	Glyma.13g215400	32875517	32877386	GO:0030154	cell differentiation
Lin13.1	Glyma.13g215400	32875517	32877386	GO:0045893	positive regulation of transcription DNA-dependent
Lin13.1	Glyma.13g216600	32978123	32979687	GO:0006457	protein folding
Lin13.1	Glyma.13g216600	32978123	32979687	GO:0009266	response to temperature stimulus
Lin13.1	Glyma.13g216600	32978123	32979687	GO:0009408	response to heat
Lin13.1	Glyma.13g216600	32978123	32979687	GO:0009615	response to virus
Lin13.1	Glyma.13g216600	32978123	32979687	GO:0009617	response to bacterium
Lin13.1	Glyma.13g216600	32978123	32979687	GO:0009644	response to high light intensity
Lin13.1	Glyma.13g216600	32978123	32979687	GO:0016567	protein ubiquitination
Lin13.1	Glyma.13g216600	32978123	32979687	GO:0034976	response to endoplasmic reticulum stress
Lin13.1	Glyma.13g216600	32978123	32979687	GO:0042542	response to hydrogen peroxide
Lin13.1	Glyma.13g216600	32978123	32979687	GO:0046686	response to cadmium ion
Lin13.1	Glyma.13g216700	32980066	32981039	GO:0008150	biological_process
Prot13.1	Glyma.13g239300	34970872	34976701	GO:0000003	reproduction
Prot13.1	Glyma.13g239300	34970872	34976701	GO:0006075	13-beta-D-glucan biosynthetic process
Prot13.1	Glyma.13g239300	34970872	34976701	GO:0006944	cellular membrane fusion
Prot13.1	Glyma.13g239300	34970872	34976701	GO:0007623	circadian rhythm
Prot13.1	Glyma.13g239300	34970872	34976701	GO:0009555	pollen development
Prot13.1	Glyma.13g239300	34970872	34976701	GO:0009556	microsporogenesis
Prot13.1	Glyma.13g239300	34970872	34976701	GO:0052543	callose deposition in cell wall
Prot13.1	Glyma.13g245200	35444517	35444785	GO:0006412	translation
Lin13.2	Glyma.13g263600	36687392	36692071	GO:0006468	protein phosphorylation
Lin13.2	Glyma.13g263600	36687392	36692071	GO:0007165	signal transduction
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0000165	MAPKKK cascade
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0001666	response to hypoxia
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0002679	respiratory burst involved in defense response
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0006355	regulation of transcription DNA-dependent
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0006612	protein targeting to membrane
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0006944	cellular membrane fusion
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009409	response to cold
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009595	detection of biotic stimulus
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009617	response to bacterium
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009625	response to insect
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009627	systemic acquired resistance

Table 4.8 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009697	salicylic acid biosynthetic process
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009751	response to salicylic acid stimulus
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009753	response to jasmonic acid stimulus
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009759	indole glucosinolate biosynthetic process
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009814	defense response incompatible interaction
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009862	systemic acquired resistance salicylic acid mediated signaling pathway
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009863	salicylic acid mediated signaling pathway
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009864	induced systemic resistance jasmonic acid mediated signaling pathway
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0009867	jasmonic acid mediated signaling pathway
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0010120	camalexin biosynthetic process
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0010200	response to chitin
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0010310	regulation of hydrogen peroxide metabolic process
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0010363	regulation of plant-type hypersensitive response
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0019684	photosynthesis light reaction
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0031347	regulation of defense response
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0031348	negative regulation of defense response
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0035304	regulation of protein dephosphorylation
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0042742	defense response to bacterium
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0043069	negative regulation of programmed cell death
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0043900	regulation of multi-organism process
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0045088	regulation of innate immune response
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0045892	negative regulation of transcription DNA-dependent
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0050832	defense response to fungus
Lin13.2	Glyma.13g267400	37010187	37012198	GO:0051707	response to other organism
Lin13.2	Glyma.13g267400	37010187	37012198	GO:1900056	
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0000165	MAPKKK cascade
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0001666	response to hypoxia
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0002679	respiratory burst involved in defense response
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0006355	regulation of transcription DNA-dependent
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0006612	protein targeting to membrane
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0006944	cellular membrane fusion
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009409	response to cold
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009595	detection of biotic stimulus
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009617	response to bacterium
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009625	response to insect
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009627	systemic acquired resistance
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009697	salicylic acid biosynthetic process
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009751	response to salicylic acid stimulus
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009753	response to jasmonic acid stimulus
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009759	indole glucosinolate biosynthetic process
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009814	defense response incompatible interaction
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009862	systemic acquired resistance salicylic acid mediated signaling pathway
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009863	salicylic acid mediated signaling pathway
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009864	induced systemic resistance jasmonic acid mediated signaling pathway
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0009867	jasmonic acid mediated signaling pathway
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0010120	camalexin biosynthetic process
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0010200	response to chitin
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0010310	regulation of hydrogen peroxide metabolic process
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0010363	regulation of plant-type hypersensitive response
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0019684	photosynthesis light reaction
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0031347	regulation of defense response
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0031348	negative regulation of defense response
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0035304	regulation of protein dephosphorylation
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0042742	defense response to bacterium
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0043069	negative regulation of programmed cell death
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0043900	regulation of multi-organism process
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0045088	regulation of innate immune response
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0045892	negative regulation of transcription DNA-dependent
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0050832	defense response to fungus
Lin13.2	Glyma.13g267700	37035749	37037544	GO:0051707	response to other organism
Lin13.2	Glyma.13g267700	37035749	37037544	GO:1900056	
Lin13.2	Glyma.13g267900	37049135	37051046	GO:0006541	glutamine metabolic process
Lin13.2	Glyma.13g267900	37049135	37051046	GO:0046900	tetrahydrofolylpolyglutamate metabolic process
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0000278	mitotic cell cycle
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0000724	double-strand break repair via homologous recombination
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0006261	DNA-dependent DNA replication
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0006270	DNA-dependent DNA replication initiation
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0006271	DNA strand elongation involved in DNA replication
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0006275	regulation of DNA replication
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0006306	DNA methylation
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0006355	regulation of transcription DNA-dependent
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0009555	pollen development
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0016444	somatic cell DNA recombination
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0016568	chromatin modification
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0031047	gene silencing by RNA

Table 4.8 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0043687	post-translational protein modification
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0045893	positive regulation of transcription DNA-dependent
Lin13.2	Glyma.13g268300	37065946	37070409	GO:0051726	regulation of cell cycle
Lin13.2	Glyma.13g268500	37082792	37087335	GO:0000394	RNA splicing via endonucleolytic cleavage and ligation
Lin13.2	Glyma.13g268500	37082792	37087335	GO:0008150	biological_process
Lin13.2	Glyma.13g268500	37082792	37087335	GO:0009086	methionine biosynthetic process
Lin13.2	Glyma.13g268500	37082792	37087335	GO:0009616	virus induced gene silencing
Lin13.2	Glyma.13g268500	37082792	37087335	GO:0010050	vegetative phase change
Lin13.2	Glyma.13g269800	37187223	37191486	GO:0008150	biological_process
Lin13.2	Glyma.13g269900	37192682	37196249	GO:0006569	tryptophan catabolic process
Lin13.2	Glyma.13g269900	37192682	37196249	GO:0006779	porphyrin biosynthetic process
Lin13.2	Glyma.13g269900	37192682	37196249	GO:0009684	indoleacetic acid biosynthetic process
Lin13.2	Glyma.13g269900	37192682	37196249	GO:0046686	response to cadmium ion
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0006096	glycolysis
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0006833	water transport
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0006972	hyperosmotic response
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0007030	Golgi organization
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0009266	response to temperature stimulus
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0009651	response to salt stress
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0009750	response to fructose stimulus
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0019288	isopentenyl diphosphate biosynthetic process mevalonate-independent pathway
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0019344	cysteine biosynthetic process
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0032880	regulation of protein localization
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0042744	hydrogen peroxide catabolic process
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0046686	response to cadmium ion
Lin13.2	Glyma.13g270600	37265741	37269626	GO:0048528	post-embryonic root development
Lin13.2	Glyma.13g270900	37283330	37287169	GO:0006970	response to osmotic stress
Lin13.2	Glyma.13g270900	37283330	37287169	GO:0009651	response to salt stress
Lin13.2	Glyma.13g270900	37283330	37287169	GO:0009737	response to abscisic acid stimulus
Lin13.2	Glyma.13g270900	37283330	37287169	GO:0009789	positive regulation of abscisic acid mediated signaling pathway
Lin13.2	Glyma.13g271100	37313142	37323355	GO:0006355	regulation of transcription DNA-dependent
Lin13.2	Glyma.13g271100	37313142	37323355	GO:0007267	cell-cell signaling
Lin13.2	Glyma.13g271100	37313142	37323355	GO:0009616	virus induced gene silencing
Lin13.2	Glyma.13g271100	37313142	37323355	GO:0010267	production of ta-siRNAs involved in RNA interference
Lin13.2	Glyma.13g271100	37313142	37323355	GO:0035196	production of miRNAs involved in gene silencing by miRNA
Lin13.2	Glyma.13g271100	37313142	37323355	GO:0043971	histone H3-K18 acetylation
Lin13.2	Glyma.13g271100	37313142	37323355	GO:0043972	histone H3-K23 acetylation
Lin13.2	Glyma.13g271100	37313142	37323355	GO:0044030	regulation of DNA methylation
Lin13.2	Glyma.13g271100	37313142	37323355	GO:0044154	histone H3-K14 acetylation
Lin13.2	Glyma.13g271100	37313142	37323355	GO:0080188	
Lin13.2	Glyma.13g271600	37349043	37349282	GO:0007165	signal transduction
Lin13.2	Glyma.13g271800	37378417	37382805	GO:0008150	biological_process
Lin13.2	Glyma.13g271800	37378417	37382805	GO:0010413	glucuronoxylan metabolic process
Lin13.2	Glyma.13g271800	37378417	37382805	GO:0045492	xylan biosynthetic process
Lin13.2	Glyma.13g272000	37395588	37400547	GO:0007030	Golgi organization
Lin13.2	Glyma.13g272000	37395588	37400547	GO:0007033	vacuole organization
Lin13.2	Glyma.13g272000	37395588	37400547	GO:0042744	hydrogen peroxide catabolic process
Lin13.2	Glyma.13g272000	37395588	37400547	GO:0055074	calcium ion homeostasis
Lin13.2	Glyma.13g272000	37395588	37400547	GO:0071472	cellular response to salt stress
Lin13.2	Glyma.13g272100	37405827	37414114	GO:0007030	Golgi organization
Lin13.2	Glyma.13g272100	37405827	37414114	GO:0007033	vacuole organization
Lin13.2	Glyma.13g272100	37405827	37414114	GO:0042744	hydrogen peroxide catabolic process
Lin13.2	Glyma.13g272100	37405827	37414114	GO:0055074	calcium ion homeostasis
Lin13.2	Glyma.13g272100	37405827	37414114	GO:0071472	cellular response to salt stress
Lin13.2	Glyma.13g272300	37430794	37435987	GO:0007030	Golgi organization
Lin13.2	Glyma.13g272300	37430794	37435987	GO:0007033	vacuole organization
Lin13.2	Glyma.13g272300	37430794	37435987	GO:0042744	hydrogen peroxide catabolic process
Lin13.2	Glyma.13g272300	37430794	37435987	GO:0055074	calcium ion homeostasis
Lin13.2	Glyma.13g272300	37430794	37435987	GO:0071472	cellular response to salt stress
Lin13.2	Glyma.13g272600	37457745	37464257	GO:0008150	biological_process
Lin13.2	Glyma.13g273000	37489719	37492339	GO:0010207	photosystem II assembly
Lin13.2	Glyma.13g273200	37505000	37506646	GO:0008150	biological_process
Lin13.2	Glyma.13g273300	37510973	37519007	GO:0000226	microtubule cytoskeleton organization
Lin13.2	Glyma.13g273300	37510973	37519007	GO:0000911	cytokinesis by cell plate formation
Lin13.2	Glyma.13g273300	37510973	37519007	GO:0009887	organ morphogenesis
Lin13.2	Glyma.13g273300	37510973	37519007	GO:0009888	tissue development
Lin13.2	Glyma.13g273300	37510973	37519007	GO:0010638	positive regulation of organelle organization
Lin13.2	Glyma.13g273300	37510973	37519007	GO:0033044	regulation of chromosome organization
Lin13.2	Glyma.13g273600	37528790	37531199	GO:0008150	biological_process
Lin13.2	Glyma.13g274300	37580356	37583227	GO:0006355	regulation of transcription DNA-dependent
Lin13.2	Glyma.13g274300	37580356	37583227	GO:0007275	multicellular organismal development
Lin13.2	Glyma.13g274300	37580356	37583227	GO:0009965	leaf morphogenesis
Lin13.2	Glyma.13g274300	37580356	37583227	GO:0010072	primary shoot apical meristem specification
Lin13.2	Glyma.13g274300	37580356	37583227	GO:0010160	formation of organ boundary
Lin13.2	Glyma.13g274300	37580356	37583227	GO:0010223	secondary shoot formation

