2008

Genetic diversity in cultivated and wild Hordeum species

Lucía Gutiérrez
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/rtd
Part of the Agricultural Science Commons, Agronomy and Crop Sciences Commons, Bioinformatics Commons, and the Genetics and Genomics Commons

Recommended Citation
Genetic diversity in cultivated and wild *Hordeum* species

by

Lucía Gutiérrez

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Co-Majors: Plant Breeding; Ecology and Evolutionary Biology

Program of Study Committee:
Jean-Luc Jannink, Co-major Professor
John D. Nason, Co-major Professor
Philip Dixon
Kendall R. Lamkey
Dean C. Adams

Iowa State University
Ames, Iowa
2008

Copyright © Lucía Gutiérrez, 2008. All rights reserved
To Valentín,

Santi and Mati,

and my parents.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS v

RESUMEN - ABSTRACT IN SPANISH xi

ABSTRACT xii

CHAPTER I. INTRODUCTION 1
  Plant Genetic Diversity 1
  Measurements of Population Structure 2
  Overall versus Pairwise Comparisons 4
  *Hordeum* genus 5
  Dissertation Objectives 6
  Dissertation Organization 7
  References 8

CHAPTER II. A NEW METHOD OF $Q_{ST}$-$F_{ST}$ CONTRASTS IDENTIFIES SELECTION MOSAICS IN WHICH SUMMARY $Q_{ST}$-$F_{ST}$ COMPARISONS FAIL TO REJECT NEUTRALITY 16
  Abstract 16
  Introduction 17
  Materials and Methods 19
  Results and Discussion 23
  Literature Cited 27
  List of Figures 28
  Table 1. $Q_{ST}$, $F_{ST}$, and $\Delta (Q_{ST} - F_{ST})$ estimates 30
  Figure 1. Example of non-neutral and neutral cases with overall $Q_{ST} = F_{ST}$ 31
  Figure 2. Observed and expected pair-wise $Q_{ST}^u$ and $F_{ST}^u$ in simulated data 32
  Figure 3. Observed and expected pair-wise $Q_{ST}^u$ and $F_{ST}^u$ in wild barley 33

CHAPTER III: OVERALL AND PAIR-WISE COMPARISONS OF QUANTITATIVE AND GENETIC POPULATION STRUCTURE IN WILD BARLEY (*Hordeum spontaneum* K. Koch) 34
  Abstract 34
  Introduction 35
  Materials and Methods 38
  Results 46
  Discussion 48
  Acknowledgments 54
  References 54
  Table 1. Description of populations sampled 61
  Table 2. Overall and pair-wise estimates of $Q_{ST}$ and $F_{ST}$ 62
CHAPTER IV. MORPHOLOGICAL GENETIC DIVERSITY OF WORLDWIDE BARLEY AND MEGA-TARGETS OF SELECTION

Abstract
Introduction
Materials and Methods
Results
Discussion
Acknowledgements
References
List of Figures

Table 1. Mean and diversity of breeding program for plot level traits
Table 2. Mean and diversity of breeding program for plant level traits
Table 3. Discriminant traits among two- and six-row and breeding programs
Table 4. Ten best genotypes for some variables of breeding interest
Figure 1. Principal component analysis of means of breeding programs
Figure 2. Cluster dendogram of performance of breeding programs
Figure 3. Cluster dendogram of difference of means and grain yield difference

CHAPTER V: EVOLUTIONARY HISTORY OF AGRONOMICALLY RELEVANT TRAITS IN BARLEY

Abstract
Introduction
Materials and Methods
Results
Discussion
References
List of Figures

Table 1. Overall and pair-wise comparison of $Q_{ST}$ and $F_{ST}$
Figure 1. Pair-wise $Q_{STij}$-$F_{STij}$ for non-neutral traits
Figure 2. Pair-wise $QSTij$-$FSTij$ for divergent traits
Figure 3. Classification of traits for study of domestication

CHAPTER VI: GENERAL CONCLUSIONS

Final thought
ACKNOWLEDGEMENTS

I am extremely grateful to my advisors, Jean-Luc Jannink and John Nason, for their guidance and support. Jean-Luc has been the most supportive advisor I will ever have. He was available even if it meant having to leave everything else to meet my coming deadlines. He would come down from his higher level thoughts to explain things in plain English. He was very understanding and would always listen first. John provided a new perspective into my work, always thinking of creative ideas to pursue, and being very careful in the written communication of our thoughts. I would also like to thank my other POS Committee members, Phillip Dixon, Kendall Lamkey, and Dean Adams, for their support, advice, and insight in various aspects of the research.

Financial support for this research project came from the people of the State of Iowa through their taxes, The Land Institute Natural Systems Agriculture Graduate Fellowship, the Raymond Baker Center for Plant Breeding at Iowa State University, and in-kind support from the Monsanto Company and INIA-Uruguay.

I had the luck of receiving help from many technicians who made my life so much easier. George Patrick and Ron Skrdla from Jannink’s group at ISU were very supportive throughout the research; they helped me design the field work, find better field techniques, deal with the growing habits of the different species, get the greenhouse work done, harvest the wild species, and so many other things, but foremost, they were always very understanding of my rushing needs. Valerie Buffard from the Molecular Breeding group at Monsanto taught me how to run molecular markers, helped me with the zillions of them I ran, and was always available to answer my questions and to find solutions. Coco and Polaco from the Barley Breeding Group at INIA were always available for helping me with the non-stop concerns I had with my field plots in Uruguay.

I am extremely thankful to all the people at INIA La Estanzuela in Uruguay who always answered the door when I kept coming with questions or requesting their assistance in yet another thing. And when I say ‘all’, I really mean it; everyone in La Estanzuela helped with this project. Thanks to Eduardo De La Rosa, the Director, for opening INIA doors to my research. Especially, I received tremendous field support from
Juan E. Diaz and the Barley Breeding Group, Silvia German, Silvia Pereyra and the Plant Pathology Group, Federico Condon and the Plant Genetic Resources and Seed Group. Thanks Roberto Diaz, Monica Rebuffo, Jaime Diaz, Silvina Stewart, Martha Diaz, Ruben Verges, Sergio Ceretta, Alberto Fassio, and Daniel Vazquez for greenhouse, field equipment, and general support. Thanks to Schenzer, Pino, Nolla, Gabriela, Victoria, and all the fantastic administrative staff; to Coco, Polaco, Fernando, Leo, Roberto, Richard, Ramallo, Mauricio and all the other technicians from different projects that helped me with different things; to Luciana and Sonia from ‘La Cantina’; to all the hourly workers that eventually helped in the research and especially to Laura, Noelia, Pablo, Carlos, and Mariela. Thanks to my student Alejandro Nopitch for his commitment and dedication to the project. Thanks to my mates from ‘the pabellon’ with whom I shared great asados: Daniela, Alejandro, Alejandro, Alejandro, Bernardo, Augusto, Alvaro, Leonardo, Maria Eugenia, and ‘la veteca’. Thanks to John Grierson, Nicolas Gutierrez and David Acker for arranging the collaborative work with INIA.

Field work was done the Uruguayan way, ‘with help from everyone and their dog’. Other students, friends, and family members helped me with fieldwork at different stages: Alejandro helped with field support; Javier and Estefani helped with seed preparation; Alejandro and Augusto helped with harvest; Erwin, Alvaro and Fernando helped with fieldwork in Young; Leti, Jaime, Cristina, Sebastian, Martin, Andrew, Ignacio, Nicolas, Marco, and German participated in the ‘get Young done’ weekend; Leti, Santiago, Gustavo, Martin, Juan Pablo, Diego, Miguel, and Alejandro helped in the ‘get me out of here’ weekend in La Estanzuela; Ina helped with logistics; Walter Losa, Rosina Brasesco, Santiago Salaberry, and the ‘Sociedad de Fomento Rural de Rio Negro’ helped with Young experiments; Gabriel, Jajo, Bernardo, Valentin, Alejandra, Alicia, Andrea and Marina helped with greenhouse work.

My lab work would not have been possible without the support from Monsanto. I am very thankful to Ted Crosbie and Sam Eatington for their support and encouragement. I am also thankful to Valerie, Chad, John, Dario, Bob, Chris, Anju, and all other members of the Ankeny lab who made my genotyping experience possible and enjoyable. Thanks
to The Land Institute staff, Wes Jackson, Jerry Glover, Cindy Cox, Lee DeHaan, David
van Tassel, and Stan Cox, and all the Natural Systems Agriculture fellows and
participants of the Summer Workshops 2003-2006.

I would like to thank all breeders who provided seed for the research: J. Eglinton, C.
Li, J. Helm, B. Rossnagel, I. Matus, J.E. Diaz, B. Cooper, D. Obert, K. Smith, J.
Lalic. And other researchers who provided helpful comments at the initial stages of the
research: Pat Hayes, Brian Stephenson, Eviatar Nevo, Ted Crosbie, Sam Eathington, and
Wes Jackson. I would like to thank John McKay and Robert Latta for writing the article
that provided the initial inspiration for this work.

Having the support of fellow graduate students made the whole experience more
enjoyable. I am thankful to the students in the Jannink’s Lab: Fred Iutzi, Murli Gogula,
Dong Hong Pei, Jin Long, Alona Chernyshowa, Shengqiang Zhong, Yoon-Sup So,
Massiel Orellana, and Julia Olmstead with whom I shared many hours of field work and
meetings. Thanks Alona for always offering a smile, and thanks Stu for your kindness. I
am also thankful to the students in the Brummer’s Lab: Heathcliffe Riday, Baldomero
Alarcon, Joseph Robins, Julia Olmstead, Muhammet Sakiroglu, and Xue-Hui Lee for the
fun despite the many hours of computer work. Thanks Julia for those great talks that
always gave me perspective. Thanks Heathcliffe for sharing your SAS knowledge and for
setting the standard. Thanks Baldo for making me experience my first winter in Iowa’s
field washing roots with freezing temperatures. And thank you all for the baby boom lab
experience. I am also thankful to the students and post-docs in the Nason’s Lab: Kristy
Bernhard, Young-Jin Chun, Kristy Halverson, Rodney Dyer, and John Stireman. Thanks
Kristy B. for sharing the experience.

Many other ISU faculty members contributed to enrich my Ph.D. experience. Thanks
to Charlie Brummer for bringing a different perspective. Thanks to Mike Lee for all his
support and encouragement, especially with the seminars. Thanks to Jode Edwards for an
amazing support with SAS and WINBUGS programming, for having the door open for
my questions, and for always providing me with the right answers. I am also thankful to
Gretchen Zdorkowski, Matt Liebman, Mary Wiedenhoeft, Ricardo Salvador, Neal Flora, Jan Flora, Fred Kirschenmann, John Pesek, Rick Exner, Bill Clark, and Brian Wilsey. Thanks to Deb Lewis for help in finding herbarium specimens for the identification of *Hordeum* species in Iowa. Thanks to the students in the EEB 585 Field Trip to Canada who helped me with collection of *H. jubatum* seed and data, and especially to Tricia and Jay. I am also grateful to Jaci Severson in the Agronomy Department for making life easier by being on top of everything and making sure I did not miss any deadline or form, she is the best. I would also like to thank all the great people in the Agronomy Department.

I would not be here if it was not for my previous mentors. My father, Jaime Gutierrez, has been my first and most illuminating mentor; I am forever grateful of all the wisdom he has shared. I would also like to thank my undergraduate advisors for their guidance: Jorge Monza, Jorge Franco, and Tabare Abadie. I am also very thankful to all the other undergraduate professors in the Facultad de Agronomia of the Universidad de la Republica for the great education they provided me.

I would not have survived in Ames this five and a half years if it was not for all the many friends with whom I shared my life: Laly, Mario and Alicia, Ale and Jajo, Maro and Andrea, Gretchen and Todd, Monica and Pablo, Pedro and Marina, Natalia and Facundo, Julia and Philip, Ann and Christopher, Gabriel, Esteban and Monica, Vicente, Majd, ‘el portugues’, Emmily and Isac, Cassi, Andy, John, Jaime and Lali, Maria and Ale, Javi, Mariana and Vale, Mollie and the twins, Bea, Christie and the other moms, Kristy, Leti, Sarah, Juane and Feca, Alejandro and Bernardo, Pete, Carlos, Mary, Sofia and Marilu, Louis and Rohini, Karina and Francisco, Mario and Lula, and all my friends from MATES and KINDERMATES. I would also like to thank The Police, Los Fabulosos Cadillacs and Ann Zimmerman for their music that kept us working in the field past dusk.

To my parents, Jaime and Laura, for their unconditional love and support all these years and for all their wisdom and advice. And to my mom especially, for always being there caring, truly understanding, and giving the most loving support we need. To my
siblings Javi and Maria, because ‘se que puedo contrar contigo, no hasta dos ni hasta
diez, sino contra contigo…’ To Mariana and Ale for being Family. To my beautiful niece
Valentina for brightening our days. I would like to thank all my extended family who
followed and supported me from whichever part of the world they or I were. To my in-
laws, Daniel and Serrana, Cuca, Cati and Pancho. To Potota, Leti, Sofia and Marilu, Ina,
and the Picassos. To the Gutis for being there all the way, and for asking, questioning,
and always loving.

Finally, I would like to thank Valentín for making me complete. I am truly inspired
and profoundly moved by your dedication and love. To Santiago and Matías for smiling
at me every day and making my life full. You are my sunshine.
Cebada “Ceibo”

Quiero ser borde de tus ensayos
para cuidarte y protegerte,
barrera de tus dificultades,
guardar del viento, que no te vuelques,
que tus ideas no se dispersen;
Cebada “Ceibo” pa’ contenerte.

Quiero llenar espacios vacíos,
recordándote de la querencia;
en tus distancias ser cercanía
y acompañarte en la impermanencia;
que esos missing no sean extraños
Cebada “Ceibo” de tus ausencias.

Quiero ser siempre tu fiel testigo
plantado en medio’e tus tratamientos,
ver como crecen esas plantitas,
como progresa tu experimento,
dar testimonio de tus cosechas:
Cebada “Ceibo” pa’ estar presente.

En esta tesis de nuestras vidas:
semilla, espiga, sol, lluvia y viento
quiero ser borde pa’ contenerte,
testigo de tu campo diverso,
releno vivo pa’ acompañarte,
quiero ser siempre Cebada “Ceibo”.

Valentín Picasso
XII-2005
RESUMEN - ABSTRACT IN SPANISH

La selección fenotípica en poblaciones puede ser probada a través de la comparación de la variación en caracteres cuantitativos (Q\textsubscript{ST}) con la variación en marcadores moleculares neutros (F\textsubscript{ST}). Un carácter se considera bajo selección divergente si Q\textsubscript{ST} > F\textsubscript{ST}, selección convergente si Q\textsubscript{ST} < F\textsubscript{ST}, y neutro si Q\textsubscript{ST} = F\textsubscript{ST}. Esta metodología asume que la selección es homogénea entre las poblaciones, pero puede fallar cuando no lo es. Los objetivos generales de este trabajo son mejorar la metodología para comparar Q\textsubscript{ST} y F\textsubscript{ST}, y estudiar la relación entre Q\textsubscript{ST} y F\textsubscript{ST} en cebada cultivada y silvestre. Desarrollamos una metodología que detecta presiones de selección heterogénea entre las poblaciones. La metodología consiste en la simulación de Q\textsubscript{ST} esperado bajo neutralidad para cada par de poblaciones y la utilización de propiedades de la distribución de Q\textsubscript{ST} y F\textsubscript{ST}. Los métodos tradicionales de comparación de Q\textsubscript{ST} y F\textsubscript{ST} no son capaces de detectar mezclas de selección divergente y convergente y por lo tanto la metodología que proponemos permitirá mejores estimaciones del tipo de selección. Utilizamos metodología tradicional y nueva para estudiar la historia evolutiva de caracteres cuantitativos en cebada silvestre. Evaluamos 14 caracteres cuantitativos y 56 marcadores moleculares (microsatélites) en 280 genotipos provenientes de 23 poblaciones de cebada silvestre. El uso combinado de estimadores permitió entender mejor los procesos evolutivos que afectan a las poblaciones. Finalmente proponemos el uso de Q\textsubscript{ST} y F\textsubscript{ST} en cebada cultivada. Evaluamos 20 caracteres cuantitativos y 66 microsatélites en 353 genotipos provenientes de 23 programas de mejoramiento de cebada que representan la distribución mundial. A través de la estimación conjunta de la diversidad a nivel molecular y fenotípico, identificamos programas de mejoramiento que podrían beneficiarse del intercambio de germoplasma, y programas de mejoramiento con características únicas deseables para conservar. Sin embargo, esta metodología debe utilizarse con precaución en caracteres relacionados con el fitness, caracteres con importante interacción genotipo por ambiente, y en germoplasma con estructura impuesta artificialmente. En resumen, desarrollamos una metodología que es efectiva en detectar selección heterogénea y proporcionamos gran evidencia de la ocurrencia de este fenómeno tanto en cebada cultivada como silvestre.
ABSTRACT

Phenotypic selection can be tested by comparing the inter-population variation in quantitative traits ($Q_{ST}$) against inter-population variation in neutral molecular markers ($F_{ST}$). A trait is considered under divergent selection when $Q_{ST} > F_{ST}$, under stabilizing selection when $Q_{ST} < F_{ST}$, and selectively neutral when $Q_{ST} = F_{ST}$. This approach assumes the form of selection to be consistent across populations and can fail to detect selection mosaics in which this consistency is lacking. The overall objectives of this work are to improve the methodology for comparing $Q_{ST}$ and $F_{ST}$, and to study the relationship between $Q_{ST}$ and $F_{ST}$ in cultivated and wild barley. First, we develop a methodology to detect situations in which the form of selection is heterogenous. The methodology simulates pair-wise $Q_{ST}$ under neutrality and focuses on the distribution of $Q_{ST}$ and $F_{ST}$. Our results indicate that studies employing traditional $Q_{ST}$ and $F_{ST}$ methods are insensitive to mixtures of disruptive and stabilizing selection and that more powerful inference concerning the form of selection may be gained via the methods introduced here. Second, we use overall and pair-wise estimates to study the evolutionary history of quantitative traits in wild barley. We evaluate 14 traits and 56 SSR markers in 280 genotypes from 23 populations of wild barley from the Fertile Crescent. We find that a combination of overall and pair-wise comparisons is better suited than the former alone to understand the evolutionary processes that shape populations. Third, we propose the use of $Q_{ST}$ and $F_{ST}$ studies in cultivated barley. We use 66 polymorphic SSR and 20 quantitative traits in 353 genotypes of barley from 23 breeding programs distributed worldwide. By simultaneously estimating population structure at morphological traits and neutral molecular markers, we identify compatible breeding programs for germplasm exchange, and breeding programs with unique characteristics worth preserving. However, caution is advised in the use of this methodology for fitness-related traits, traits with important GxE, and germplasm with strong artificially-imposed structure especially in an artificial selection context as in breeding programs. In summary, we develop a methodology that is effective for detecting selection mosaics and we provided ample evidence of the occurrence of this phenomenon in both wild and cultivated barley.
CHAPTER I. INTRODUCTION

Plant Genetic Diversity

Plant genetic diversity is a key component of any ecosystem (Frankel et al., 1995). Genetic diversity is also essential for breeding purposes and for developing more sustainable agricultural systems. Sustainable agricultural systems require genetic diversity to be more stable to changes in the environment (Brummer, 1998; Duvick et al., 2004; Stuthman, 2002). Additionally, breeding programs require genetic diversity as both an insurance against unforeseeable changes in the environment (i.e. diseases, attacks by pest, and inclement weather) and to maintain genetic progress (Brown-Guedira et al., 2000; Gepts, 2006; Rasmusson, 2001). Maintaining every single allele just in case it may be needed in the future is neither possible nor desirable. Trying to broaden the diversity of a breeding program to account for the unknown will only slow genetic progress because the selection intensity on the traits of interest would be very small. Germplasm banks can maintain genetic diversity that is not immediately needed in the breeding program. When new variation is needed because the environment has changed, specific genetic variation can be brought into the breeding program. However, using a meaningful measurement of genetic diversity is challenging (Kim and Ward, 2000; Purvis and Hector, 2000).