Table 4.8 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Lin13.2	Glyma.13g274300	37580356	37583227	GO:0048366	leaf development
Lin13.2	Glyma.13g274300	37580356	37583227	GO:0048504	regulation of timing of organ formation
Lin13.2	Glyma.13g274600	37622655	37623689	GO:0009791	post-embryonic development
Lin13.2	Glyma.13g275600	37706574	37710287	GO:0006855	drug transmembrane transport
Lin13.2	Glyma.13g275600	37706574	37710287	GO:0009636	response to toxin
Lin13.2	Glyma.13g275600	37706574	37710287	GO:0055085	transmembrane transport
Lin13.2	Glyma.13g276000	37747241	37754663	GO:0008152	metabolic process
Lin13.2	Glyma.13g276000	37747241	37754663	GO:0019509	L-methionine salvage from methylthioadenosine
Lin13.2	Glyma.13g276000	37747241	37754663	GO:0071267	L-methionine salvage
Lin13.2	Glyma.13g276400	37787475	37792273	GO:0008150	biological_process
Lin13.2	Glyma.13g276700	37813197	37819453	GO:0006468	protein phosphorylation
Lin13.2	Glyma.13g277500	37883863	37887320	GO:0008150	biological_process
Lin13.2	Glyma.13g277800	37899444	37901121	GO:0007017	microtubule-based process
Lin13.2	Glyma.13g278200	37930619	37937728	GO:0008150	biological_process
Lin13.2	Glyma.13g278300	37941278	37948343	GO:0008150	biological_process
Lin13.2	Glyma.13g278400	37951890	37960116	GO:0006952	defense response
Lin13.2	Glyma.13g278400	37951890	37960116	GO:0008219	cell death
Lin13.2	Glyma.13g278600	37972164	37977113	GO:0005975	carbohydrate metabolic process
Lin13.2	Glyma.13g278600	37972164	37977113	GO:0009809	lignin biosynthetic process
Prot13.2	Glyma.13g290600	39081781	39082309	GO:0008150	biological_process
Prot13.2	Glyma.13g294800	39420006	39423121	GO:0008150	biological_process
Prot13.2	Glyma.13g294900	39424826	39427503	GO:0009058	biosynthetic process
Prot13.2	Glyma.13g294900	39424826	39427503	GO:0010189	vitamin E biosynthetic process
Prot13.2	Glyma.13g296700	39572455	39573737	GO:0008150	biological_process
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0006811	ion transport
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0006816	calcium ion transport
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0006944	cellular membrane fusion
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0007030	Golgi organization
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0007033	vacuole organization
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0009556	microsporogenesis
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0009651	response to salt stress
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0009845	seed germination
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0010119	regulation of stomatal movement
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0019722	calcium-mediated signaling
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0052543	callose deposition in cell wall
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0055085	transmembrane transport
Prot13.2	Glyma.13g296900	39576148	39594946	GO:0080141	regulation of jasmonic acid biosynthetic process
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0000003	reproduction
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0006075	13-beta-D-glucan biosynthetic process
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0006612	protein targeting to membrane
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0006820	anion transport
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0006862	nucleotide transport
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0006888	ER to Golgi vesicle-mediated transport
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0006944	cellular membrane fusion
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0006952	defense response
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0009555	pollen development
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0009556	microsporogenesis
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0009620	response to fungus
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0009863	salicylic acid mediated signaling pathway
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0009870	defense response signaling pathway resistance gene-dependent
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0009965	leaf morphogenesis
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0010150	leaf senescence
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0010363	regulation of plant-type hypersensitive response
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0015696	ammonium transport
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0015802	basic amino acid transport
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0042742	defense response to bacterium
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0043069	negative regulation of programmed cell death
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0043090	amino acid import
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0043269	regulation of ion transport
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0050832	defense response to fungus
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0052542	defense response by callose deposition
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0052543	callose deposition in cell wall
Prot13.2	Glyma.13g297100	39600226	39605930	GO:0052544	defense response by callose deposition in cell wall
Prot13.2	Glyma.13g297200	39606796	39610638	GO:0008150	biological_process
Prot13.2	Glyma.13g297500	39632130	39640099	GO:0008150	biological_process
Prot13.2	Glyma.13g297800	39650874	39653153	GO:0008150	biological_process
Prot13.2	Glyma.13g297800	39650874	39653153	GO:0009560	embryo sac egg cell differentiation
Prot13.2	Glyma.13g298000	39657343	39660414	GO:0008150	biological_process
Prot13.2	Glyma.13g299400	39766325	39767241	GO:0008150	biological_process
Prot13.2	Glyma.13g299400	39766325	39767241	GO:0031540	regulation of anthocyanin biosynthetic process
Prot13.2	Glyma.13g303500	40051398	40053944	GO:0009793	embryo development ending in seed dormancy
Prot13.2	Glyma.13g308500	40413291	40419050	GO:0006810	transport
Prot13.2	Glyma.13g308500	40413291	40419050	GO:0006874	cellular calcium ion homeostasis
Prot13.2	Glyma.13g308500	40413291	40419050	GO:0007186	G-protein coupled receptor protein signaling pathway
Prot13.2	Glyma.13g308500	40413291	40419050	GO:0009416	response to light stimulus

Table 4.9 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 14 based on Glyma 2.0 position.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Yld14	Glyma.14g012900	971984	978686	GO:0000271	polysaccharide biosynthetic process
Yld14	Glyma.14g012900	971984	978686	GO:0009832	plant-type cell wall biogenesis
Yld14	Glyma.14g012900	971984	978686	GO:0030244	cellulose biosynthetic process
Yld14	Glyma.14g012900	971984	978686	GO:0046482	para-aminobenzoic acid metabolic process
Yld14	Glyma.14g013100	983795	987274	GO:0006355	regulation of transcription DNA-dependent
Yld14	Glyma.14g013100	983795	987274	GO:0006364	rRNA processing
Yld14	Glyma.14g013100	983795	987274	GO:0006399	tRNA metabolic process
Yld14	Glyma.14g013100	983795	987274	GO:0009658	chloroplast organization
Yld14	Glyma.14g013100	983795	987274	GO:0045036	protein targeting to chloroplast
Yld14	Glyma.14g013500	1009519	1014238	GO:0008152	metabolic process
Yld14	Glyma.14g014000	1029711	1032256	GO:0006520	cellular amino acid metabolic process
Yld14	Glyma.14g014000	1029711	1032256	GO:0006535	cysteine biosynthetic process from serine
Yld14	Glyma.14g014000	1029711	1032256	GO:0008152	metabolic process
Yld14	Glyma.14g014000	1029711	1032256	GO:0048573	photoperiodism flowering
Yld14	Glyma.14g014100	1033723	1036812	GO:0000271	polysaccharide biosynthetic process
Yld14	Glyma.14g014100	1033723	1036812	GO:0006816	calcium ion transport
Yld14	Glyma.14g014100	1033723	1036812	GO:0007030	Golgi organization
Yld14	Glyma.14g014100	1033723	1036812	GO:0009651	response to salt stress
Yld14	Glyma.14g014100	1033723	1036812	GO:0010029	regulation of seed germination
Yld14	Glyma.14g014100	1033723	1036812	GO:0010119	regulation of stomatal movement
Yld14	Glyma.14g014100	1033723	1036812	GO:0030003	cellular cation homeostasis
Yld14	Glyma.14g014100	1033723	1036812	GO:0030007	cellular potassium ion homeostasis
Yld14	Glyma.14g014100	1033723	1036812	GO:0051260	protein homooligomerization
Yld14	Glyma.14g014100	1033723	1036812	GO:0070838	divalent metal ion transport
Yld14	Glyma.14g014100	1033723	1036812	GO:0071805	potassium ion transmembrane transport
Yld14	Glyma.14g014200	1042883	1043942	GO:0008150	biological_process
Yld14	Glyma.14g014900	1077772	1080343	GO:0008150	biological_process
Yld14	Glyma.14g015700	1130025	1137147	GO:0008150	biological_process
Yld14	Glyma.14g016100	1150980	1152275	GO:0006355	regulation of transcription DNA-dependent
Yld14	Glyma.14g016100	1150980	1152275	GO:0006417	regulation of translation
Yld14	Glyma.14g016100	1150980	1152275	GO:0009657	plastid organization
Yld14	Glyma.14g016100	1150980	1152275	GO:0009965	leaf morphogenesis
Yld14	Glyma.14g016100	1150980	1152275	GO:0030154	cell differentiation
Yld14	Glyma.14g016100	1150980	1152275	GO:0045893	positive regulation of transcription DNA-dependent
Yld14	Glyma.14g016200	1154838	1160434	GO:0006355	regulation of transcription DNA-dependent
Yld14	Glyma.14g016200	1154838	1160434	GO:0009611	response to wounding
Yld14	Glyma.14g016200	1154838	1160434	GO:0009616	virus induced gene silencing
Yld14	Glyma.14g016200	1154838	1160434	GO:0009961	response to 1-aminocyclopropane-1-carboxylic acid
Yld14	Glyma.14g016200	1154838	1160434	GO:0010050	vegetative phase change
Yld14	Glyma.14g016200	1154838	1160434	GO:0045893	positive regulation of transcription DNA-dependent
Yld14	Glyma.14g016300	1170626	1173884	GO:0002679	respiratory burst involved in defense response
Yld14	Glyma.14g016300	1170626	1173884	GO:0006355	regulation of transcription DNA-dependent
Yld14	Glyma.14g016300	1170626	1173884	GO:0006612	protein targeting to membrane
Yld14	Glyma.14g016300	1170626	1173884	GO:0006944	cellular membrane fusion
Yld14	Glyma.14g016300	1170626	1173884	GO:0009409	response to cold
Yld14	Glyma.14g016300	1170626	1173884	GO:0009863	salicylic acid mediated signaling pathway
Yld14	Glyma.14g016300	1170626	1173884	GO:0010200	response to chitin
Yld14	Glyma.14g016300	1170626	1173884	GO:0010363	regulation of plant-type hypersensitive response
Yld14	Glyma.14g016300	1170626	1173884	GO:0030968	endoplasmic reticulum unfolded protein response
Yld14	Glyma.14g016300	1170626	1173884	GO:0035556	intracellular signal transduction
Yld14	Glyma.14g016300	1170626	1173884	GO:0043069	negative regulation of programmed cell death
Yld14	Glyma.14g016300	1170626	1173884	GO:0050832	defense response to fungus
Yld14	Glyma.14g016400	1177818	1180839	GO:0006468	protein phosphorylation
Yld14	Glyma.14g016400	1177818	1180839	GO:0010260	organ senescence
Yld14	Glyma.14g016800	1206662	1212734	GO:0006260	DNA replication
Yld14	Glyma.14g016800	1206662	1212734	GO:0006270	DNA-dependent DNA replication initiation
Yld14	Glyma.14g016800	1206662	1212734	GO:0006275	regulation of DNA replication
Yld14	Glyma.14g016800	1206662	1212734	GO:0006306	DNA methylation
Yld14	Glyma.14g016800	1206662	1212734	GO:0006342	chromatin silencing
Yld14	Glyma.14g016800	1206662	1212734	GO:0006346	methylation-dependent chromatin silencing
Yld14	Glyma.14g016800	1206662	1212734	GO:0007018	microtubule-based movement
Yld14	Glyma.14g016800	1206662	1212734	GO:0008283	cell proliferation
Yld14	Glyma.14g016800	1206662	1212734	GO:0010089	xylem development
Yld14	Glyma.14g016800	1206662	1212734	GO:0016246	RNA interference
Yld14	Glyma.14g016800	1206662	1212734	GO:0044036	cell wall macromolecule metabolic process
Yld14	Glyma.14g016800	1206662	1212734	GO:0051567	histone H3-K9 methylation
Yld14	Glyma.14g016800	1206662	1212734	GO:0051726	regulation of cell cycle
Yld14	Glyma.14g017000	1218416	1221786	GO:0008150	biological_process
Yld14	Glyma.14g017100	1229583	1236592	GO:0006071	glycerol metabolic process
Yld14	Glyma.14g017100	1229583	1236592	GO:0006629	lipid metabolic process
Yld14	Glyma.14g017100	1229583	1236592	GO:0030643	cellular phosphate ion homeostasis
Yld14	Glyma.14g017600	1257629	1260935	GO:0006184	GTP catabolic process
Yld14	Glyma.14g017600	1257629	1260935	GO:0007017	microtubule-based process
Yld14	Glyma.14g017600	1257629	1260935	GO:0009220	pyrimidine ribonucleotide biosynthetic process
Yld14	Glyma.14g017600	1257629	1260935	GO:0009637	response to blue light
Yld14	Glyma.14g017600	1257629	1260935	GO:0009658	chloroplast organization
Yld14	Glyma.14g017600	1257629	1260935	GO:0009902	chloroplast relocation

Table 4.9 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Yld14	Glyma.14g017600	1257629	1260935	GO:0010020	chloroplast fission
Yld14	Glyma.14g017600	1257629	1260935	GO:0043572	plastid fission
Yld14	Glyma.14g017600	1257629	1260935	GO:0051258	protein polymerization
Yld14	Glyma.14g018000	1281413	1289765	GO:0006355	regulation of transcription DNA-dependent
Yld14	Glyma.14g018100	1299276	1303945	GO:0009944	polarity specification of adaxial/abaxial axis
Yld14	Glyma.14g018100	1299276	1303945	GO:0010413	glucuronoxylan metabolic process
Yld14	Glyma.14g018100	1299276	1303945	GO:0045492	xylan biosynthetic process
Yld14	Glyma.14g018500	1333663	1336745	GO:0000272	polysaccharide catabolic process
Yld14	Glyma.14g018500	1333663	1336745	GO:0000910	cytokinesis
Yld14	Glyma.14g018500	1333663	1336745	GO:0001578	microtubule bundle formation
Yld14	Glyma.14g018500	1333663	1336745	GO:0005982	starch metabolic process
Yld14	Glyma.14g018500	1333663	1336745	GO:0007020	microtubule nucleation
Yld14	Glyma.14g018500	1333663	1336745	GO:0008283	cell proliferation
Yld14	Glyma.14g018500	1333663	1336745	GO:0009664	plant-type cell wall organization
Yld14	Glyma.14g018500	1333663	1336745	GO:0019344	cysteine biosynthetic process
Yld14	Glyma.14g018500	1333663	1336745	GO:0030244	cellulose biosynthetic process
Yld14	Glyma.14g018500	1333663	1336745	GO:0031116	positive regulation of microtubule polymerization
Yld14	Glyma.14g018500	1333663	1336745	GO:0043622	cortical microtubule organization
Yld14	Glyma.14g018500	1333663	1336745	GO:0046785	microtubule polymerization
Yld14	Glyma.14g018500	1333663	1336745	GO:0048193	Golgi vesicle transport
Yld14	Glyma.14g018500	1333663	1336745	GO:0048528	post-embryonic root development
Yld14	Glyma.14g018500	1333663	1336745	GO:0051302	regulation of cell division
Yld14	Glyma.14g018500	1333663	1336745	GO:0051322	anaphase
Yld14	Glyma.14g018700	1346104	1352242	GO:0000023	maltose metabolic process
Yld14	Glyma.14g018700	1346104	1352242	GO:0006457	protein folding
Yld14	Glyma.14g018700	1346104	1352242	GO:0009902	chloroplast relocation
Yld14	Glyma.14g018700	1346104	1352242	GO:0010027	thylakoid membrane organization
Yld14	Glyma.14g018700	1346104	1352242	GO:0010103	stomatal complex morphogenesis
Yld14	Glyma.14g018700	1346104	1352242	GO:0015031	protein transport
Yld14	Glyma.14g018700	1346104	1352242	GO:0016117	carotenoid biosynthetic process
Yld14	Glyma.14g018700	1346104	1352242	GO:0016556	mRNA modification
Yld14	Glyma.14g018700	1346104	1352242	GO:0019252	starch biosynthetic process
Yld14	Glyma.14g018700	1346104	1352242	GO:0019288	isopentenyl diphosphate biosynthetic process mevalonate-independent pathway
Yld14	Glyma.14g018700	1346104	1352242	GO:0034660	ncRNA metabolic process
Yld14	Glyma.14g018900	1360753	1361467	GO:0006869	lipid transport
Yld14	Glyma.14g019400	1384203	1396383	GO:0006355	regulation of transcription DNA-dependent
Yld14	Glyma.14g019400	1384203	1396383	GO:0010093	specification of floral organ identity
Yld14	Glyma.14g019400	1384203	1396383	GO:0048440	carpel development
Yld14	Glyma.14g019400	1384203	1396383	GO:0048441	petal development
Yld14	Glyma.14g019400	1384203	1396383	GO:0048443	stamen development
Yld14	Glyma.14g019400	1384203	1396383	GO:0048481	ovule development
Yld14	Glyma.14g020300	1442687	1444895	GO:0042545	cell wall modification
Yld14	Glyma.14g021200	1489197	1494157	GO:0006396	RNA processing
Yld14	Glyma.14g021300	1497987	1512889	GO:0042732	D-xylose metabolic process
Yld14	Glyma.14g022000	1553534	1554901	GO:0006886	intracellular protein transport
Yld14	Glyma.14g022400	1578265	1594980	GO:0000710	meiotic mismatch repair
Yld14	Glyma.14g022400	1578265	1594980	GO:0006200	ATP catabolic process
Yld14	Glyma.14g022400	1578265	1594980	GO:0006298	mismatch repair
Yld14	Glyma.14g022400	1578265	1594980	GO:0006302	double-strand break repair
Yld14	Glyma.14g022400	1578265	1594980	GO:0006312	mitotic recombination
Yld14	Glyma.14g022400	1578265	1594980	GO:0007059	chromosome segregation
Yld14	Glyma.14g022400	1578265	1594980	GO:0007062	sister chromatid cohesion
Yld14	Glyma.14g022400	1578265	1594980	GO:0007129	synapsis
Yld14	Glyma.14g022400	1578265	1594980	GO:0007131	reciprocal meiotic recombination
Yld14	Glyma.14g022400	1578265	1594980	GO:0009691	cytokinin biosynthetic process
Yld14	Glyma.14g022400	1578265	1594980	GO:0010332	response to gamma radiation
Yld14	Glyma.14g022400	1578265	1594980	GO:0032204	regulation of telomere maintenance
Yld14	Glyma.14g022400	1578265	1594980	GO:0032504	multicellular organism reproduction
Yld14	Glyma.14g022400	1578265	1594980	GO:0042138	meiotic DNA double-strand break formation
Yld14	Glyma.14g022400	1578265	1594980	GO:0043247	telomere maintenance in response to DNA damage
Yld14	Glyma.14g022400	1578265	1594980	GO:0045132	meiotic chromosome segregation
Yld14	Glyma.14g022400	1578265	1594980	GO:0045143	homologous chromosome segregation
Yld14	Glyma.14g022400	1578265	1594980	GO:0051026	chiasma assembly
Yld14	Glyma.14g022600	1604907	1612521	GO:0000956	nuclear-transcribed mRNA catabolic process
Yld14	Glyma.14g022600	1604907	1612521	GO:0008150	biological_process
Yld14	Glyma.14g023600	1691809	1693795	GO:0006094	gluconeogenesis
Yld14	Glyma.14g023600	1691809	1693795	GO:0006096	glycolysis
Yld14	Glyma.14g023600	1691809	1693795	GO:0009651	response to salt stress
Yld14	Glyma.14g023600	1691809	1693795	GO:0019761	glucosinolate biosynthetic process
Yld14	Glyma.14g023600	1691809	1693795	GO:0046686	response to cadmium ion
Yld14	Glyma.14g023900	1707404	1710941	GO:0006464	protein modification process
Yld14	Glyma.14g023900	1707404	1710941	GO:0006546	glycine catabolic process
Yld14	Glyma.14g023900	1707404	1710941	GO:0009416	response to light stimulus
Yld14	Glyma.14g023900	1707404	1710941	GO:0019538	protein metabolic process
Yld14	Glyma.14g023900	1707404	1710941	GO:0034599	cellular response to oxidative stress
Yld14	Glyma.14g023900	1707404	1710941	GO:0055114	oxidation-reduction process
Yld14	Glyma.14g024000	1718621	1723131	GO:0007155	cell adhesion
Yld14	Glyma.14g024000	1718621	1723131	GO:0008150	biological_process
Yld14	Glyma.14g024000	1718621	1723131	GO:0010090	trichome morphogenesis
Yld14	Glyma.14g024000	1718621	1723131	GO:0045010	actin nucleation
Yld14	Glyma.14g024000	1718621	1723131	GO:0048765	root hair cell differentiation
Yld14	Glyma.14g024000	1718621	1723131	GO:0071555	cell wall organization
Yld14	Glyma.14g024100	1724080	1727463	GO:0008150	biological_process
Yld14	Glyma.14g024100	1724080	1727463	GO:0009640	photomorphogenesis

Table 4.9 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Yld14	Glyma.14g024100	1724080	1727463	GO:0009793	embryo development ending in seed dormancy
Yld14	Glyma.14g024100	1724080	1727463	GO:0009845	seed germination
Yld14	Glyma.14g024100	1724080	1727463	GO:0009909	regulation of flower development
Yld14	Glyma.14g024100	1724080	1727463	GO:0009933	meristem structural organization
Yld14	Glyma.14g024100	1724080	1727463	GO:0010162	seed dormancy
Yld14	Glyma.14g024100	1724080	1727463	GO:0010182	sugar mediated signaling pathway
Yld14	Glyma.14g024100	1724080	1727463	GO:0010228	vegetative to reproductive phase transition of meristem
Yld14	Glyma.14g024100	1724080	1727463	GO:0016567	protein ubiquitination
Yld14	Glyma.14g024100	1724080	1727463	GO:0019915	lipid storage
Yld14	Glyma.14g024100	1724080	1727463	GO:0050826	response to freezing
Yld14	Glyma.14g024200	1727864	1733875	GO:0006457	protein folding
Yld14	Glyma.14g024300	1740362	1745905	GO:0006606	protein import into nucleus
Yld14	Glyma.14g024400	1746860	1749862	GO:0006952	defense response
Yld14	Glyma.14g024400	1746860	1749862	GO:0007165	signal transduction
Yld14	Glyma.14g024500	1751063	1753753	GO:0006952	defense response
Yld14	Glyma.14g024500	1751063	1753753	GO:0007165	signal transduction
Yld14	Glyma.14g024600	1755937	1778895	GO:0001510	RNA methylation
Yld14	Glyma.14g024600	1755937	1778895	GO:0009220	pyrimidine ribonucleotide biosynthetic process
Yld14	Glyma.14g024700	1781035	1783960	GO:0015824	proline transport
Yld14	Glyma.14g024800	1785084	1786355	GO:0008150	biological_process
Yld14	Glyma.14g025400	1827049	1830067	GO:0008150	biological_process
Yld14	Glyma.14g025800	1849711	1854842	GO:0006007	glucose catabolic process
Yld14	Glyma.14g025800	1849711	1854842	GO:0006810	transport
Yld14	Glyma.14g025800	1849711	1854842	GO:0006863	purine base transport
Yld14	Glyma.14g025800	1849711	1854842	GO:0007033	vacuole organization
Yld14	Glyma.14g025800	1849711	1854842	GO:0009553	embryo sac development
Yld14	Glyma.14g025800	1849711	1854842	GO:0009624	response to nematode
Yld14	Glyma.14g025800	1849711	1854842	GO:0009790	embryo development
Yld14	Glyma.14g025800	1849711	1854842	GO:0010152	pollen maturation
Yld14	Glyma.14g025800	1849711	1854842	GO:0015713	phosphoglycerate transport
Yld14	Glyma.14g025800	1849711	1854842	GO:0015714	phosphoenolpyruvate transport
Yld14	Glyma.14g025800	1849711	1854842	GO:0015760	glucose-6-phosphate transport
Yld14	Glyma.14g025800	1849711	1854842	GO:0034389	lipid particle organization
Yld14	Glyma.14g025800	1849711	1854842	GO:0035436	triose phosphate transmembrane transport
Yld14	Glyma.14g025800	1849711	1854842	GO:0055085	transmembrane transport
Yld14	Glyma.14g026300	1887286	1898057	GO:0006468	protein phosphorylation
Yld14	Glyma.14g026300	1887286	1898057	GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway
Yld14	Glyma.14g026500	1906384	1909125	GO:0055114	oxidation-reduction process
Yld14	Glyma.14g026600	1908407	1912887	GO:0000394	RNA splicing via endonucleolytic cleavage and ligation
Yld14	Glyma.14g026600	1908407	1912887	GO:0006744	ubiquinone biosynthetic process
Yld14	Glyma.14g026600	1908407	1912887	GO:0008152	metabolic process
Yld14	Glyma.14g026600	1908407	1912887	GO:0009086	methionine biosynthetic process
Yld14	Glyma.14g026600	1908407	1912887	GO:0009793	embryo development ending in seed dormancy
Yld14	Glyma.14g026600	1908407	1912887	GO:0019243	methylglyoxal catabolic process to D-lactate
Yld14	Glyma.14g026700	1915319	1921138	GO:0006468	protein phosphorylation
Yld14	Glyma.14g026700	1915319	1921138	GO:0009966	regulation of signal transduction
Yld14	Glyma.14g027300	1975415	1977737	GO:0006355	regulation of transcription DNA-dependent
Yld14	Glyma.14g027300	1975415	1977737	GO:0010093	specification of floral organ identity
Yld14	Glyma.14g027300	1975415	1977737	GO:0048440	carpel development
Yld14	Glyma.14g027300	1975415	1977737	GO:0048441	petal development
Yld14	Glyma.14g027300	1975415	1977737	GO:0048443	stamen development
Yld14	Glyma.14g027300	1975415	1977737	GO:0048481	ovule development
Yld14	Glyma.14g027600	2004420	2006136	GO:0055114	oxidation-reduction process
Yld14	Glyma.14g027700	2010502	2012080	GO:0005975	carbohydrate metabolic process
Yld14	Glyma.14g027700	2010502	2012080	GO:0006073	cellular glucan metabolic process
Yld14	Glyma.14g029200	2136060	2141071	GO:0000271	polysaccharide biosynthetic process
Yld14	Glyma.14g029200	2136060	2141071	GO:0009832	plant-type cell wall biogenesis
Yld14	Glyma.14g029200	2136060	2141071	GO:0009846	pollen germination
Yld14	Glyma.14g029200	2136060	2141071	GO:0030244	cellulose biosynthetic process
Yld14	Glyma.14g031600	2300463	2308888	GO:0000226	microtubule cytoskeleton organization
Yld14	Glyma.14g031600	2300463	2308888	GO:0000911	cytokinesis by cell plate formation
Yld14	Glyma.14g032100	2331596	2334547	GO:0008150	biological_process
Yld14	Glyma.14g032700	2372881	2381492	GO:0006355	regulation of transcription DNA-dependent
Yld14	Glyma.14g032700	2372881	2381492	GO:0009725	response to hormone stimulus
Yld14	Glyma.14g032700	2372881	2381492	GO:0009733	response to auxin stimulus
Yld14	Glyma.14g032700	2372881	2381492	GO:0009855	determination of bilateral symmetry
Yld14	Glyma.14g032700	2372881	2381492	GO:0009886	post-embryonic morphogenesis
Yld14	Glyma.14g032700	2372881	2381492	GO:0009887	organ morphogenesis
Yld14	Glyma.14g032700	2372881	2381492	GO:0009908	flower development
Yld14	Glyma.14g032700	2372881	2381492	GO:0009909	regulation of flower development
Yld14	Glyma.14g032700	2372881	2381492	GO:0010051	xylem and phloem pattern formation
Yld14	Glyma.14g032700	2372881	2381492	GO:0010154	fruit development
Yld14	Glyma.14g032700	2372881	2381492	GO:0048439	flower morphogenesis
Yld14	Glyma.14g032700	2372881	2381492	GO:0048481	ovule development
Yld14	Glyma.14g032700	2372881	2381492	GO:0048507	meristem development
Yld14	Glyma.14g032700	2372881	2381492	GO:0048519	negative regulation of biological process
Yld14	Glyma.14g036800	2752036	2776640	GO:0008150	biological_process
Yld14	Glyma.14g037200	2800964	2806874	GO:0007015	actin filament organization
Yld14	Glyma.14g037200	2800964	2806874	GO:0010119	regulation of stomatal movement
Yld14	Glyma.14g037200	2800964	2806874	GO:0032880	regulation of protein localization
Ste14	Glyma.14g124400	19061201	19065418	GO:0006468	protein phosphorylation
Ste14	Glyma.14g124400	19061201	19065418	GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway
Ste14	Glyma.14g124400	19061201	19065418	GO:0010075	regulation of meristem growth
Ste14	Glyma.14g125000	19605857	19606066	GO:0008150	biological_process