Different types of data have been used to attempt genetic characterization, including morphology (Ortiz et al., 2002), pedigree information (Cox et al., 1985; Delannay et al., 1983; Rasmusson and Phillips, 1997; Smith et al., 2004), and molecular markers (Donini et al., 2000; Kim and Ward, 2000; Koebner et al., 2003; Malysheva-Otto et al., 2007; Ordon et al., 2005; Russell et al., 1997). Molecular marker techniques were first used in 1966 when Lewontin and Hubby applied allozyme electrophoresis to study genetic diversity (Hubby and Lewontin, 1966; Lewontin and Hubby, 1966). The use of allozymes was rapidly adopted, and by the end of the 1970s and 1980s, hundreds of plant species had been characterized for their allozyme diversity (Hamrick and Godt, 1990). In the following years, PCR-based techniques for genetic marker analysis became available and...
were rapidly adopted in evolutionary and conservation studies because they are relatively inexpensive, fast, and involve less invasive sampling than other methods (Crandall et al., 2000). DNA-based marker diversity has increasingly been used in lieu of the direct measurement of morphological variation. However, the use of both molecular marker and morphological information should be used to obtain a more complete summary of the distribution of diversity among and within populations. Gepts (2006) states that genetic resources conservation and utilization is facilitated by the understanding of the genotypic basis of agriculturally important traits, and specifically by understanding the evolutionary, ecological and anthropogenic mechanisms that underlie the current characteristics of traits. He goes further to say that understanding the distribution of genetic diversity responsible for trait variation within and among populations becomes the ‘holy grail’ of the science of genetic resources.

**Measurements of Population Structure**

Different measurements have been developed for estimating among population variation. $F_{ST}$ is a commonly used measure of population structure at presumably neutral molecular marker loci and was first introduced by Wright in 1951, based on the concept of fixation indexes (1943; Wright, 1951; 1965). Several estimators of $F_{ST}$ have been developed including $G_{ST}$ (Nei, 1972; 1978), $\Phi_{ST}$ (Excoffier et al., 1992), $R_{ST}$ (Slatkin, 1993), and most notably $\theta_{ST}$ (Cockerham and Weir, 1983) in which $F_{ST}$ for an allele is related to the within- and among-population variance components as:

$$F_{ST} = \frac{\sigma^2_b}{\sigma^2_b + \sigma^2_w}$$  \[1\]

where $\sigma^2_b$ is the between- and $\sigma^2_w$ is the within-population variance in allele frequency. In 1993, Spitze (1993) applied this variance partitioning approach to obtain the standardized among-population variance for a quantitative trait as:

$$Q_{ST} = \frac{(1+f)\sigma^2_b}{(1+f)\sigma^2_b + 2\sigma^2_w},$$  \[2\]

where $f$ is the within-population inbreeding coefficient (Bonin et al., 1996). Defined in this way, when phenotypic variation is neutral with respect to natural selection and the
genetic component of phenotypic variance is completely additive, $Q_{ST}$ has the same expected value as $F_{ST}$.

Understanding the origin and maintenance of phenotypic diversity is one of the central issues of evolutionary biology (Brown, 1989; Frankel et al., 1995; Lynch et al., 1999) and is key for genetic resources conservation and utilization (Brown and Clegg, 1983; Brown, 1978; Gepts, 2006; Nevo, 1992; Nevo et al., 1979). The relative importance of natural selection versus the neutral processes of mutation, gene flow, and genetic drift in determining patterns of phenotypic diversity is still unclear (Gomez-Mestre and Tejedo, 2004; Gould and Johnston, 1972; Johannesson et al., 2004; Rousset, 1997; Slatkin, 1993), but can be tested empirically by comparing genetic differentiation among populations for quantitative traits ($Q_{ST}$) against a null model of no selection represented by differentiation at neutral molecular markers ($F_{ST}$, Crnokrak and Merila, 2002; Lande, 1992; Lynch et al., 1999; McKay and Latta, 2002; Reed and Frankham, 2001). If $Q_{ST} = F_{ST}$, migration, mutation, and drift cannot be ruled out as the sole causes of quantitative trait differentiation among populations (Gomez-Mestre and Tejedo, 2004; Kuittinen et al., 1997; Lynch et al., 1999). In contrast, local directional or disruptive selection is expected to lead to $Q_{ST} > F_{ST}$, while stabilizing selection should lead to $Q_{ST} < F_{ST}$ (Morgan et al., 2001; Petit et al., 2001; Storz, 2002).

Valid comparisons of $Q_{ST}$ and $F_{ST}$ rest upon assumptions that are potentially restrictive. Fortunately, recent studies indicate that good results can be obtained even when these assumptions are not fully met (Bonnin et al., 1996; Goudet and Buchi, 2006). First, although only purely neutral molecular markers should be used for $F_{ST}$ estimation (Merilä and Crnokrak, 2001), because estimation is typically over numerous loci it is robust to deviations from neutrality at individual loci (e.g. Allendorf and Seeb, 2000). Second, the contrast of $Q_{ST}$ and $F_{ST}$ assumes that $F_{ST}$ is at migration-drift equilibrium (Hedrick, 1999) however, it has been demonstrated that $F_{ST}$ achieves equilibrium faster than its components (Crow and Aoki, 1984; Slatkin, 1993). Third, only purely additive effects should be used to estimate $Q_{ST}$ (Wright, 1951) and it has been argued that epistasis could bias $Q_{ST}$ upward (Lynch et al., 1999) or downward (Whitlock, 1999), and
that dominance could also bias $Q_{ST}$ upward (Lopez-Fanjul et al., 2003; Whitlock, 1999) or downward (Lopez-Fanjul et al., 2003; Porcher et al., 2006; Whitlock, 1999). However, dominance and some epistatic interactions can be controlled for using appropriate experimental designs (Lynch and Walsh, 1998), with parental half-sib designs best estimating additive effects in outcrossing species and selfed families offering the best approach in selfing species (Goudet and Buchi, 2006). Fourth, in estimating $Q_{ST}$, the need for panmixis and Hardy-Weinberg equilibrium within populations (Lande, 1992; Yang et al., 1996) has been relaxed to permit incorporation of selfing species (Bonnin et al., 1996). Fifth, low precision of $Q_{ST}$ estimates can be addressed using better estimators, (e.g. Bayesian $Q_{ST}$; O’Hara and Merila, 2005), and more populations (O’Hara and Merila, 2005). In sum, the $Q_{ST}$ versus $F_{ST}$ approach has been widely used to test for phenotypic selection in a diversity of natural populations (see review by Leinonen et al., 2008).

Overall versus Pairwise Comparisons

Traditionally, comparison of $Q_{ST}$ and $F_{ST}$ has been made using summary statistics estimated over all sample populations (McKay and Latta, 2002; Merilä and Crnokrak, 2001). As measures of centrality, these summary statistics provide useful descriptors of genetic differentiation when the evolutionary mechanisms underlying differentiation are consistent across populations. $Q_{ST}$ and $F_{ST}$ may not accurately represent population-to-population patterns of variation, however, when selection is heterogeneous across populations. Indeed, the same numerical value of such summary statistics can be obtained for an infinite array of different inter-population relationships (Dyer and Nason, 2004). Consequently, when the hypothesis $Q_{ST} = F_{ST}$ is tested, false inference concerning the nature of natural selection may be reached. Of particular interest, a trait may be classified as neutral when it is not if there exists a geographic mosaic of selection pressures with different subsets of populations subject to divergent selection and to stabilizing selection. Such selection mosaics may characterize species occupying distinct but geographically proximal habitats (Dyer and Nason, 2004; Steinger et al., 2002) with the outcome that local adaptation leads to phenotypic differentiation between and phenotypic similarity
within habitat associated populations in a manner that does not reflect their spatial relationships. Such conditions may not be uncommon in nature and represent a significant but largely unappreciated challenge to the detection of phenotypic selection comparing summary estimates of $Q_{ST}$ and $F_{ST}$. In response to this challenge, we develop and apply methods utilizing population pair-wise estimates of $Q_{ST}$ and $F_{ST}$ ($Q_{ST}^{ij}$ and $F_{ST}^{ij}$, respectively) capable of identifying heterogenous selection on phenotypic traits under conditions where overall statistics fail to reject neutrality.