Table 4.10 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 17 based on Glyma 2.0 position.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Len17; Pal17.1; Ole17	Glyma.17g019600	1470065	1471500	GO:0008152	metabolic process
Len17; Pal17.1; Ole17	Glyma.17g021400	1566684	1575389	GO:0006148	inosine catabolic process
Len17; Pal17.1; Ole17	Glyma.17g021400	1566684	1575389	GO:0006154	adenosine catabolic process
Len17; Pal17.1; Ole17	Glyma.17g021400	1566684	1575389	GO:0009611	response to wounding
Len17; Pal17.1; Ole17	Glyma.17g021400	1566684	1575389	GO:0009753	response to jasmonic acid stimulus
Len17; Pal17.1; Ole17	Glyma.17g027600	2043940	2051649	GO:0006631	fatty acid metabolic process
Len17; Pal17.1; Ole17	Glyma.17g027600	2043940	2051649	GO:0006635	fatty acid beta-oxidation
Len17; Pal17.1; Ole17	Glyma.17g027600	2043940	2051649	GO:0007031	peroxisome organization
Len17; Pal17.1; Ole17	Glyma.17g027600	2043940	2051649	GO:0008152	metabolic process
Len17; Pal17.1; Ole17	Glyma.17g027600	2043940	2051649	GO:0009062	fatty acid catabolic process
Len17; Pal17.1; Ole17	Glyma.17g027600	2043940	2051649	GO:0009407	toxin catabolic process
Len17; Pal17.1; Ole17	Glyma.17g027600	2043940	2051649	GO:0043161	proteasomal ubiquitin-dependent protein catabolic process
Len17; Pal17.1; Ole17	Glyma.17g027600	2043940	2051649	GO:0051788	response to misfolded protein
Len17; Pal17.1; Ole17	Glyma.17g027600	2043940	2051649	GO:0055114	oxidation-reduction process
Len17; Pal17.1; Ole17	Glyma.17g027600	2043940	2051649	GO:0080129	proteasome core complex assembly
Len17; Pal17.1; Ole17	Glyma.17g029000	2147897	2150720	GO:0008150	biological_process
Len17; Pal17.1; Ole17	Glyma.17g029100	2151135	2153011	GO:0008150	biological_process
Pal17.1; Ole17	Glyma.17g031700	2318958	2327179	GO:0006897	endocytosis
Pal17.1; Ole17	Glyma.17g031700	2318958	2327179	GO:0007033	vacuole organization
Pal17.1; Ole17	Glyma.17g031700	2318958	2327179	GO:0044090	positive regulation of vacuole organization
Pal17.1; Ole17	Glyma.17g031700	2318958	2327179	GO:0046907	intracellular transport
Pal17.1; Ole17	Glyma.17g031700	2318958	2327179	GO:0070536	protein K63-linked deubiquitination
Pal17.1; Ole17	Glyma.17g031700	2318958	2327179	GO:0071108	protein K48-linked deubiquitination
Pal17.1; Ole17	Glyma.17g031700	2318958	2327179	GO:0090316	positive regulation of intracellular protein transport
Pal17.1; Ole17	Glyma.17g031800	2331626	2336016	GO:0006083	acetate metabolic process
Pal17.1; Ole17	Glyma.17g031800	2331626	2336016	GO:0006097	glyoxylate cycle
Pal17.1; Ole17	Glyma.17g031800	2331626	2336016	GO:0008152	metabolic process
Pal17.1; Ole17	Glyma.17g031800	2331626	2336016	GO:0019605	butyrate metabolic process
Pal17.1; Ole17	Glyma.17g032400	2376469	2391870	GO:0008150	biological_process
Pal17.1; Ole17	Glyma.17g032400	2376469	2391870	GO:0009630	gravitropism
Pal17.1; Ole17	Glyma.17g032400	2376469	2391870	GO:0009639	response to red or far red light
Pal17.1; Ole17	Glyma.17g032900	2408152	2410553	GO:0008150	biological_process
Pal17.1; Ole17	Glyma.17g033000	2412135	2414981	GO:0006865	amino acid transport
Pal17.1; Ole17	Glyma.17g033000	2412135	2414981	GO:0008150	biological_process
Pal17.1; Ole17	Glyma.17g033200	2425245	2431485	GO:0006261	DNA-dependent DNA replication
Pal17.1; Ole17	Glyma.17g033200	2425245	2431485	GO:0008150	biological_process
Pal17.1; Ole17	Glyma.17g033200	2425245	2431485	GO:0009560	embryo sac egg cell differentiation
Pal17.1; Ole17	Glyma.17g034100	2492326	2494702	GO:0006635	fatty acid beta-oxidation
Pal17.1; Ole17	Glyma.17g034100	2492326	2494702	GO:0008152	metabolic process
Pal17.2	Glyma.17g085100	6564410	6568425	GO:0000023	maltose metabolic process
Pal17.2	Glyma.17g085100	6564410	6568425	GO:0016126	sterol biosynthetic process
Pal17.2	Glyma.17g085100	6564410	6568425	GO:0019252	starch biosynthetic process
Pal17.2	Glyma.17g085100	6564410	6568425	GO:0019745	pentacyclic triterpenoid biosynthetic process
Pal17.2	Glyma.17g085100	6564410	6568425	GO:0019761	glucosinolate biosynthetic process
Pal17.2	Glyma.17g085100	6564410	6568425	GO:0055114	oxidation-reduction process
Pal17.2	Glyma.17g085600	6609185	6610514	GO:0045892	negative regulation of transcription DNA-dependent
Pal17.2	Glyma.17g091000	7090659	7095786	GO:0008152	metabolic process
Pal17.2	Glyma.17g091000	7090659	7095786	GO:0046686	response to cadmium ion
Pal17.2	Glyma.17g091000	7090659	7095786	GO:0055114	oxidation-reduction process
Pal17.2	Glyma.17g095800	7499629	7509122	GO:0006511	ubiquitin-dependent protein catabolic process
Pal17.2	Glyma.17g098300	7744245	7753347	GO:0009407	toxin catabolic process
Pal17.2	Glyma.17g098300	7744245	7753347	GO:0010583	response to cyclopentenone
Pal17.2	Glyma.17g099300	7838198	7840795	GO:0006355	regulation of transcription DNA-dependent
Pal17.2	Glyma.17g099300	7838198	7840795	GO:0008284	positive regulation of cell proliferation
Pal17.2	Glyma.17g099300	7838198	7840795	GO:0048573	photoperiodism flowering
Pal17.2	Glyma.17g100100	7899983	7902689	GO:0008150	biological_process
Pal17.2	Glyma.17g100300	7907773	7909350	GO:0000741	karyogamy
Pal17.2	Glyma.17g100300	7907773	7909350	GO:0006606	protein import into nucleus
Pal17.2	Glyma.17g100300	7907773	7909350	GO:0009560	embryo sac egg cell differentiation
Pal17.2	Glyma.17g100300	7907773	7909350	GO:0009640	photomorphogenesis
Pal17.2	Glyma.17g100300	7907773	7909350	GO:0009790	embryo development
Pal17.2	Glyma.17g100300	7907773	7909350	GO:0010388	cullin deneddylation
Pal17.2	Glyma.17g100300	7907773	7909350	GO:0017126	nucleologenesis
Pal17.2	Glyma.17g100300	7907773	7909350	GO:0048825	cotyledon development
Pal17.2	Glyma.17g100300	7907773	7909350	GO:0051301	cell division
Pal17.2	Glyma.17g100300	7907773	7909350	GO:0051302	regulation of cell division
Pal17.2	Glyma.17g100600	7917204	7919329	GO:0005975	carbohydrate metabolic process
Pal17.2	Glyma.17g100600	7917204	7919329	GO:0006108	malate metabolic process
Pal17.2	Glyma.17g100600	7917204	7919329	GO:0009409	response to cold
Pal17.2	Glyma.17g100600	7917204	7919329	GO:0044262	cellular carbohydrate metabolic process
Pal17.2	Glyma.17g100600	7917204	7919329	GO:0055114	oxidation-reduction process
Pal17.2	Glyma.17g102800	8045209	8056758	GO:0016043	cellular component organization
Pal17.2	Glyma.17g102800	8045209	8056758	GO:0030036	actin cytoskeleton organization

Table 4.11 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 1 mb from associated QTL on soybean chromosome 19 based on Glyma 2.0 position.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Lin19; Ole19; Len19	Glyma.19g016400	1631472	1639538	GO:0006200	ATP catabolic process
Lin19; Ole19; Len19	Glyma.19g016400	1631472	1639538	GO:0006810	transport
Lin19; Ole19; Len19	Glyma.19g016400	1631472	1639538	GO:0010048	vernalization response
Lin19; Ole19; Len19	Glyma.19g016400	1631472	1639538	GO:0010315	auxin efflux
Lin19; Ole19; Len19	Glyma.19g016400	1631472	1639538	GO:0010540	basipetal auxin transport
Lin19; Ole19; Len19	Glyma.19g016400	1631472	1639538	GO:0010541	acropetal auxin transport
Lin19; Ole19; Len19	Glyma.19g016400	1631472	1639538	GO:0043481	anthocyanin accumulation in tissues in response to UV light
Lin19; Ole19; Len19	Glyma.19g016400	1631472	1639538	GO:0048440	carpel development
Lin19; Ole19; Len19	Glyma.19g016400	1631472	1639538	GO:0055085	transmembrane transport
Lin19; Ole19; Len19	Glyma.19g016600	1691184	1698603	GO:0006200	ATP catabolic process
Lin19; Ole19; Len19	Glyma.19g016600	1691184	1698603	GO:0006810	transport
Lin19; Ole19; Len19	Glyma.19g016600	1691184	1698603	GO:0010048	vernalization response
Lin19; Ole19; Len19	Glyma.19g016600	1691184	1698603	GO:0010315	auxin efflux
Lin19; Ole19; Len19	Glyma.19g016600	1691184	1698603	GO:0010540	basipetal auxin transport
Lin19; Ole19; Len19	Glyma.19g016600	1691184	1698603	GO:0010541	acropetal auxin transport
Lin19; Ole19; Len19	Glyma.19g016600	1691184	1698603	GO:0043481	anthocyanin accumulation in tissues in response to UV light
Lin19; Ole19; Len19	Glyma.19g016600	1691184	1698603	GO:0048440	carpel development
Lin19; Ole19; Len19	Glyma.19g016600	1691184	1698603	GO:0055085	transmembrane transport
Lin19; Ole19; Len19	Glyma.19g016700	1728554	1736171	GO:0006200	ATP catabolic process
Lin19; Ole19; Len19	Glyma.19g016700	1728554	1736171	GO:0006810	transport
Lin19; Ole19; Len19	Glyma.19g016700	1728554	1736171	GO:0010048	vernalization response
Lin19; Ole19; Len19	Glyma.19g016700	1728554	1736171	GO:0010315	auxin efflux
Lin19; Ole19; Len19	Glyma.19g016700	1728554	1736171	GO:0010540	basipetal auxin transport
Lin19; Ole19; Len19	Glyma.19g016700	1728554	1736171	GO:0010541	acropetal auxin transport
Lin19; Ole19; Len19	Glyma.19g016700	1728554	1736171	GO:0043481	anthocyanin accumulation in tissues in response to UV light
Lin19; Ole19; Len19	Glyma.19g016700	1728554	1736171	GO:0048440	carpel development
Lin19; Ole19; Len19	Glyma.19g016700	1728554	1736171	GO:0055085	transmembrane transport
Lin19; Ole19; Len19	Glyma.19g017200	1785304	1799141	GO:0006094	gluconeogenesis
Lin19; Ole19; Len19	Glyma.19g017200	1785304	1799141	GO:0006096	glycolysis
Lin19; Ole19; Len19	Glyma.19g017200	1785304	1799141	GO:0009744	response to sucrose stimulus
Lin19; Ole19; Len19	Glyma.19g017200	1785304	1799141	GO:0009813	flavonoid biosynthetic process
Lin19; Ole19; Len19	Glyma.19g017200	1785304	1799141	GO:0009817	defense response to fungus incompatible interaction
Lin19; Ole19; Len19	Glyma.19g017200	1785304	1799141	GO:0010224	response to UV-B
Lin19; Ole19; Len19	Glyma.19g017200	1785304	1799141	GO:0046686	response to cadmium ion
Lin19; Ole19; Len19	Glyma.19g017300	1814297	1816879	GO:0006468	protein phosphorylation
Lin19; Ole19; Len19	Glyma.19g017300	1814297	1816879	GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway
Lin19; Ole19; Len19	Glyma.19g017300	1814297	1816879	GO:0030968	endoplasmic reticulum unfolded protein response
Lin19; Ole19; Len19	Glyma.19g019400	2066612	2071520	GO:0008150	biological_process
Lin19; Ole19; Len19	Glyma.19g019800	2100931	2104966	GO:0006468	protein phosphorylation
Lin19; Ole19; Len19	Glyma.19g020300	2166516	2169372	GO:0008150	biological_process
Lin19; Ole19; Len19	Glyma.19g020400	2170639	2177437	GO:0000394	RNA splicing via endonucleolytic cleavage and ligation
Lin19; Ole19; Len19	Glyma.19g020400	2170639	2177437	GO:0006355	regulation of transcription DNA-dependent
Lin19; Ole19; Len19	Glyma.19g020400	2170639	2177437	GO:0006366	transcription from RNA polymerase II promoter
Lin19; Ole19; Len19	Glyma.19g020400	2170639	2177437	GO:0007062	sister chromatid cohesion
Lin19; Ole19; Len19	Glyma.19g020400	2170639	2177437	GO:0007131	reciprocal meiotic recombination
Lin19; Ole19; Len19	Glyma.19g020400	2170639	2177437	GO:0033044	regulation of chromosome organization
Lin19; Ole19; Len19	Glyma.19g020400	2170639	2177437	GO:0042138	meiotic DNA double-strand break formation
Lin19; Ole19; Len19	Glyma.19g020400	2170639	2177437	GO:0045132	meiotic chromosome segregation
Lin19; Ole19; Len19	Glyma.19g021700	2404294	2414947	GO:0009165	nucleotide biosynthetic process
Lin19; Ole19; Len19	Glyma.19g022000	2451273	2455587	GO:0006412	translation
Pal19	Glyma.19g158700	41920067	41925098	GO:0032957	inositol trisphosphate metabolic process
Pal19	Glyma.19g158700	41920067	41925098	GO:0046854	phosphatidylinositol phosphorylation
Pal19	Glyma.19g158700	41920067	41925098	GO:0046855	inositol phosphate dephosphorylation
Pal19	Glyma.19g159400	42015735	42025479	GO:0000096	sulfur amino acid metabolic process
Pal19	Glyma.19g159400	42015735	42025479	GO:0000394	RNA splicing via endonucleolytic cleavage and ligation
Pal19	Glyma.19g159400	42015735	42025479	GO:0007267	cell-cell signaling
Pal19	Glyma.19g159400	42015735	42025479	GO:0008652	cellular amino acid biosynthetic process
Pal19	Glyma.19g159400	42015735	42025479	GO:0009069	serine family amino acid metabolic process
Pal19	Glyma.19g159400	42015735	42025479	GO:0009086	methionine biosynthetic process
Pal19	Glyma.19g159400	42015735	42025479	GO:0009616	virus induced gene silencing
Pal19	Glyma.19g159400	42015735	42025479	GO:0009855	determination of bilateral symmetry
Pal19	Glyma.19g159400	42015735	42025479	GO:0010014	meristem initiation
Pal19	Glyma.19g159400	42015735	42025479	GO:0010050	vegetative phase change
Pal19	Glyma.19g159400	42015735	42025479	GO:0010073	meristem maintenance
Pal19	Glyma.19g159400	42015735	42025479	GO:0010267	production of ta-siRNAs involved in RNA interference
Pal19	Glyma.19g159400	42015735	42025479	GO:0010364	regulation of ethylene biosynthetic process
Pal19	Glyma.19g159400	42015735	42025479	GO:0035196	production of miRNAs involved in gene silencing by miRNA
Pal19	Glyma.19g159400	42015735	42025479	GO:0042545	cell wall modification
Pal19	Glyma.19g159500	42029589	42032652	GO:0006355	regulation of transcription DNA-dependent
Pal19	Glyma.19g159500	42029589	42032652	GO:0009407	toxin catabolic process
Pal19	Glyma.19g164100	42496453	42497543	GO:0006355	regulation of transcription DNA-dependent
Pal19	Glyma.19g164100	42496453	42497543	GO:0006952	defense response
Pal19	Glyma.19g164100	42496453	42497543	GO:0009867	jasmonic acid mediated signaling pathway
Pal19	Glyma.19g164100	42496453	42497543	GO:0009873	ethylene mediated signaling pathway
Pal19	Glyma.19g164800	42559312	42561641	GO:0009737	response to abscisic acid stimulus

Table 4.11 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Pal19	Glyma.19g164800	42559312	42561641	GO:0010162	seed dormancy
Pal19	Glyma.19g164800	42559312	42561641	GO:0010431	seed maturation
Pal19	Glyma.19g164800	42559312	42561641	GO:0071215	cellular response to abscisic acid stimulus
Pal19	Glyma.19g164900	42567390	42570310	GO:0009737	response to abscisic acid stimulus
Pal19	Glyma.19g164900	42567390	42570310	GO:0010162	seed dormancy
Pal19	Glyma.19g164900	42567390	42570310	GO:0010431	seed maturation
Pal19	Glyma.19g164900	42567390	42570310	GO:0071215	cellular response to abscisic acid stimulus
Pal19	Glyma.19g165000	42574160	42577319	GO:0006886	intracellular protein transport
Pal19	Glyma.19g165000	42574160	42577319	GO:0006891	intra-Golgi vesicle-mediated transport
Pal19	Glyma.19g165100	42581234	42582938	GO:0009737	response to abscisic acid stimulus
Pal19	Glyma.19g165100	42581234	42582938	GO:0009793	embryo development ending in seed dormancy
Pal19	Glyma.19g165100	42581234	42582938	GO:0010239	chloroplast mRNA processing
Pal19	Glyma.19g165100	42581234	42582938	GO:0016226	iron-sulfur cluster assembly
Pal19	Glyma.19g165100	42581234	42582938	GO:0042793	transcription from plastid promoter
Pal19	Glyma.19g165100	42581234	42582938	GO:0045893	positive regulation of transcription DNA-dependent
Pal19	Glyma.19g165100	42581234	42582938	GO:1901259	
Pal19	Glyma.19g165200	42589885	42594875	GO:0016567	protein ubiquitination
Pal19	Glyma.19g165500	42636072	42644217	GO:0006355	regulation of transcription DNA-dependent
Pal19	Glyma.19g165500	42636072	42644217	GO:0006486	protein glycosylation
Pal19	Glyma.19g165500	42636072	42644217	GO:0009630	gravitropism
Pal19	Glyma.19g165500	42636072	42644217	GO:0010228	vegetative to reproductive phase transition of meristem
Pal19	Glyma.19g165500	42636072	42644217	GO:0010413	glucuronoxylan metabolic process
Pal19	Glyma.19g165500	42636072	42644217	GO:0016926	protein desumoylation
Pal19	Glyma.19g165500	42636072	42644217	GO:0043687	post-translational protein modification
Pal19	Glyma.19g165500	42636072	42644217	GO:0045492	xylan biosynthetic process
Pal19	Glyma.19g165500	42636072	42644217	GO:0045893	positive regulation of transcription DNA-dependent
Pal19	Glyma.19g165500	42636072	42644217	GO:0050665	hydrogen peroxide biosynthetic process
Pal19	Glyma.19g166700	42781678	42785575	GO:0007015	actin filament organization
Pal19	Glyma.19g166700	42781678	42785575	GO:0009932	cell tip growth
Pal19	Glyma.19g166700	42781678	42785575	GO:0010053	root epidermal cell differentiation
Pal19	Glyma.19g166800	42787961	42800459	GO:0006629	lipid metabolic process
Pal19	Glyma.19g166800	42787961	42800459	GO:0010150	leaf senescence
Pal19	Glyma.19g166800	42787961	42800459	GO:0016127	sterol catabolic process
Pal19	Glyma.19g166800	42787961	42800459	GO:0034434	sterol esterification
Pal19	Glyma.19g167000	42811010	42814859	GO:0006857	oligopeptide transport
Pal19	Glyma.19g167000	42811010	42814859	GO:0009611	response to wounding
Pal19	Glyma.19g167000	42811010	42814859	GO:0009737	response to abscisic acid stimulus
Pal19	Glyma.19g167000	42811010	42814859	GO:0009751	response to salicylic acid stimulus
Pal19	Glyma.19g167000	42811010	42814859	GO:0009753	response to jasmonic acid stimulus
Pal19	Glyma.19g167000	42811010	42814859	GO:0042538	hyperosmotic salinity response
Pal19	Glyma.19g167000	42811010	42814859	GO:0042742	defense response to bacterium
Pal19	Glyma.19g167000	42811010	42814859	GO:0042938	dipeptide transport
Pal19	Glyma.19g167000	42811010	42814859	GO:0042939	tripeptide transport
Pal19	Glyma.19g167000	42811010	42814859	GO:0043201	response to leucine
Pal19	Glyma.19g167000	42811010	42814859	GO:0080052	response to histidine
Pal19	Glyma.19g167000	42811010	42814859	GO:0080053	response to phenylalanine
Pal19	Glyma.19g168100	42894772	42904538	GO:0006396	RNA processing
Pal19	Glyma.19g168100	42894772	42904538	GO:0006397	mRNA processing
Pal19	Glyma.19g168100	42894772	42904538	GO:0048510	regulation of timing of transition from vegetative to reproductive phase
Pal19	Glyma.19g168200	42908207	42915046	GO:0006499	N-terminal protein myristoylation
Pal19	Glyma.19g168200	42908207	42915046	GO:0006810	transport
Pal19	Glyma.19g168200	42908207	42915046	GO:0006913	nucleocytoplasmic transport
Pal19	Glyma.19g168200	42908207	42915046	GO:0009220	pyrimidine ribonucleotide biosynthetic process
Pal19	Glyma.19g169300	43008455	43010658	GO:0000165	MAPKKK cascade
Pal19	Glyma.19g169300	43008455	43010658	GO:0006612	protein targeting to membrane
Pal19	Glyma.19g169300	43008455	43010658	GO:0006855	drug transmembrane transport
Pal19	Glyma.19g169300	43008455	43010658	GO:0007165	signal transduction
Pal19	Glyma.19g169300	43008455	43010658	GO:0009414	response to water deprivation
Pal19	Glyma.19g169300	43008455	43010658	GO:0009595	detection of biotic stimulus
Pal19	Glyma.19g169300	43008455	43010658	GO:0009697	salicylic acid biosynthetic process
Pal19	Glyma.19g169300	43008455	43010658	GO:0009723	response to ethylene stimulus
Pal19	Glyma.19g169300	43008455	43010658	GO:0009733	response to auxin stimulus
Pal19	Glyma.19g169300	43008455	43010658	GO:0009737	response to abscisic acid stimulus
Pal19	Glyma.19g169300	43008455	43010658	GO:0009738	abscisic acid mediated signaling pathway
Pal19	Glyma.19g169300	43008455	43010658	GO:0009751	response to salicylic acid stimulus
Pal19	Glyma.19g169300	43008455	43010658	GO:0009753	response to jasmonic acid stimulus
Pal19	Glyma.19g169300	43008455	43010658	GO:0009862	systemic acquired resistance salicylic acid mediated signaling pathway
Pal19	Glyma.19g169300	43008455	43010658	GO:0009867	jasmonic acid mediated signaling pathway
Pal19	Glyma.19g169300	43008455	43010658	GO:0010193	response to ozone
Pal19	Glyma.19g169300	43008455	43010658	GO:0010200	response to chitin
Pal19	Glyma.19g169300	43008455	43010658	GO:0010310	regulation of hydrogen peroxide metabolic process
Pal19	Glyma.19g169300	43008455	43010658	GO:0010363	regulation of plant-type hypersensitive response
Pal19	Glyma.19g169300	43008455	43010658	GO:0015692	lead ion transport
Pal19	Glyma.19g169300	43008455	43010658	GO:0031348	negative regulation of defense response
Pal19	Glyma.19g169300	43008455	43010658	GO:0042538	hyperosmotic salinity response
Pal19	Glyma.19g169300	43008455	43010658	GO:0042742	defense response to bacterium
Pal19	Glyma.19g169300	43008455	43010658	GO:0043069	negative regulation of programmed cell death
Pal19	Glyma.19g169300	43008455	43010658	GO:0043900	regulation of multi-organism process
Pal19	Glyma.19g169300	43008455	43010658	GO:0046865	terpenoid transport
Pal19	Glyma.19g169300	43008455	43010658	GO:0050832	defense response to fungus
Pal19	Glyma.19g169300	43008455	43010658	GO:0080168	abscisic acid transport
Pal19	Glyma.19g169400	43013724	43021681	GO:0000165	MAPKKK cascade

Table 4.11 Continued.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Pal19	Glyma.19g169400	43013724	43021681	GO:0006612	protein targeting to membrane
Pal19	Glyma.19g169400	43013724	43021681	GO:0006855	drug transmembrane transport
Pal19	Glyma.19g169400	43013724	43021681	GO:0007165	signal transduction
Pal19	Glyma.19g169400	43013724	43021681	GO:0009414	response to water deprivation
Pal19	Glyma.19g169400	43013724	43021681	GO:0009595	detection of biotic stimulus
Pal19	Glyma.19g169400	43013724	43021681	GO:0009697	salicylic acid biosynthetic process
Pal19	Glyma.19g169400	43013724	43021681	GO:0009723	response to ethylene stimulus
Pal19	Glyma.19g169400	43013724	43021681	GO:0009733	response to auxin stimulus
Pal19	Glyma.19g169400	43013724	43021681	GO:0009737	response to abscisic acid stimulus
Pal19	Glyma.19g169400	43013724	43021681	GO:0009738	abscisic acid mediated signaling pathway
Pal19	Glyma.19g169400	43013724	43021681	GO:0009751	response to salicylic acid stimulus
Pal19	Glyma.19g169400	43013724	43021681	GO:0009753	response to jasmonic acid stimulus
Pal19	Glyma.19g169400	43013724	43021681	GO:0009862	systemic acquired resistance salicylic acid mediated signaling pathway
Pal19	Glyma.19g169400	43013724	43021681	GO:0009867	jasmonic acid mediated signaling pathway
Pal19	Glyma.19g169400	43013724	43021681	GO:0010193	response to ozone
Pal19	Glyma.19g169400	43013724	43021681	GO:0010200	response to chitin
Pal19	Glyma.19g169400	43013724	43021681	GO:0010310	regulation of hydrogen peroxide metabolic process
Pal19	Glyma.19g169400	43013724	43021681	GO:0010363	regulation of plant-type hypersensitive response
Pal19	Glyma.19g169400	43013724	43021681	GO:0015692	lead ion transport
Pal19	Glyma.19g169400	43013724	43021681	GO:0031348	negative regulation of defense response
Pal19	Glyma.19g169400	43013724	43021681	GO:0042538	hyperosmotic salinity response
Pal19	Glyma.19g169400	43013724	43021681	GO:0042742	defense response to bacterium
Pal19	Glyma.19g169400	43013724	43021681	GO:0043069	negative regulation of programmed cell death
Pal19	Glyma.19g169400	43013724	43021681	GO:0043900	regulation of multi-organism process
Pal19	Glyma.19g169400	43013724	43021681	GO:0046865	terpenoid transport
Pal19	Glyma.19g169400	43013724	43021681	GO:0050832	defense response to fungus
Pal19	Glyma.19g169400	43013724	43021681	GO:0080168	abscisic acid transport
Pal19	Glyma.19g170900	43217844	43223544	GO:0008150	biological_process
Pal19	Glyma.19g171000	43229301	43230858	GO:0006629	lipid metabolic process
Pal19	Glyma.19g171000	43229301	43230858	GO:0010103	stomatal complex morphogenesis
Pal19	Glyma.19g171000	43229301	43230858	GO:0016556	mRNA modification
Pal19	Glyma.19g172100	43314273	43316701	GO:0009408	response to heat
Pal19	Glyma.19g172100	43314273	43316701	GO:0009644	response to high light intensity
Pal19	Glyma.19g172100	43314273	43316701	GO:0042538	hyperosmotic salinity response
Pal19	Glyma.19g172100	43314273	43316701	GO:0042542	response to hydrogen peroxide
Pal19	Glyma.19g172300	43325742	43328166	GO:0006810	transport
Pal19	Glyma.19g172600	43333204	43337622	GO:0006623	protein targeting to vacuole
Pal19	Glyma.19g172600	43333204	43337622	GO:0006812	cation transport
Pal19	Glyma.19g172600	43333204	43337622	GO:0006885	regulation of pH
Pal19	Glyma.19g172600	43333204	43337622	GO:0030007	cellular potassium ion homeostasis
Pal19	Glyma.19g172600	43333204	43337622	GO:0030104	water homeostasis
Pal19	Glyma.19g172600	43333204	43337622	GO:0035725	sodium ion transmembrane transport
Pal19	Glyma.19g172600	43333204	43337622	GO:0055085	transmembrane transport
Pal19	Glyma.19g175700	43569883	43572105	GO:0002237	response to molecule of bacterial origin
Pal19	Glyma.19g175700	43569883	43572105	GO:0010015	root morphogenesis
Pal19	Glyma.19g175900	43587364	43593960	GO:0006754	ATP biosynthetic process
Pal19	Glyma.19g175900	43587364	43593960	GO:0006812	cation transport
Pal19	Glyma.19g175900	43587364	43593960	GO:0006816	calcium ion transport
Pal19	Glyma.19g175900	43587364	43593960	GO:0008152	metabolic process
Pal19	Glyma.19g176100	43606806	43609648	GO:0000398	nuclear mRNA splicing via spliceosome
Pal19	Glyma.19g176100	43606806	43609648	GO:0006355	regulation of transcription DNA-dependent
Pal19	Glyma.19g176100	43606806	43609648	GO:0008380	RNA splicing
Pal19	Glyma.19g176100	43606806	43609648	GO:0030422	production of siRNA involved in RNA interference
Pal19	Glyma.19g176100	43606806	43609648	GO:0035196	production of miRNAs involved in gene silencing by miRNA
Pal19	Glyma.19g176100	43606806	43609648	GO:0043687	post-translational protein modification
Pal19	Glyma.19g176100	43606806	43609648	GO:0045893	positive regulation of transcription DNA-dependent
Pal19	Glyma.19g176500	43635433	43637342	GO:0008150	biological_process
Pal19	Glyma.19g177000	43654763	43656193	GO:0010731	protein glutathionylation
Pal19	Glyma.19g177700	43697819	43705695	GO:0008150	biological_process
Pal19	Glyma.19g177800	43708336	43711024	GO:0008150	biological_process
Pal19	Glyma.19g178000	43718461	43721894	GO:0006355	regulation of transcription DNA-dependent
Pal19	Glyma.19g178000	43718461	43721894	GO:0007623	circadian rhythm
Pal19	Glyma.19g178000	43718461	43721894	GO:0009651	response to salt stress
Pal19	Glyma.19g178000	43718461	43721894	GO:0009723	response to ethylene stimulus
Pal19	Glyma.19g178000	43718461	43721894	GO:0009733	response to auxin stimulus
Pal19	Glyma.19g178000	43718461	43721894	GO:0009737	response to abscisic acid stimulus
Pal19	Glyma.19g178000	43718461	43721894	GO:0009739	response to gibberellin stimulus
Pal19	Glyma.19g178000	43718461	43721894	GO:0009751	response to salicylic acid stimulus
Pal19	Glyma.19g178000	43718461	43721894	GO:0009753	response to jasmonic acid stimulus
Pal19	Glyma.19g178000	43718461	43721894	GO:0032922	circadian regulation of gene expression
Pal19	Glyma.19g178000	43718461	43721894	GO:0043966	histone H3 acetylation
Pal19	Glyma.19g178000	43718461	43721894	GO:0046686	response to cadmium ion
Pal19	Glyma.19g178000	43718461	43721894	GO:0048573	photoperiodism flowering
Pal19	Glyma.19g178000	43718461	43721894	GO:0048574	long-day photoperiodism flowering
Pal19	Glyma.19g179100	43824144	43826915	GO:0009832	plant-type cell wall biogenesis
Pal19	Glyma.19g179100	43824144	43826915	GO:0009834	secondary cell wall biogenesis
Pal19	Glyma.19g179100	43824144	43826915	GO:0010413	glucuronoxylan metabolic process
Pal19	Glyma.19g179100	43824144	43826915	GO:0010417	glucuronoxylan biosynthetic process
Pal19	Glyma.19g179100	43824144	43826915	GO:0042546	cell wall biogenesis
Pal19	Glyma.19g179100	43824144	43826915	GO:0044036	cell wall macromolecule metabolic process
Pal19	Glyma.19g179100	43824144	43826915	GO:0045492	xylan biosynthetic process
Pal19	Glyma.19g179300	43833754	43838961	GO:0009220	pyrimidine ribonucleotide biosynthetic process