*Hordeum* genus


For several reasons, the genus *Hordeum*, and particularly cultivated barley (*Hordeum vulgare* L. ssp. *vulgare*) and wild barley (*Hordeum vulgare* ssp. *spontanuem* C. Koch), the ancestor of cultivated barley (Harlan, 1971; Harlan and Zohary, 1966; Nevo, 1992), offer an ideal model system for the study of molecular marker and quantitative trait diversity. First, the genus shows a high degree of biological diversity, with adaptations to several environmental conditions (Hayes et al., 2003) and high levels of genetic variation and population differentiation (Brown, 1992; Brown and Clegg, 1983; Nevo, 1992; Turpeinen et al., 2001). Information on population structure at both macro-geographic and micro-geographic scales is available for wild barley (Baek et al., 2003; Nevo, 1992; 1998a, b; 2001; Nevo et al., 1983; 1979; 1997; 1981; 1984; 1986a, b, c, d, e). Consequently, sufficient power for $F_{ST}$ studies and $F_{ST}$-$Q_{ST}$ comparisons is available. Second, a range of selection pressures are expected to operate across the geographical ranges of both cultivated and wild barley. Cultivated barley was one of the first crops to be domesticated 10,000 years ago (Harlan, 1971) and has undergone intensive breeding for more than one century (van Hintum, 1994). Because of this breeding, different selection pressures are expected for different traits. A range of selection pressures is also
expected in wild barley due to its adaptation to diverse environments (Gutterman and Gozlan, 1998; Nevo, 1992; Nevo et al., 1983; Nevo et al., 1979; Volis et al., 2002; Whabi and Gregory, 1989), and substantial phenotypic variation has been found among populations (Ivandic et al., 2003; Nevo et al., 1979; Nevo et al., 1984; van Rijn et al., 2000; Vanhala et al., 2004). Third, both species are easy to work with; they are diploid, short-lived, self-fertilizing annuals (Bothmer et al., 1995; Brown and Marshall, 1981; Harlan and Zohary, 1966), and their life history characteristics and reproductive biology are well studied (Harlan and Zohary, 1966; Nevo, 1992). Finally, both species have long been the objects of research and have well-developed research tools. Cultivated barley is the fifth most important cereal crop in the world (Nevo, 1992), and a large collection of rapid inexpensive PCR-based DNA markers (e.g. SSRs) are available (Ramsay et al., 2000). Wild barley has also been considerably studied especially at the Institute of Evolution, University of Haifa in Israel (Nevo, 1992). Additionally, many molecular markers developed for cultivated barley can also be used in wild barley (Bothmer, 2003) because they have the same genome (Bothmer et al., 1995; Harlan, 1971; Harlan and de Wet, 1971).

**Dissertation Objectives**

The overall objectives of this dissertation are to improve the methodology for comparing $Q_{ST}$ and $F_{ST}$, and to study the relationship between $Q_{ST}$ and $F_{ST}$ in two *Hordeum* species: cultivated barley, and its ancestor, wild barley. We accomplish these overall objectives by pursuing three research goals. First, we develop and apply new methodology for using population pair-wise $Q_{ST}$ and $F_{ST}$ contrasts to study selection on phenotypic variation. The working hypothesis is that because the evolutionary processes that shape variation across populations operate at the population-to-population level, our focus should be on the distribution of pair-wise $Q_{ST}$ and $F_{ST}$ estimates. Indeed, the development of our pair-wise approach represents a response to the observation that distilling variation down into overall summary statistics may result in vital information being lost and wrong conclusions reached. Second, we combine the use of overall and
pair-wise estimates of $Q_{ST}$ and $F_{ST}$ to study the evolutionary history of quantitative traits in wild barley. The working hypothesis is that the use of pair-wise comparisons will provide a better understanding of the population structure by correctly assessing deviations from neutrality in traits that are classified as neutral by overall statistics. Third, we employ comparison of $Q_{ST}$ and $F_{ST}$ studies in cultivated barley. The working hypothesis is that $Q_{ST}$ and $F_{ST}$ studies can aid in genetic resources conservation and utilization by providing an understanding of the structure of the diversity across breeding programs.

**Dissertation Organization**

In Chapter II, we provide a methodology addressing mosaics of disruptive and stabilizing selection where comparison of overall $Q_{ST}$ and $F_{ST}$ estimates could be misleading as to the strength and form of natural selection acting on quantitative trait variation. Specifically, we first introduce methods focusing on the distribution of pair-wise estimates of genetic differentiation among populations by simulating phenotypic data under the hypothesis of selective neutrality. Specifically, the data is simulated from the observed within-population phenotypic variance and observed among-population covariance structure determined from neutral molecular marker data. Second, we use simulated data to demonstrate the ability of this approach to detect a mixture of disruptive and stabilizing selection for a situation were overall estimates indicated that $Q_{ST} = F_{ST}$. Finally, we provide an empirical example in wild barley where the approach succeeds at detecting selection while summary $Q_{ST}$ and $F_{ST}$ comparisons do not.

In the following chapters, we report results from an extensive field experiment conducted in two locations in Uruguay during the year 2005, where 280 genotypes of wild barley from 23 natural populations from the Fertile Crescent, and 353 genotypes of cultivated barley from 23 breeding programs distributed worldwide were evaluated for a total of 20 morphological traits. In Chapter III, we study the patterns emerging from $F_{ST}$ and $Q_{ST}$ comparisons on a broad range of functionally different quantitative traits in wild barley. Specifically, we
first test the hypothesis of neutral phenotypic evolution in wild barley. Second, we assess whether traits that were classified as neutral by overall statistics were still neutral when pair-wise comparisons were also evaluated. Finally, we discuss the use of pair-wise comparisons to identify diverging populations (i.e. populations that experienced $Q_{ST}^{ij} > F_{ST}^{ij}$ for comparisons of most traits).

In Chapter IV, we quantify worldwide genotypic diversity of cultivated barley for morphological traits and use this information to provide a methodology to aid in germplasm exchange between breeding programs. Specifically, we first characterize genetic diversity at morphological traits of advanced inbred lines of cultivated barley. Second, we describe the diversity of the breeding programs for those traits. Third, we develop a data-driven method identifying groups of breeding programs that would benefit from germplasm exchange in cultivated barley.

In Chapter V, we study the evolutionary history of agriculturally relevant traits in cultivated barley. Specifically, we first test the hypothesis of neutral evolution of 19 traits. Second, we describe the patterns of divergence across 19 breeding programs of cultivated barley distributed worldwide. Third, we compare the patterns of selection of cultivated with those obtained in wild barley. Finally, we discuss the relevance of $Q_{ST}-F_{ST}$ studies for breeding purposes, genetic resources conservation and utilization purposes, and for understanding the process of domestication.

References


Gutterman, Y., and S. Gozlan. 1998. Amounts of winter or summer rain triggering germination and ‘the point of no return’ of seedling desiccation tolerance, of some Hordeum spontaneum local ecotypes in Israel. Plant and Soil 204:223-234.


CHAPTER II. A NEW METHOD OF $Q_{ST}$-$F_{ST}$ CONTRASTS IDENTIFIES SELECTION MOSAICS IN WHICH SUMMARY $Q_{ST}$-$F_{ST}$ COMPARISONS FAIL TO REJECT NEUTRALITY

A paper submitted to *Genetics*

Lucía Gutiérrez, *,1 Jean-Luc Jannink, † and John D. Nason‡

Abstract

Phenotypic selection is often tested by comparing the standardized inter-population variation in quantitative traits ($Q_{ST}$) against variation in neutral molecular markers ($F_{ST}$). This common approach assumes the form of selection to be consistent across populations, but can fail to detect selection when it is not. In particular, as summary statistics, $Q_{ST}$ can equal $F_{ST}$ when different subsets of populations are under disruptive ($Q_{ST} > F_{ST}$) and stabilizing selection ($Q_{ST} < F_{ST}$), a situation likely to occur in nature. Here we present methodology to detect these situations that focuses on the distribution of population pairwise $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ estimates. We apply the method to both simulated and empirical data sets, the latter identifying non-neutral situations in wild barley despite the equality of overall $Q_{ST}$ and $F_{ST}$ estimates. These results indicate that studies employing traditional $Q_{ST}$ and $F_{ST}$ methods may be relatively insensitive to geographic mosaics of disruptive and stabilizing selection and that more powerful inference concerning the form of selection acting within and among populations may be gained via the methods introduced here.

* Department of Agronomy, Iowa State University, Ames, Iowa 50011-1010, † USDA-ARS, U.S. Plant, Soil, and Nutrition Laboratory, Ithaca, New York 14583, and ‡ Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, Iowa 50011-1010

Running head: Methodology to detect selection

Keywords: $F_{ST}$, $Q_{ST}$, pair-wise, stabilizing, disruptive selection

1 Corresponding author: Lucía Gutiérrez, Department of Agronomy, Iowa State University, Ames, 1203 Agronomy Hall, Ames, Iowa 50011-1010. Phone: (515)-294-6795. Fax: (515)-294-8146. E-mail: luciag@iastate.edu
Introduction

The relative importance of natural selection and neutral evolution can be tested empirically by comparing genetic differentiation among populations for quantitative traits ($Q_{ST}$) against a null model of no selection represented by differentiation at neutral molecular markers ($F_{ST}$, CRNOKRAK and MERILA 2002; McKAY and LATTA, 2002). If $Q_{ST} = F_{ST}$, migration, mutation, and drift cannot be ruled out as the sole causes of quantitative trait differentiation (GOMEZ-MESTRE and TEJEDO 2004; KUITTINEN et al. 1997; LYNCH et al. 1999). In contrast, local directional or disruptive selection is expected to lead to $Q_{ST} > F_{ST}$ while stabilizing selection causes $Q_{ST} < F_{ST}$ (MORGAN et al. 2001; PETIT et al. 2001; STORZ 2002). Traditionally, such comparisons are made using $Q_{ST}$ and $F_{ST}$ statistics estimated from a set of sample populations. As measures of centrality, these summary statistics provide useful descriptors of genetic differentiation when the evolutionary mechanisms underlying differentiation are consistent across populations. $Q_{ST}$ and $F_{ST}$ may not accurately represent population-to-population patterns of variation, however, when type of selection are heterogeneous across populations. Indeed, the same numerical value of such summary statistics can be obtained for an infinite array of different inter-population relationships (DYER and NASON 2004). Consequently, when the hypothesis of $Q_{ST} = F_{ST}$ is tested, false inference concerning the nature of natural selection may be reached.

Of particular interest, a trait may be falsely classified as neutral if there exists a geographic mosaic of selection pressures with different subsets of populations subject to divergent selection and to stabilizing selection. Such selection mosaics may characterize species occupying distinct but geographically proximal habitats with the outcome that local adaptation leads to phenotypic differentiation between and phenotypic similarity within habitat-associated populations in a manner that does not reflect their spatial relationships. This scenario is illustrated in Figure 1 for a hypothetical organism occurring in adjacent but replicated mountain and valley habitats. If selection favors similar phenotypic optima within and different optima between environments (Figure 1a), then the average phenotypic differentiation among populations ($Q_{ST}$) may be statistically
indistinguishable from differentiation at neutral markers ($F_{ST}$) driven by non-selective forces (e.g., isolation by distance). A focus on the distribution of pair-wise estimates of genetic differentiation among populations, however, offers greater insight into the presence and nature of phenotypic selection. While for each pair of populations $i$ and $j$ the difference $\Delta_{ij} = Q_{ST}^{ij} - F_{ST}^{ij}$ may have expectation zero, relative to neutral expectations excessively large positive and negative values of $\Delta_{ij}$ are indicative of disruptive and stabilizing selection, respectively, and would inflate the variance in $\Delta_{ij}$. In contrast, under a purely neutral scenario (Figure 1b), overall and pair-wise estimates have expectation $Q_{ST} = F_{ST}$, the distribution of $\Delta_{ij}$ should not exhibit excess extreme values, and, consequently, the variance of $\Delta_{ij}$ should not be overly large.

The aim of this study is to provide methodology addressing geographic mosaics of disruptive and stabilizing selection where comparison of overall $Q_{ST}$ and $F_{ST}$ estimates could be misleading as to the strength and form of natural selection acting on quantitative trait variation. In particular, we introduce and evaluate methods focusing on the distribution of pair-wise estimates of genetic differentiation among populations. Central to this approach is a procedure for simulating phenotypic individuals under the hypothesis of selective neutrality, with inter-population covariation determined by differentiation at neutral genetic markers and intra-population variation determined by observed phenotypic values. These individuals are used to construct a null distribution of $\Delta_{ij}$ values, which ultimately serves to evaluate the signals of both disruptive and stabilizing phenotypic selection across populations. We begin by describing the procedural steps in our approach to detecting geographical selection mosaics. Using simulated data, we then demonstrate the ability of this approach to detect a mixture of disruptive and stabilizing selection for a situation were overall $Q_{ST} = F_{ST}$. Finally, we provide an empirical example in wild barley in which our approach succeeds at detecting selection while summary $Q_{ST}$ and $F_{ST}$ comparisons do not.
Materials and Methods

Methodology: Population pair-wise $Q_{ST}^{ij}$ (and $F_{ST}^{ij}$) estimates are not independent and, consequently, classical statistical approaches cannot be used to determine the statistical significance of their deviation from selective neutrality either individually or in aggregate. We addressed this problem by employing a simulation-based procedure to construct the distribution of $Q_{ST}^{ij}$ expected under neutrality. For $Q_{ST}^{ij}$ values to reflect neutrality, an estimate of the underlying population genetic structure must be available, and this pattern of inter-population differentiation preserved when simulating phenotypes. We assume this neutral structure to be reflected in the matrix of population pair-wise covariance calculated from genetic distances estimated from a battery of presumably neutral genetic markers. For our purposes these distances must be Euclidean and were calculated in the manner of AMOVA (EXCOFFIER et al 1992; EXCOFFIER and SMOUSE 1994; SMOUSE and PEAKALL 1999). Specifically, we constructed a square matrix ($D$) of inter-population distances with elements $d_{ij}^2$ calculated as the average of the squared multi-locus genetic distances between all pairs of individuals belonging to populations $i$ and $j$. In order to incorporate this neutral pattern of differentiation into the simulation of phenotypes, $D$ was translated into a covariance matrix $C$ following GOWER (1966) with inter-population covariances calculated as

$$C_{ij} = \left( -\frac{1}{2} \right) \left[ d_{ij}^2 - \frac{\sum_{j=1}^{p} d_{ij}^2}{P} - \frac{\sum_{i=1}^{p} d_{ij}^2}{P} + \frac{\sum_{i\neq j} d_{ij}^2}{P^2} \right],$$

where $P$ is the number of sample populations. The variance within each population ($c_{ii}$) was obtained as the sum of squares within populations divided by its degrees of freedom following EXCOFFIER et al (1992). We then combined $c_{ii}$ and $c_{ij}$ into a neutral genetic variance-covariance matrix
Phenotypic population means reflecting the inter-population covariance structure at the neutral molecular markers scaled to the observed phenotypic variance, and within-population variances representing the observed phenotypic variance within each population were used to obtain $Q^p_{ST}$. To obtain phenotypic mean values for each population we used the square-root matrix of the phenotypic variance-covariance matrix adjusted by the structure in neutral markers. We first standardized $C$ to obtain the correlation matrix ($R$). We then obtained a phenotypic variance-covariance matrix ($VCV$) that represents the inter-population structure observed in the molecular markers (in the form of $R$) and the observed within population phenotypic standard deviations ($s_{Pi}$) as follows:

$$V_{CV} = \begin{bmatrix} s_{P1} & 0 & ... & 0 \\ 0 & s_{P2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & s_{Pn} \end{bmatrix} \begin{bmatrix} 1 & r_{12} & ... & r_{1n} \\ r_{21} & 1 & ... & r_{2n} \\ ... & ... & ... & ... \\ r_{n1} & r_{n2} & ... & 1 \end{bmatrix} \begin{bmatrix} s_{P1} & 0 & ... & 0 \\ 0 & s_{P2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & s_{Pn} \end{bmatrix}$$

$$= \begin{bmatrix} s_{P1}^2 & \text{cov}_{12} & ... & \text{cov}_{1n} \\ \text{cov}_{21} & s_{P2}^2 & ... & \text{cov}_{2n} \\ ... & ... & ... & ... \\ \text{cov}_{n1} & \text{cov}_{n2} & ... & s_{Pn}^2 \end{bmatrix},$$

where $s_{Pi}$ is the observed phenotypic standard deviation of population $i$. The resulting matrix ($V_{CV}$) contains the observed phenotypic variance within populations on the diagonal ($s_{Pi}^2$) and the phenotypic covariance among populations expected under neutrality on the off-diagonals ($\text{cov}_{ij}$). We then obtained the square root of the phenotypic variance-covariance matrix ($S-V_{CV}$) as an eigenvalue decomposition: $S-V_{CV} =$
\( \mathbf{V}^* \mathbf{E}^* \mathbf{t}(\mathbf{V}), \) where \( \mathbf{V} \) is the matrix of eigenvectors, \( \mathbf{E} \) is a diagonal matrix containing the square root of the eigenvalues, and \( \mathbf{t}(\mathbf{V}) \) is the transpose of \( \mathbf{V} \). We multiplied \( \mathbf{S-VCV} \) by a vector of random variables \( (x_i) \) obtained from a standard normal distribution to obtain a resulting vector of the phenotypic mean values for each population \( (\bar{X}_p) \) conditional on the inter-population structure of the molecular markers:

\[
\bar{X}_p = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{bmatrix}.
\]

Least square estimates of the among population variance (VB) were obtained from the difference in mean estimates. We calculated \( Q_{ST}^{ij}.e \) expected under neutrality in the traditional manner as the ratio of the between and total population variances:

\[
Q_{ST}^{ij}.e = \frac{(1 + f)VB}{(1 + F)VB + VW},
\]

where \( f \) is the species-level estimate of the within-population inbreeding coefficient (i.e., \( F_{IS} \)), VB is the phenotypic variance between expected under neutrality, and VW is the observed phenotypic variance within populations.

Finally, to provide a valid test of neutrality (given the non-independence structure of the data), we obtained the distribution of \( \Delta_{ij}.e \) expected under neutrality and its variance. Population pair-wise estimates \( F_{ST}^{ij} \) were obtained from the observed genetic marker data using the methods of AMOVA (\( \Phi_{ST}^{ij} \), EXCOFFIER et al. 1992) and the program GENALEX (PEAKALL and SMOUSE 2006). From these we then calculated \( \Delta_{ij}.e = Q_{ST}^{ij}.e - F_{ST}^{ij} \) and the variance of \( \Delta_{ij}.e \) (\( \text{Var}[\Delta_{ij}.e] \)) across population pairs. For each quantitative trait, this process was repeated for the full set of sample populations 10,000 times to generate a distribution of \( \text{Var}(\Delta_{ij}.e) \) from which critical values at the 5% level for the variance in \( \Delta_{ij} \) under neutrality were obtained. For a given trait we were then able to reject the null hypothesis of neutrality if the observed variance in \( \Delta_{ij} \) (\( \text{Var}[\Delta_{ij}.o] \)) fell outside of these critical values, where \( \Delta_{ij}.o = Q_{ST}^{ij}.o - F_{ST}^{ij} \).
Simulation examples: To demonstrate our approach and its ability to detect selection under situations where overall $Q_{ST} = F_{ST}$, data sets were simulated in accord with the non-neutral scenario represented in Figure 1a. Populations were created occurring in one of two distinct habitats. Within the same type of habitat, populations experienced similar environments but were geographically distant from each other. In contrast, populations in different habitats could be geographically proximate. Differentiation at molecular markers was neutral and proportional to geographical distance, whereas phenotypic differentiation was primarily determined by adaptation to local environmental conditions. This not un-natural situation resulted in population pairs representing three general classes of $\Delta_{ij}$ values: $\Delta_{ij} >> 0$ for pairs where phenotypic divergence was larger than expected under neutrality (proximal populations in different habitats), $\Delta_{ij} << 0$ for pairs where phenotypic divergence was less than expected under neutrality (distant populations in the similar habitats), and $\Delta_{ij} \approx 0$ for the remainder. Importantly, this scenario also results in a geographic selection mosaic in which overall $Q_{ST} \approx F_{ST}$.