Table 4.12 Gene models and gene ontology (GO) IDs and descriptions for biological processes located within +/- 5 mb from stearic acid QTL on soybean chromosome 14 based on Glyma 2.0 position.

QTL	Gene Model	Start	Stop	Annotation ID	Annotation Description
Ste14	Glyma.14g118700	16245387	16247738	GO:0006468	protein phosphorylation
Ste14	Glyma.14g118800	16247988	16257186	GO:0006468	protein phosphorylation
Ste14	Glyma.14g119700	16750343	16751203	GO:0006662	glycerol ether metabolic process
Ste14	Glyma.14g119700	16750343	16751203	GO:0006979	response to oxidative stress
Ste14	Glyma.14g119700	16750343	16751203	GO:0043085	positive regulation of catalytic activity
Ste14	Glyma.14g119700	16750343	16751203	GO:0045454	cell redox homeostasis
Ste14	Glyma.14g119800	16751592	16763348	GO:0007275	multicellular organismal development
Ste14	Glyma.14g119800	16751592	16763348	GO:0009790	embryo development
Ste14	Glyma.14g119800	16751592	16763348	GO:0009909	regulation of flower development
Ste14	Glyma.14g119800	16751592	16763348	GO:0042744	hydrogen peroxide catabolic process
Ste14	Glyma.14g119800	16751592	16763348	GO:0048467	gynoecium development
Ste14	Glyma.14g119800	16751592	16763348	GO:0048481	ovule development
Ste14	Glyma.14g120000	16912795	16915379	GO:0008150	biological_process
Ste14	Glyma.14g120000	16912795	16915379	GO:0009827	plant-type cell wall modification
Ste14	Glyma.14g120000	16912795	16915379	GO:0009860	pollen tube growth
Ste14	Glyma.14g120000	16912795	16915379	GO:0048610	cellular process involved in reproduction
Ste14	Glyma.14g120000	16912795	16915379	GO:0048868	pollen tube development
Ste14	Glyma.14g120200	16972083	16974184	GO:0006629	lipid metabolic process
Ste14	Glyma.14g120400	17050528	17052976	GO:0000256	allantoin catabolic process
Ste14	Glyma.14g120400	17050528	17052976	GO:0000394	RNA splicing via endonucleolytic cleavage and ligation
Ste14	Glyma.14g120400	17050528	17052976	GO:0006508	proteolysis
Ste14	Glyma.14g120400	17050528	17052976	GO:0008152	metabolic process
Ste14	Glyma.14g120400	17050528	17052976	GO:0009086	methionine biosynthetic process
Ste14	Glyma.14g120700	17134656	17136748	GO:0008150	biological_process
Ste14	Glyma.14g121000	17429115	17432802	GO:0006614	SRP-dependent cotranslational protein targeting to membrane
Ste14	Glyma.14g121000	17429115	17432802	GO:0006617	SRP-dependent cotranslational protein targeting to membrane signal sequence recognition
Ste14	Glyma.14g121000	17429115	17432802	GO:0010267	production of ta-siRNAs involved in RNA interference
Ste14	Glyma.14g121000	17429115	17432802	GO:0019288	isopentenyl diphosphate biosynthetic process mevalonate-independent pathway
Ste14	Glyma.14g121000	17429115	17432802	GO:0035196	production of miRNAs involved in gene silencing by miRNA
Ste14	Glyma.14g121000	17429115	17432802	GO:0045038	protein import into chloroplast thylakoid membrane
Ste14	Glyma.14g121000	17429115	17432802	GO:0051607	defense response to virus
Ste14	Glyma.14g121100	17454361	17456871	GO:0008150	biological_process
Ste14	Glyma.14g121500	17529809	17534637	GO:0008150	biological_process
Ste14	Glyma.14g121600	17544093	17547984	GO:0006281	DNA repair
Ste14	Glyma.14g121600	17544093	17547984	GO:0010100	negative regulation of photomorphogenesis
Ste14	Glyma.14g121600	17544093	17547984	GO:0048608	reproductive structure development
Ste14	Glyma.14g121700	17550558	17554820	GO:0007129	synapsis
Ste14	Glyma.14g121700	17550558	17554820	GO:0007131	reciprocal meiotic recombination
Ste14	Glyma.14g124400	19061201	19065418	GO:0006468	protein phosphorylation
Ste14	Glyma.14g124400	19061201	19065418	GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway
Ste14	Glyma.14g124400	19061201	19065418	GO:0010075	regulation of meristem growth
Ste14	Glyma.14g125000	19605857	19606066	GO:0008150	biological_process
Ste14	Glyma.14g128200	21029187	21032368	GO:0009851	auxin biosynthetic process
Ste14	Glyma.14g128200	21029187	21032368	GO:0009911	positive regulation of flower development
Ste14	Glyma.14g128200	21029187	21032368	GO:0010229	inflorescence development
Ste14	Glyma.14g128200	21029187	21032368	GO:0022603	regulation of anatomical structure morphogenesis
Ste14	Glyma.14g128200	21029187	21032368	GO:0048825	cotyledon development
Ste14	Glyma.14g128200	21029187	21032368	GO:0048827	phyllome development
Ste14	Glyma.14g128200	21029187	21032368	GO:0055114	oxidation-reduction process
Ste14	Glyma.14g128200	21029187	21032368	GO:2000024	regulation of leaf development
Ste14	Glyma.14g129000	21674410	21678350	GO:0008150	biological_process
Ste14	Glyma.14g131900	22580087	22586928	GO:0008150	biological_process
Ste14	Glyma.14g133400	23111745	23115790	GO:0000278	mitotic cell cycle
Ste14	Glyma.14g133400	23111745	23115790	GO:0000724	double-strand break repair via homologous recombination
Ste14	Glyma.14g133400	23111745	23115790	GO:0000911	cytokinesis by cell plate formation
Ste14	Glyma.14g133400	23111745	23115790	GO:0006260	DNA replication
Ste14	Glyma.14g133400	23111745	23115790	GO:0006261	DNA-dependent DNA replication
Ste14	Glyma.14g133400	23111745	23115790	GO:0006270	DNA-dependent DNA replication initiation
Ste14	Glyma.14g133400	23111745	23115790	GO:0006275	regulation of DNA replication
Ste14	Glyma.14g133400	23111745	23115790	GO:0006281	DNA repair
Ste14	Glyma.14g133400	23111745	23115790	GO:0006306	DNA methylation
Ste14	Glyma.14g133400	23111745	23115790	GO:0008283	cell proliferation
Ste14	Glyma.14g133400	23111745	23115790	GO:0009555	pollen development
Ste14	Glyma.14g133400	23111745	23115790	GO:0009909	regulation of flower development
Ste14	Glyma.14g133400	23111745	23115790	GO:0010389	regulation of G2/M transition of mitotic cell cycle
Ste14	Glyma.14g133400	23111745	23115790	GO:0016444	somatic cell DNA recombination
Ste14	Glyma.14g133400	23111745	23115790	GO:0016458	gene silencing
Ste14	Glyma.14g133400	23111745	23115790	GO:0016568	chromatin modification
Ste14	Glyma.14g133400	23111745	23115790	GO:0031047	gene silencing by RNA
Ste14	Glyma.14g133400	23111745	23115790	GO:0034968	histone lysine methylation
Ste14	Glyma.14g133400	23111745	23115790	GO:0051567	histone H3-K9 methylation
Ste14	Glyma.14g133400	23111745	23115790	GO:0051726	regulation of cell cycle

Figures

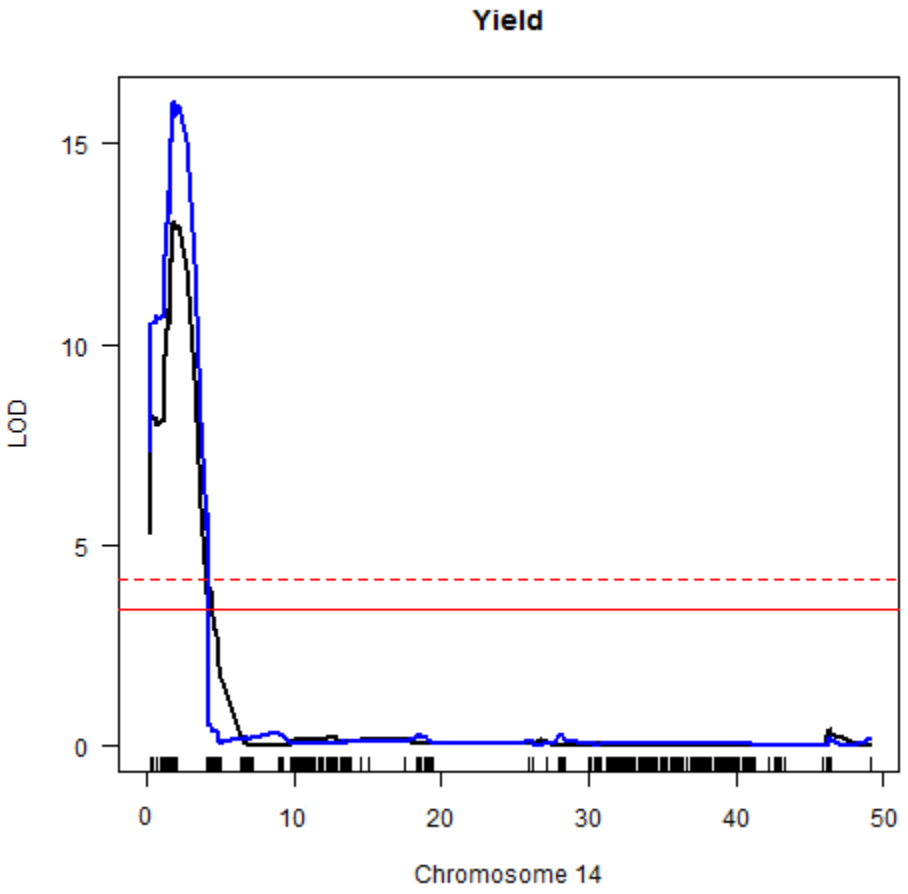

Figure 4.1 Quantitative trait loci (QTL) for yield detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).

Figure 4.2 Quantitative trait loci (QTL) for palmitic acid detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).