Multilocus molecular genotypes were generated for each individual assuming linkage equilibrium among 10 neutral, codominant loci, each with 5 alleles (comparable to many studies using microsatellite data). Allele frequencies were arbitrarily selected to reflect the neutral, isolation-by-distance relationship among populations described above, and to provide an overall $F_{ST} \approx 0.5$ in accord with our empirical wild barley example described below. The phenotype of each individual was sampled from a random normal distribution with arbitrary means and standard deviations such that overall $Q_{ST} = F_{ST}$ and the pairwise $Q^i_{ST}$ reflect the selection mosaic described in Figure 1 (i.e., small $Q^i_{ST}$ between pairs of populations with similar environmental conditions and large $Q^i_{ST}$ between pairs of populations with different environmental conditions).

In accord with our wild barley example we set the number of populations at $N = 4$ (simulations with 12 populations yielded similar results; data not shown). We varied the number of individuals within populations ($n = 25$, and 75) with samples of 25 individuals per population being typical of many population genetic studies. We also varied the average within-population inbreeding coefficient ($f = 0$ and $f = 0.97$). An inbreeding
coefficient of $f = 0$ represents outcrossing populations that are very common in nature, whereas $f = 0.97$ represents an extreme level of inbreeding, a situation common to wild barley and to other predominantly selfing plant species.

An empirical example from wild barley: To further demonstrate that our pair-wise approach can detect selection where the traditional approach does not, we provide an example involving four populations of wild barley (*Hordeum spontaneum* K. Koch) from the Fertile Crescent. Populations were: Ashqelon (34.60°E, 31.63°N), Talpiyyot (35.25°E, 31.75°N), Bar Giyyora (35.08°E, 31.72°N), and Rosh Pinna (35.52°E, 32.95°N). Ten individuals from each of Ashqelon, Bar Giyyora, and Rosh Pinna populations, and 20 individuals from Talpiyyot were collected by E. Nevo and collaborators (NEVO, 1979). Phenotypic evaluations of flag leaf width were conducted in common garden experiments at two locations in Uruguay during the year 2005. Progeny from selfed-individuals were planted in hill-plots in a row-column design. Flag leaf width measured at 2.5 cm from the ligulae was recorded for five stems in each plot. REML variance components of populations and genotypes within populations were estimated and used in overall and population pair-wise $Q_{ST}$ estimation. The estimation of $Q_{ST}$ was performed in the R package (IHAKA and GENTLEMAN 1996).

An estimate of $F_{ST}$ ($Φ_{ST}$) was obtained over 56 SSR loci using AMOVA (EXCOFFIER et al. 1992) in GENALEX (PEAKALL and SMOUSE 2006). Euclidean genetic distances were also calculated in GENALEX. Simulation of $Q_{ST}^{ij}$, $Δ_{ij}$, and $\text{Var}(Δ_{ij})$ expected under neutrality was performed using the R package (IHAKA and GENTLEMAN 1996).

Results and Discussion

Simulated example: Overall $Q_{ST}$ does not differ significantly from $F_{ST}$ for any of the model data sets (Table 1). This traditional comparison of summary statistics thus leads to the false conclusion of phenotypic neutrality. Analysis of pair-wise estimates of $Q_{ST}^{ij}$ and $F_{ST}^{ij}$, in contrast, leads us to correctly reject the hypothesis of neutrality for each of the model data sets. Specifically, the variance in $Δ_{ij}$ observed is significantly larger than
the variance in $\Delta_{ij}$ simulated under neutrality (Table 1), reflecting the model structure in which some pairs of populations were constructed to be under stabilizing selection, and others under disruptive selection (Figure 2).

In general, our pair-wise approach testing the variance in $\Delta_{ij}$ performed well for all combinations of model parameters examined (Table 1). We found that there was a small bias in $\Delta_{ij}$ estimates for small population sizes, however, the ability to detect significant results did not change for 5, 10, 25, 50, and 75 individuals within populations, as similar critical values were found (data not shown). Furthermore, population sizes of published $Q_{ST}$-$F_{ST}$ studies range from 2 to 81 individuals with an average of 10 and a median of 7 (LEINONEN et al. 2008), a range well represented with the model data sets.

The pair-wise $\Delta_{ij}$ approach also performed well for inbreeding coefficients of both $f = 0$ and $f = 0.97$, with the null hypothesis of neutrality properly rejected regardless of population size. There was a slightly higher bias in $\Delta_{ij}$ for data sets at $f = 0.97$, but it decreased rapidly with increasing population size (data not shown). Given the extremes of inbreeding represented, we expect the pair-wise approach to perform well at any inbreeding level.

We simulated 4 and 12 populations in the model data sets, the former to reflect the situation found in the wild barley example. Even though some caution about estimations of $Q_{ST}$ from small numbers of populations has been raised (OHARA and MERILA, 2005), the methodology required in these studies makes it common to study relatively few populations. Indeed, a recent review (LEINONEN et al. 2008) indicates 21 studies of overall $Q_{ST}$-$F_{ST}$ that examined four or fewer populations. Consequently, the number of populations in our model data sets is not uncommon. Furthermore, we also tested the pair-wise approach with simulations of twelve populations (data not shown) and were able to detect non-neutrality in those situations as we did for four populations. In sum, the pair-wise $\Delta_{ij}$ approach performed well under a range of population sample numbers commonly used in $Q_{ST}$-$F_{ST}$ studies.

In simulating model data sets, we chose parameters values representative of many
Q_{ST}-F_{ST} studies. For example, because a recent review (LEINONEN et al. 2008) found that 51 out of 77 studies used 10 or less loci, we used 10 loci in our simulation models. The overall Q_{ST} and F_{ST} values we used also are representative of numerous Q_{ST}-F_{ST} studies (LEINONEN et al. 2008).

In general, our model data sets demonstrate that situations may exist in nature in which overall Q_{ST}-F_{ST} values may reflect the structure present in the neutral molecular marker data and quantitative traits, but fail to provide evidence of selection when it does in fact exist. Under the same conditions, focus on the variance in $\Delta_{ij} = Q_{ij}^{ST} - F_{ij}^{ST}$, in contrast, provides an effective means of differentiating mosaics of disruptive and stabilizing selection from neutrality. By varying model parameters, our results also demonstrate that the approach is capable of detecting such selection mosaics under a range of sample population numbers and within population sample sizes.

**Empirical example from wild barley:** The general patterns and conclusions obtained from the model data sets were also obtained for the empirical example from wild barley. As with the model data sets, in wild barley, overall Q_{ST} was not significantly different from F_{ST} (Table 1), indicating that variation in the trait in question, flag leaf width, is neutral. The observed variance in $\Delta_{ij}$, however, was larger than expected under neutrality (Table 1), consistent with the trait being under disruptive selection in some populations and under stabilizing selection in others (Figure 3).

Due to lack of independence, we cannot statistically test the significance of the deviation from neutrality of individual values of $\Delta_{ij}$. Having constructed the distribution of $\Delta_{ij}$ expected under neutrality, however, we have a means of drawing attention to unusual population pairs potentially worthy of further study. In particular, the distribution of $\Delta_{ij}$ enables us to put critical values on the null hypothesis of neutrality, and population pairs falling beyond upper and lower critical values constructed with a conservative Bonferroni-corrected p-value are likely candidates for disruptive or stabilizing selection, respectively. Using this approach, disruptive selection was found for the population pairs Ashqelon-Bar Giyyora, Ashqelon-Rosh Pinna, Talpiyyot-Bar Giyyora, and
Talpiyyot-Rosh Pinna, and stabilizing selection for the pair Ashqelon-Talpiyyot (Figure 3). These extreme $\Delta_{ij}$ values can be explained primarily by high or low inter-population differentiation in flag leaf width ($Q_{ST}^{ij}$) because all pairs of populations exhibited similar differentiation for molecular markers ($F_{ST}^{ij}$). The amount of genetic differentiation at molecular markers was relatively high for all pairs of populations, with the smallest differentiation being between populations Ashqelon and Bar Giyyora ($F_{ST}^{ij} = 0.376$), and the largest differentiation being between populations Ashqelon and Talpiyyot ($F_{ST}^{ij} = 0.583$). Genetic marker differentiation was not associated with geographical distance among populations (data not shown). In contrast, there were larger differences in population covariation for the quantitative trait; Ashqelon and Talpiyyot and Bar Giyyora and Rosh Pinna had relatively low $Q_{ST}^{ij}$, while the remaining pairs of populations had high $Q_{ST}^{ij}$. This distribution of $Q_{ST}^{ij}$ estimates is reflected in the mean values of flag leaf width, where individuals from Ashqelon and Talpiyyot had relatively wide flag leaves, with a mean flag leaf width of 52.4 and 47.8 mm respectively, and individuals from Bar Giyyora and Rosh Pinna populations had narrow leaves, with a mean of 27.4 and 31.7 mm respectively. These differences between the two pairs of populations are consistent with differences in mean annual rainfall and humidity, with wider- and narrower-leaved pairs of populations associated with locations of greater- and lower-water availability, respectively (NEVO et al. 1984; NEVO et al. 1979).

**Conclusions:** Many studies have used comparisons of overall $Q_{ST}$ and $F_{ST}$ to determine whether or not a trait is neutral (see LEINONEN et al. 2008 for a review). These studies largely fail to appreciate, however, that this approach assumes a homogeneous pattern of disruptive or stabilizing selection across populations and that geographical selection mosaics resulting from heterogenous mixtures of these forms of selection across different subsets of populations can result in an overall $Q_{ST} = F_{ST}$. Such outcomes, should they occur, would lead to false inference of neutrality, hindering our understanding of the relative importance of selective versus neutral processes in shaping the evolution of quantitative traits. In violation of assumptions of traditional $Q_{ST}$ and $F_{ST}$ approach,
selection mosaics combining disruptive and stabilizing selection are apparently common in nature (ENDLER 1986; THOMPSON 2005), an analytical challenge we address here by introducing methods focusing on the dispersion of $Q_{ST}$ and $F_{ST}$ contrasts as opposed solely to their central tendency. We provide simple simulated and empirical examples demonstrating situations in which our approach reveals significant evidence of selection while the contrast of overall $Q_{ST}$-$F_{ST}$ does not. Combined with growing scientific appreciation of the geographically dynamic nature of selection, our observations and results suggest that reanalysis of the many $Q_{ST}$-$F_{ST}$ studies that have failed to reject the null hypothesis of neutrality could yield new results and thus bring into clearer focus the true evolutionary patterns that shape populations.

**Literature Cited**


GOMEZ-MESTRE, I., and M. TEJEDO, 2004 Contrasting patterns of quantitative and neutral genetic variation in locally adapted populations of the Natterjack toad, Bufo calamita. Evolution 58: 2343-2352.


NEVO, E., A. BEILES, Y. GUTTERMAN, N. STORCH and D. KAPLAN, 1984 Genetic resources of wild cereals in Israel and vicinity. II. Phenotypic variation within and between populations of wild barley, Hordeum spontaneum. Euphytica 33.


**List of Figures**

**FIGURE 1.** Example of situations were populations could be either non-neutral (a) or neutral (b) when they have overall $Q_{ST} = F_{ST}$. In the non-neutral scenario, interpopulation phenotypic differentiation ($Q_{ST}^{ij}$) is driven by stabilizing and disruptive selection, favoring one phenotypic optimum in mountain populations (M1 and M2) and a different one in valley populations (V1 and V2), while differentiation at neutral markers ($F_{ST}^{ij}$) is driven by non-selective forces (e.g. isolation by distance). In this scenario, we expect mean $Q_{ST} = F_{ST}$, but with several pairs having $\Delta_{ij} > 0$ (divergent selection), and
several having $\Delta_{ij} < 0$ (stabilizing selection). Therefore, the variance of $\Delta_{ij}$ is large. In the neutral scenario, phenotypic differentiation is due to non-selective forces (i.e. genetic drift, migration and mutation). Inter-population differentiation of phenotypes ($Q_{ST}^{ij}$) is expected to be similar to differentiation of neutral markers ($F_{ST}^{ij}$). In this example $F_{ST}^{ij}$ is driven by geographical distance; populations that are close have low $F_{ST}^{ij}$ (M1-V1, and M2-V2), while populations that are far apart have large $F_{ST}^{ij}$. In this scenario, expectations are both mean $Q_{ST} = F_{ST}$ and pair-wise $Q_{ST}^{ij} = F_{ST}^{ij}$, therefore mean $\Delta_{ij} = 0$ and the variance of $\Delta_{ij}$ is small. For clarity, nomenclature was maintained between examples. See text for additional interpretations.

**FIGURE 2.** Relationship between observed (solid squares) and simulated (empty squares) pair-wise $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ values for a simulated data set with $n = 25$ individuals per population, $N = 4$ populations, and $f = 0$. Standard errors of $Q_{ST}^{ij}$ simulated under selective neutrality are represented with a grey line.

**FIGURE 3.** Relationship between observed (solid squares) and simulated (empty squares) pair-wise $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ values for the wild barley example. Standard errors of $Q_{ST}^{ij}$ simulated under selective neutrality are represented with a grey line.
Table 1. Q_{ST}, F_{ST}, and Δ (Q_{ST}-F_{ST}) estimates for the simulated non-neutral data sets (4 populations, population sizes n = 25 and 75, and inbreeding coefficients f = 0 and 0.97), and the wild barley example. In ‘Overall’ analyses conducted using traditional estimates of Q_{ST} and F_{ST}. Δ was not significant for all data sets. In ‘Pair-wise’ analyses conducted examining pairs of populations, the $F_{ST}^{ij}.o$, $Q_{ST}^{ij}.o$, and $Δ_{ij}.o$ represent pair-wise values observed, Var($Δ_{ij}.o$) is the variance in the observed $Δ_{ij}.o$, Var($Δ_{ij}.e$) is the variance in the expected $Δ_{ij}.e$ under neutrality. Critical values for the variance of $Δ_{ij}.e$ at the 5% level (c.v. Var[$Δ_{ij}.e$]) indicate significant deviation from neutrality in all data-sets.

<table>
<thead>
<tr>
<th>n</th>
<th>F</th>
<th>F_{ST}</th>
<th>Q_{ST}</th>
<th>Δ</th>
<th>$F_{ST}^{ij}.o$</th>
<th>$Q_{ST}^{ij}.o$</th>
<th>$Δ_{ij}.o$</th>
<th>Var($Δ_{ij}.o$)</th>
<th>Var($Δ_{ij}.e$)</th>
<th>c.v. Var($Δ_{ij}.e$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0</td>
<td>0.535</td>
<td>0.536</td>
<td>0.001</td>
<td>0.4480</td>
<td>0.6587</td>
<td>0.2107</td>
<td>0.3315</td>
<td>0.0693</td>
<td>[0.0082, 0.1920]</td>
</tr>
<tr>
<td>75</td>
<td>0</td>
<td>0.536</td>
<td>0.534</td>
<td>-0.002</td>
<td>0.4443</td>
<td>0.6226</td>
<td>0.1783</td>
<td>0.3971</td>
<td>0.0646</td>
<td>[0.0049, 0.1854]</td>
</tr>
<tr>
<td>25</td>
<td>0.97</td>
<td>0.380</td>
<td>0.393</td>
<td>0.013</td>
<td>0.3325</td>
<td>0.5614</td>
<td>0.2289</td>
<td>0.2991</td>
<td>0.0503</td>
<td>[0.0106, 0.1283]</td>
</tr>
<tr>
<td>75</td>
<td>0.97</td>
<td>0.194</td>
<td>0.259</td>
<td>0.065</td>
<td>0.1803</td>
<td>0.4121</td>
<td>0.2318</td>
<td>0.1830</td>
<td>0.0485</td>
<td>[0.0112, 0.1014]</td>
</tr>
<tr>
<td>Wild barley</td>
<td>0.501</td>
<td>0.688</td>
<td>0.187</td>
<td></td>
<td>0.4902</td>
<td>0.6333</td>
<td>0.1432</td>
<td>0.2217</td>
<td>0.0937</td>
<td>[0.0187, 0.1769]</td>
</tr>
</tbody>
</table>
Figure 1. Example of non-neutral and neutral cases with overall $Q_{ST} = F_{ST}$

- **Non-neutral example**
  - $Q_{ST}^H << F_{ST}^H$ for $M1-M2, V1-V2$
  - $Q_{ST}^H = F_{ST}^H$ for $M1-V2, M2-V1$
  - $Q_{ST}^H >> F_{ST}^H$ for $M1-V1, M2-V2$

- **Neutral example**
  - $Q_{ST}^H = F_{ST}^H$ for all pairs
    - low $Q_{ST}^H$ and $F_{ST}^H$ for $M1-M2, M1-V2, M2-V1$
    - high $Q_{ST}^H$ and $F_{ST}^H$ for $M1-V1, M2-V2$
Figure 2. Observed and expected pair-wise $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ in simulated data
Figure 3. Observed and expected pair-wise $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ in wild barley.
CHAPTER III: OVERALL AND PAIR-WISE COMPARISONS OF
QUANTITATIVE AND GENETIC POPULATION STRUCTURE IN WILD
BARLEY (Hordeum spontaneum K. Koch)

A paper to be submitted to Molecular Ecology

Gutiérrez, L.,¹ Nason, J.D., Jannink, J.-L.

Abstract

The relative importance of selection can be tested empirically by comparing population
differentiation for quantitative traits (Q<sub>ST</sub>) against that for neutral molecular markers
(F<sub>ST</sub>). Overall Q<sub>ST</sub> and F<sub>ST</sub> are commonly evaluated for this purpose; however, because
they are summary statistics they may not adequately reflect population-to-population-
level variation. The aim of this study was to understand patterns of differentiation for a
broad range of functionally diverse quantitative traits in wild barley. Specifically, we
used pair-wise comparisons of Q<sub>ST</sub> and F<sub>ST</sub> to test whether a combination of population
pairs under divergent selection with population pairs under stabilizing selection could
lead to a trait being to be classified as neutral, and to detect cases where selection and
gene flow act in concert. We evaluated 14 traits and 56 SSR markers in 23 populations of
wild barley from the Fertile Crescent. We found that for two traits the signal from some
pairs of populations under divergent selection and some pairs under stabilizing selection
canceled each other due to selection mosaics, such that the traits were misclassified as
neutral by overall statistics. Moreover, we identified one trait for which geographical
patterns of selection and gene flow reinforced each other. Additionally, pair-wise Q<sub>ST</sub> -
F<sub>ST</sub> comparisons were used to identify populations subject to strong divergent or
stabilizing selection. The results of this study indicate that a combination of overall and

Plant, Soil, and Nutrition Laboratory, Ithaca, New York 14583; J.D. Nason, Dept. of Ecology, Evolution
and Organismic Biology, Iowa State Univ., IA 50011.
Keywords: F<sub>ST</sub>, Q<sub>ST</sub>, pair-wise, Wild Barley, Population Structure
Corresponding Author: Lucia Gutierrez, 1203 Agronomy Hall, Iowa State University, Ames, IA 50011,
Fax: 515-294-8146, e-mail: luciag@iastate.edu
Running Title: Population structure in Wild Barley
pair-wise comparisons is better suited than the former alone to understand the evolutionary processes that shape populations.

**Introduction**

The relative importance of adaptive and non-adaptive causes of variation is central to understanding the evolution of genetic and phenotypic variation and their population structure (Storz 2002). Natural selection is responsible for adaptations to local environmental conditions and geographical differentiation of populations for quantitative traits (Gould & Johnston 1972; Johannesson et al. 2004). Selection is not, however, the only explanation for population structuring of quantitative traits, as non-selective (neutral) processes can also influence patterns of differentiation (Storz 2002). The relative importance of selection and neutral forces can be evaluated empirically by using genetic differentiation of populations for neutral molecular markers ($F_{ST}$) as a null hypothesis against which the significance of natural selection in creating population structure of quantitative traits is tested (Mc Kay & Latta 2002; Merila & Crnokrak 2001). The $Q_{ST}$ statistic, which partitions quantitative genetic variation in a manner analogous to $F_{ST}$, can be tested against $F_{ST}$ and if $Q_{ST} = F_{ST}$, non-selective processes (migration, drift and mutation) cannot be ruled out as the sole causes of quantitative trait differentiation (Gomez-Mestre & Tejedo 2004; Kuittinen et al. 1997; Lynch et al. 1999). In contrast, local directional or disruptive selection is expected to lead to $Q_{ST} > F_{ST}$, while stabilizing selection causes $Q_{ST} < F_{ST}$ (Morgan et al. 2001; Petit et al. 2001; Storz 2002).