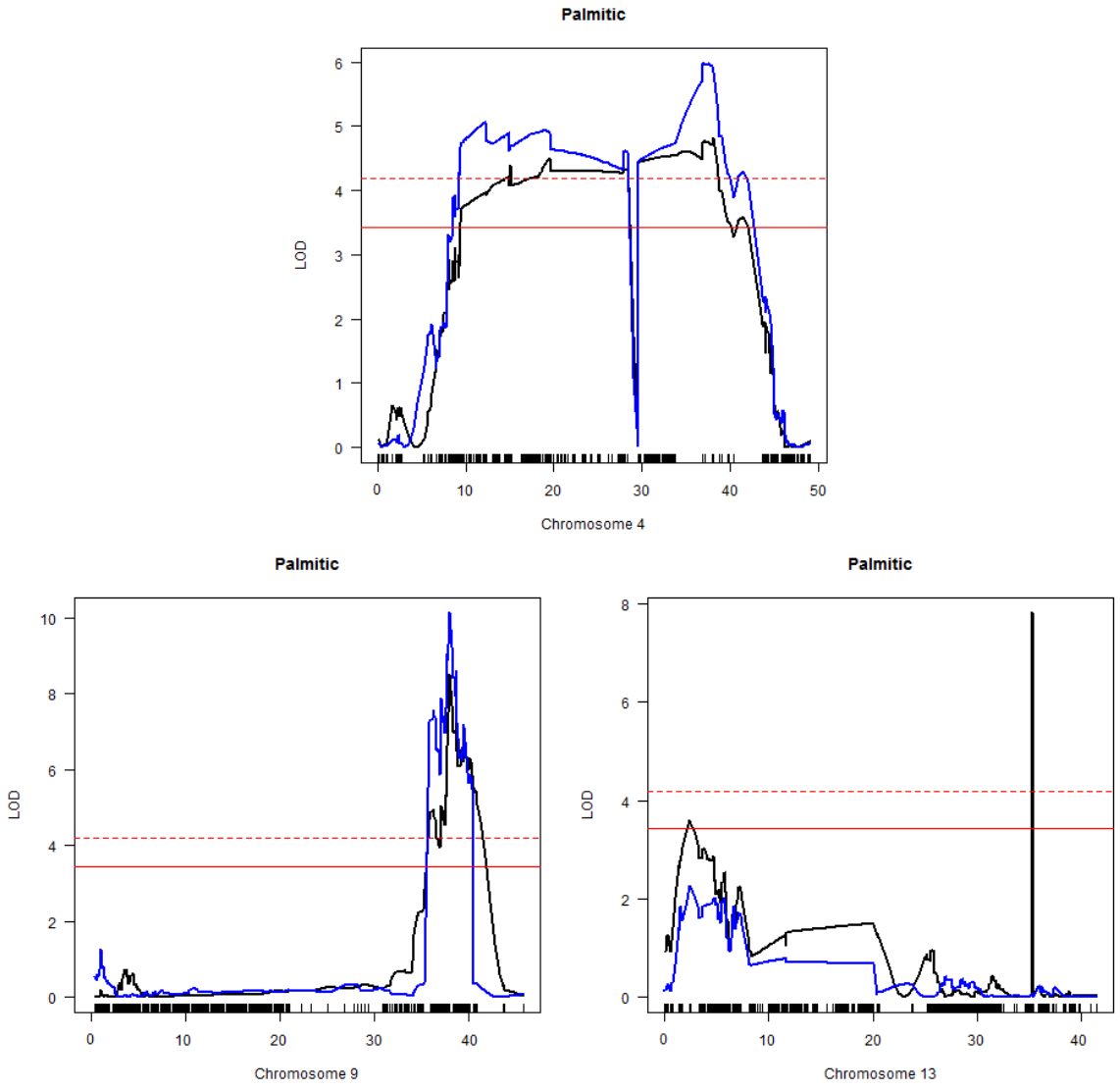

Figure 4.2 Continued.

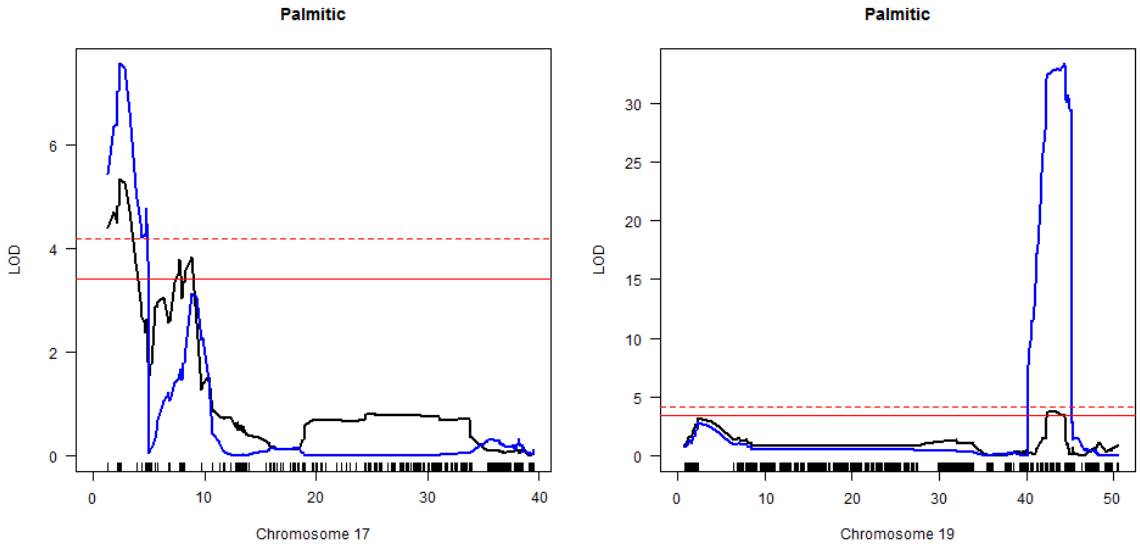

Figure 4.2 Continued.

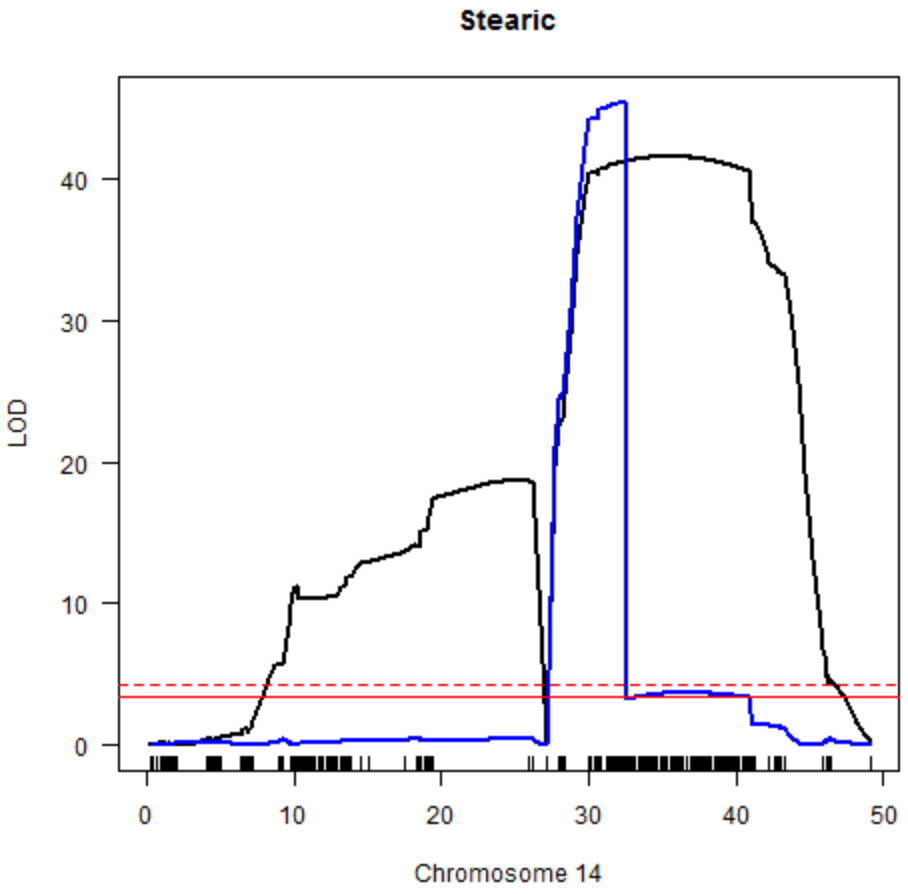

Figure 4.3 Quantitative trait loci (QTL) for stearic acid detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).

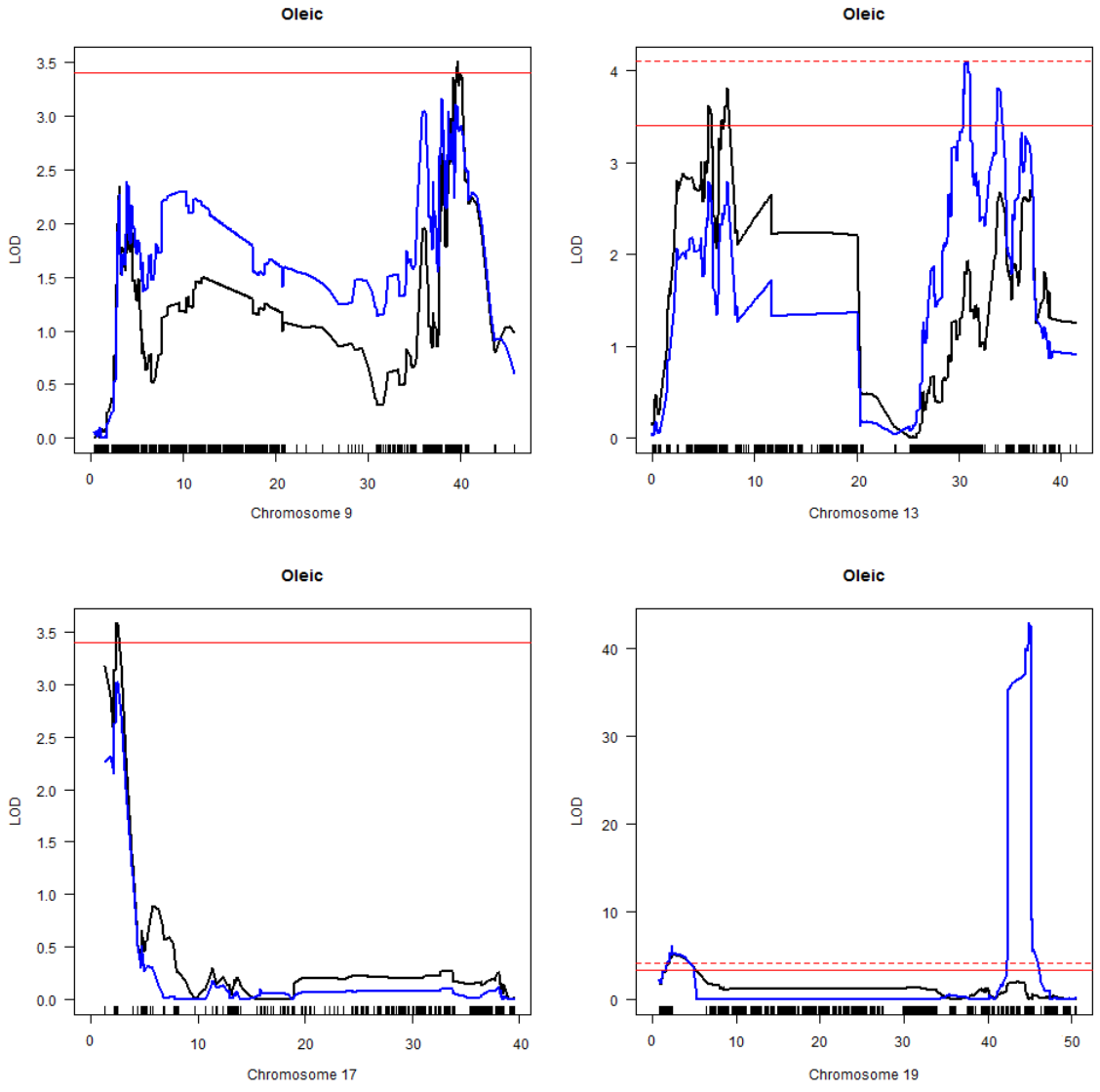

Figure 4.4 Quantitative trait loci (QTL) for oleic acid detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).

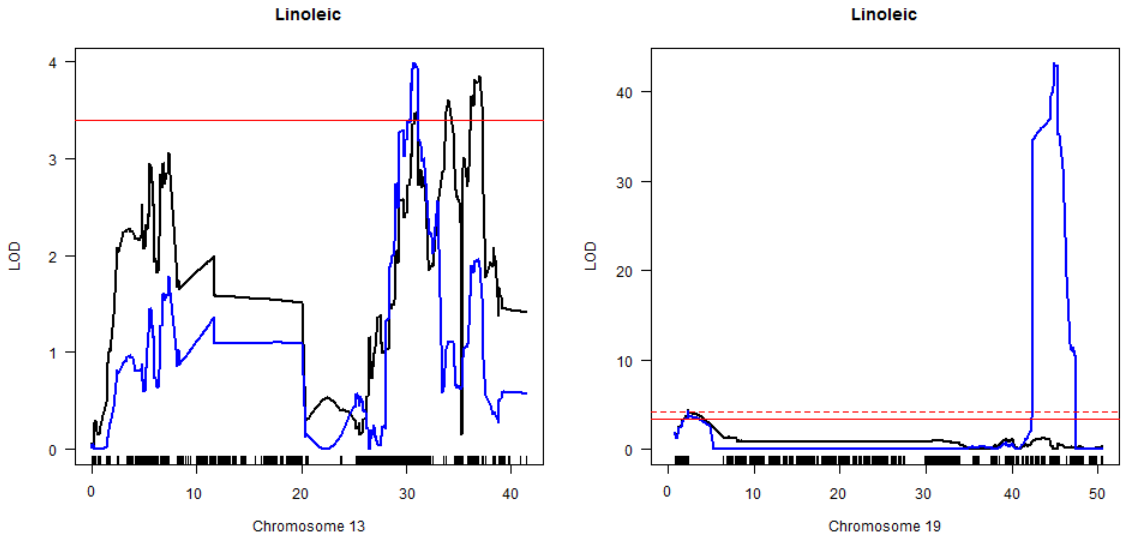

Figure 4.5 Quantitative trait loci (QTL) for linoleic acid detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).

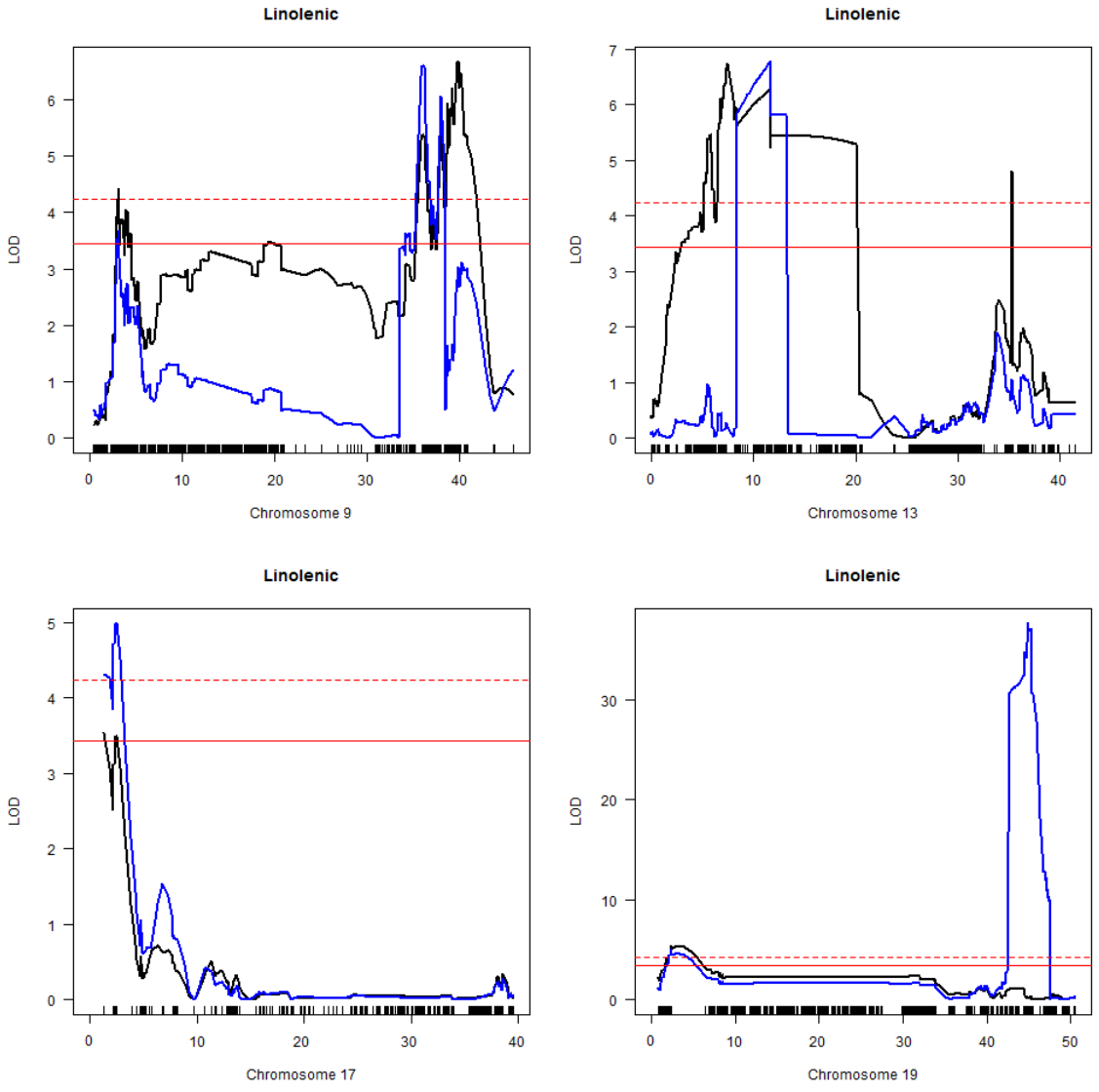

Figure 4.6 Quantitative trait loci (QTL) for linolenic acid detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).

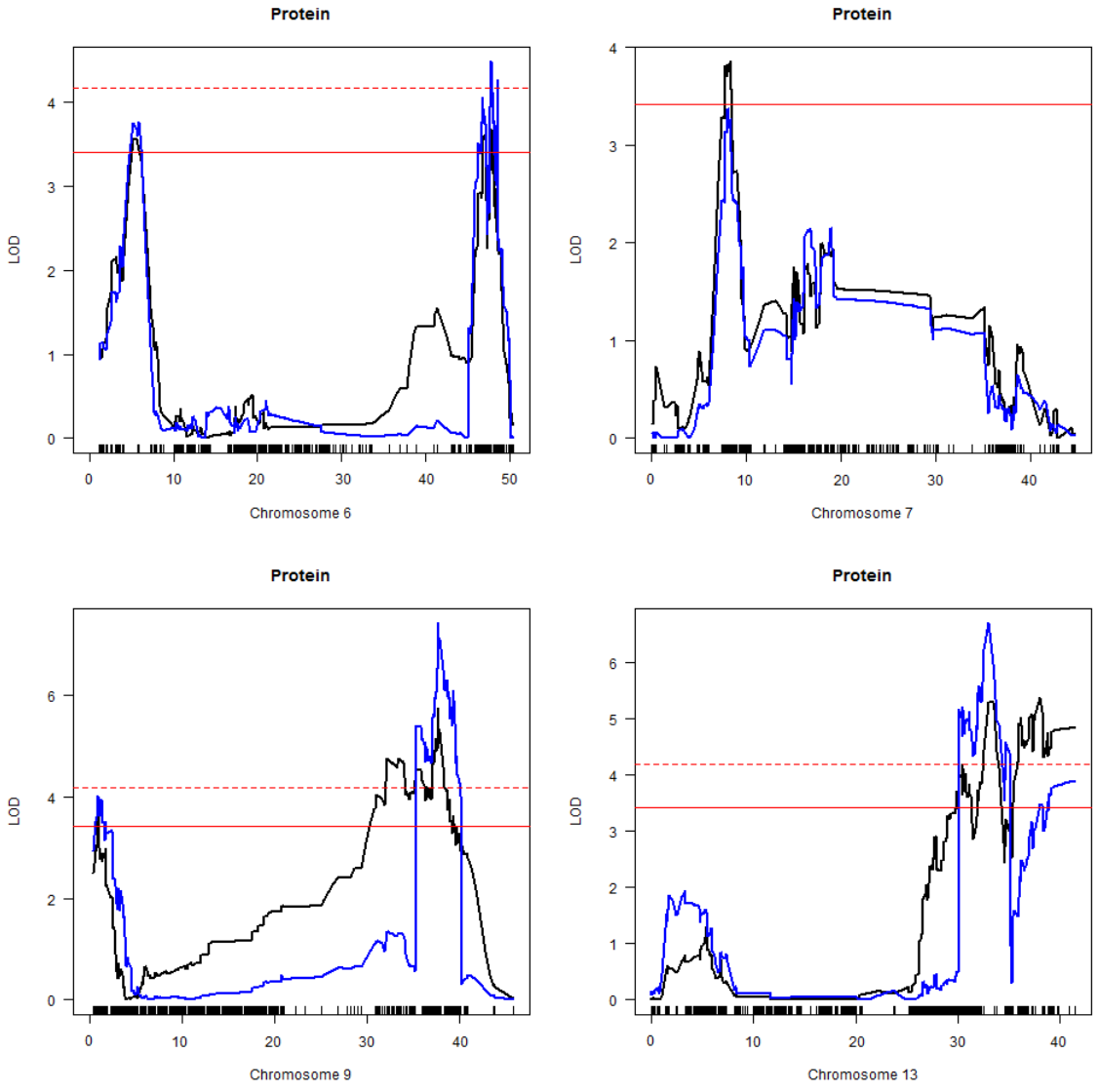

Figure 4.7 Quantitative trait loci (QTL) for protein detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).

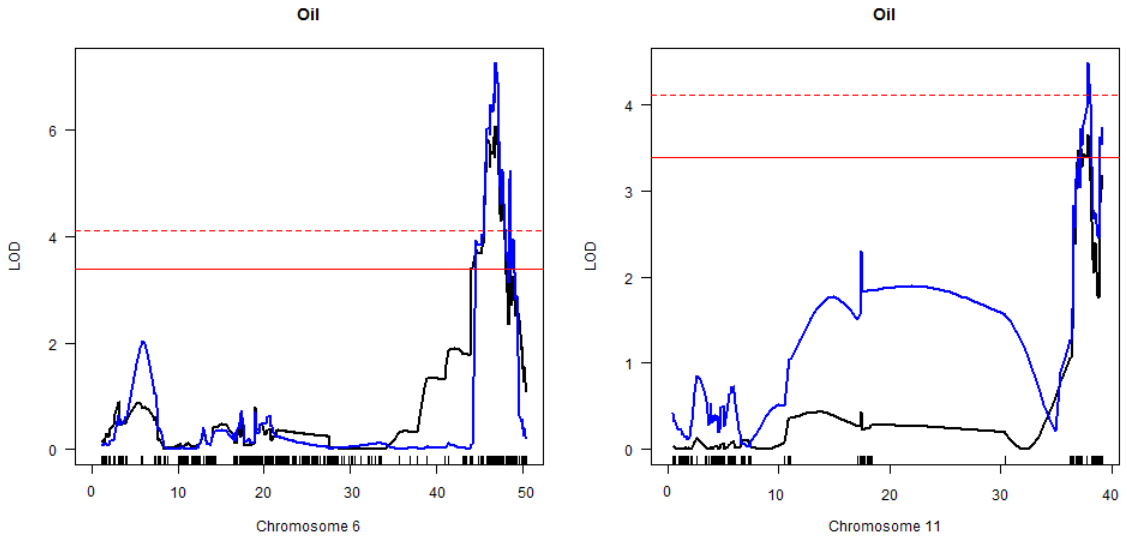

Figure 4.8 Quantitative trait loci (QTL) for oil detected in soybean population E×W-50K subset consisting of 203 F5 derived RILs using 16,718 SNPs. Phenotypic data was taken from the combined analysis using data from 2010, 2013, and 2014 field seasons. Chromosomal positions on the X-axis are in mbs from Glyma 1.01 (Schmutz et al., 2010). The solid and dashed red horizontal lines indicate the 5% and 1% LOD thresholds, respectively. The black QTL trace is from the interval mapping procedure, which included the Dt1 locus as a covariate. The blue QTL trace in from the composite interval mapping procedure, serving as a protection against 'ghost' QTL (Martinez and Curnow, 1992).

CHAPTER 5

CONCLUSIONS

Soybean [Glycine max (L.) Merrill] is the leading oilseed crop grown in the world (Sharma et al., 2012). Seed protein (~400 g kg^{-1}) and oil (~200 g kg^{-1}) are primary components of soybean. As high value traits, protein and oil are common targets for research seeking to improve the value of soybean. Additionally, with the Food and Drug Administration (FDA) recently banning partially hydrogenated oils (PHOs) in all food products (https://www.federalregister.gov/articles/2015/06/17/2015-14883/final-determination-regarding-partially-hydrogenated-oils), improving the fatty acid profile of soybean has become a major objective. For these reasons, the purpose of this research was to evaluate molecular breeding strategies for improving soybean yield, fatty acids, protein, and oil.

The first evaluation was a comparison of molecular and phenotypic breeding strategies from progeny row selections. From this it was determined that molecular strategies consistently outperformed phenotypic selections (PS) in the progeny row stage for soybean yield, fatty acids, protein, and oil. For yield, Epistacy (Holland, 1998) was the preferred selection method. For fatty acids, protein, and oil, the genomic selection (GS) (Meuwissen et al., 2001) strategies were preferred.

The second evaluation was very similar to the first, except that the trait selections were derived from replicated field trials. With the exception of yield, for which PS was the dominant method, the results were rather mixed. Overall, PS and GS methods were very comparable for fatty acids, protein, and oil; indicating that either of these methods could be useful for making improvements.

Finally, the third evaluation of molecular breeding strategies involved identifying and exploring significant genomic regions for soybean yield, fatty acids, protein, and oil. For these traits a total of 29 quantitative trait loci (QTL) were detected. Of these QTLs, three were excellent candidates for confirmed status

(http://www.soybase.org/resources/QTL.php, accessed 7/26/2015) and four were strong candidates for positional confirmations (Smallwood et al., 2014). Additionally, while more work is needed for validation, candidate genes with possible influence on soybean yield, fatty acids, protein and oil were identified. Further, pleiotropic effects between protein and oil and between the fatty acids were identified in this study.

The results from this research should be beneficial for those seeking to improve soybean yield, fatty acids, protein, and oil. Researchers looking to make selections from both progeny rows and replicated field trials can draw from these results when choosing which selection strategy to use. Additionally, the gained knowledge with regard to influential genomic regions for these traits can be applied to future efforts seeking improvement. More research seeking to implement high performing molecular breeding strategies and to identify causative genes for these and other QTLs impacting targeted traits will be important for the soybean breeding community in the future.

References

Federal Register. 2015. Final determination regarding partially hydrogenated oils. https://www.federalregister.gov/articles/2015/06/17/2015-14883/final-determination-regarding-partially-hydrogenated-oils (accessed 24 July 2015).

Holland, J.B. 1998. EPISTACY: A SAS program for detecting two-locus epistatic interactions using genetic marker information. J. Hered. 89:374–375.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.

Sharma, M., S.K. Gupta, and A.K. Mondal. 2012. Production and trade of major world oil crops. In: S. K. Gupta, editor, Technological innovations in major world oil crops, volume 1. Springer Science + Business Media LLC. p. 1-15.

Smallwood, C.J., C.N. Nyinyi, D.A. Kopsell, C.E. Sams, D.R. West, P. Chen, S.K. Kantartzi, P.B. Cregan, D.L. Hyten, and V.R. Pantalone. 2014. Detection and confirmation of quantitative trait loci for soybean seed isoflavones. Crop Sci. 54:1-12.

SoyBase and the Soybean Breeder’s Toolbox. 2007. QTL nomenclature. http://www.soybase.org/resources/QTL.php (accessed 26 July 2015).

VITA

Christopher Joseph Smallwood was born in Laurel, MD in 1984. After completing an Associate of Arts degree in General Studies from Anne Arundel Community College in 2004, he was drawn southward to the Great Smoky Mountains to continue his studies. In the spring of 2008 he received a Bachelor of Science degree in Wildlife & Fisheries Science, with a minor in Forestry from the University of Tennessee, Knoxville.

After completing his B.S. degree, Chris returned to Maryland where he worked primarily as a driver/guard in the armored car industry in Washington D.C. In the summer of 2010, he started a Master of Science program at the University of Tennessee, Knoxville in Plant Sciences, with a concentration in plant breeding. The title of his thesis was “Detection of Quantitative Trait Loci for Marker-Assisted Selection of Soybean Isoflavone Genistein”, which he completed in 2012. Upon completion of his M.S. degree, Chris transitioned immediately into his Ph.D. dissertation project.

While completing his dissertation, Chris has also served as the field manager for the University of Tennessee Soybean Breeding Program since 2013. The combined responsibilities of being a graduate student and a research associate helped Chris gain an improved understanding of both theoretical and practical components in plant breeding. After completing his dissertation, Chris intends to continue with plant improvement as a breeder in an applied program.