Population differentiation is often expressed by statistics summarizing differentiation among several populations. The use of such summary statistics to understand population structure can be misleading because they may not adequately reflect variation in evolutionary forces acting at the inter-population level. These forces can vary depending on the scale of spatial separation, pathways and barriers to gene migration, and abiotic and biotic conditions influencing selection mosaics (Dyer & Nason 2004; Rhodes et al. 1996; Steinger et al. 2002). While summary statistics such as $F_{ST}$ (Wright 1951) and $Q_{ST}$ (Spitze 1993) are useful when the evolutionary mechanisms underlying population
differentiation are consistent across populations, they may not adequately represent population-to-population variation in differentiation and the responsible evolutionary forces when these forces are heterogeneous across populations. Indeed, the same numerical value of a summary statistic can be obtained for an infinite array of different inter-population relationships (Dyer & Nason, 2004). Consequently, when the hypothesis of \( Q_{ST} = F_{ST} \) is tested, incorrect conclusions about the nature of selection on quantitative traits may be reached (Gutierrez et al. 2008). The distribution of pair-wise contrasts between \( Q_{ST} \) and \( F_{ST} \) (i.e. between \( Q_{ST}^{ij} \) and \( F_{ST}^{ij} \)), in contrast, is better suited than summary statistics for study inter-population differentiation. For example, a trait can potentially be classified as neutral when it is not if there exist subsets of populations subject to divergent selection \( (Q_{ST}^{ij} > F_{ST}^{ij}) \) and to stabilizing selection \( (Q_{ST}^{ij} < F_{ST}^{ij}) \) that when combined result in overall \( Q_{ST} = F_{ST} \). Such limitations of \( Q_{ST} \) and \( F_{ST} \) as summary statistics argue for tests of quantitative versus neutral variation that more explicitly evaluate the dispersion of pair-wise \( Q_{ST}^{ij} - F_{ST}^{ij} \) as opposed to focusing on the central tendencies of such statistics.

Gutierrez et al (2008) recently introduced a methodology based on analysis of pair-wise \( Q_{ST}^{ij} \) and \( F_{ST}^{ij} \) values to address situations in which the traditional contrast of summary statistics \( (Q_{ST} - F_{ST}) \) fails to reveal selection on quantitative traits when it does in fact occur. There is increasing scientific appreciation for the fact that the form of and strength of natural selection is often highly dynamic in space and time (Thompson 2005). In contrast, the traditional \( Q_{ST} \) versus \( F_{ST} \) approaches assumes that the form of selection is essentially homogenous over populations (i.e. entirely disruptive or entirely stabilizing). Gutierrez et al (2008) thus focused on geographical selection mosaics in which some pairs of populations were subject disruptive selection and others to stabilizing selection. They demonstrated that in selection mosaics such as these, overall \( Q_{ST} = F_{ST} \) may be rejected, and present an approach based on the dispersion of pair-wise \( Q_{ST}^{ij} - F_{ST}^{ij} \) that successfully identifies selection under these conditions.

In this paper we apply traditional and pair-wise approaches to study of phenotypic and molecular genetic variation to investigate the evolution of population structure in wild
barley. For several reasons, the genus *Hordeum* and, particularly, *H. spontaneum* (wild barley), the ancestor of cultivated barley (Harlan & Zohary 1966; Nevo 1992; Zohary 1969), offers an ideal model system for the simultaneous study of molecular markers and quantitative traits. First, the genus exhibits a high degree of molecular genetic differentiation among populations (Brown et al., 1978; Brown & Clegg 1983; Nevo 1992; Turpeinen et al. 2001). Second, a range of selection pressures is expected in wild barley due to its adaptation to diverse environments (Gutterman & Gozlan 1998; Nevo 1992; Nevo et al. 1983; 1979; Volis et al. 2002; Whabi & Gregory 1989), and substantial phenotypic variation has been found among populations (Ivandic et al. 2003; Nevo et al. 1984; 1979; Van Rijn et al. 2000; Vanhala et al. 2004). Third, wild barley is a logistically amenable plant to work with; it is a diploid, short-lived annual (Bothmer et al. 1995; Brown & Marshall 1981), and its life history characteristics and reproductive biology are well studied (Nevo 1992). Fourth, the species is highly self-fertilizing (Nevo, 1992), minimizing potential bias in the estimation of $Q_{ST}$ attributable to dominance effects (Goudet & Bouchi 2006; Lopez-Fanjul et al. 2007). Finally, wild barley is the closely related progenitor of cultivated barley (*H. vulgare*), permitting the use of a battery of neutral molecular markers (e.g., SSRs) developed for the cultivated species (Ramsay et al. 2000).

The aim of this paper is to evaluate the utility of overall and pair-wise estimates of $Q_{ST}$ and $F_{ST}$ to understand the nature of selection acting on a broad range of functionally different quantitative traits in wild barley. We begin by identifying traits indicated by the contrast in overall $Q_{ST}$ and $F_{ST}$ to be under disruptive or stabilizing selection. Traits classified as neutral by overall statistics were subsequently tested for selection using the pair-wise approach of Gutierrez et al. (2008). Overall statistics may lead to a neutral classification ($Q_{ST} = F_{ST}$) either because the trait truly evolves neutrally or because cases of divergent ($Q_{ST}^{ij} > F_{ST}^{ij}$) and stabilizing ($Q_{ST}^{ij} < F_{ST}^{ij}$) selection across different pairs of populations caused by selection mosaics could cancel each other out. We identified two traits in our study that would have been wrongly classified as neutral by overall statistics in this way. We show that two traits in wild barley appear to be falsely classified as
neutral for this reason. We also use pair-wise comparisons of $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ to detect cases where selection and gene flow act in concert, identifying one trait in wild barley for which mechanisms that cause populations to differentiate due to restricted gene flow, and mechanisms that causes them to differentiate due to differential selection pressure, reinforce each other. These results indicate that combining the pair-wise approach of Gutierrez et al. (2008) with traditional $Q_{ST}$-$F_{ST}$ contrasts provides greater power that the latter approach alone to detect selection on quantitative traits, particularly when the form of selection is not homogeneous across populations.

Materials and Methods

Study Species and Sampling Strategy

Our initial sample of wild barley consisted of a total of 280 individuals from 23 natural populations representing diverse eco-geographical regions of the Fertile Crescent (Figure 1, Table 1). For each individual, a family of five seeds was generated by selfing in the greenhouse during the year 2005. These seeds were vernalized in soil at $5^\circ C$ for 21 days and seedlings pre-germinated in the greenhouse. One seedling was used for DNA extraction to estimate $F_{ST}$, with tissue sampled at the two-leaf stage. The other four seedlings were grown in greenhouse until maturity and allowed to self, with the resulting seed harvested for use in field evaluations to estimate $Q_{ST}$.

Microsatellite amplification and molecular marker analysis

Eighty SSR markers of cultivated barley (Becker & Heun 1995; Liu et al. 1996; Ramsay et al. 2000; Saghai Maroof et al. 1994; Struss & Plieske 1998; and Smith K. pers. com.) were screened for all 280 extracted individuals of wild barley. After discarding monomorphic and inconsistently amplifying loci, 56 polymorphic markers were scored, representing good genome coverage.

Fresh plant tissue was collected in 1.4 mL tubes, freeze-dried and ground with a paint shaker. Powdered tissue was incubated with extraction buffer (0.2M Tris, pH = 7.5, 25 µM EDTA, 25M NaCL, 25% SSD from EM Science) for 30 min at $65^\circ C$. Nucleic acids
were precipitated by adding a solution of 5M of potassium acetate, pH = 7.0 (EM Science). DNA was precipitated by the addition of isopropanol. After a 70% (v/v) ethanol wash and pelleting, DNA was hydrated in TE buffer (1 mM Tris, pH = 8.0, 0.4 mM EDTA, pH = 8.0 from OmniPur). DNA concentrations were quantified and the same genomic DNA concentration was used for all genotypes. Polymerase Chain Reaction (PCR) amplifications were performed in a final reaction volume of 10 µL containing 1 µL of 10X Gold Buffer (Perkin Elmer), 1 µL of 25 mM MgCl₂, 0.5 µL of 0.4 w/v cresol red, 0.25 µL of 99% Glycerol, 0.4 µL of 5 µM primers, 0.8 µL total of 100mM dNTPs, 0.05 µL of Taq polymerase, and 4 µL of 10 ng/µL genomic DNA.

SSR markers were grouped in three categories with each category using a different PCR condition (A, B, or C) for the annealing and elongation step: (i) one hold of 10 minutes at 94°C for denaturation; (ii-A) 35 cycles of 45 seconds each at 94°C for denaturation, 55°C for annealing, and 68°C for elongation; (ii-B) 35 cycles of 45 seconds each at 94°C for denaturation, 58°C for annealing, and 68°C for elongation; (ii-C) a “touch down” procedure with 2 cycles of 30 seconds each at 94°C for denaturation, 65°C for annealing, and 72°C for elongation; 10 cycles of 30 seconds each at 94°C for denaturation, 65°C for annealing, and 72°C for elongation with the annealing temperature reduced 1°C each cycle, followed by 25 cycles of 30 seconds each at 94°C, 55°C, 72°C; (iii) two holds of 5 minutes at 72 °C; and (iv) a 4°C hold. All PCR reaction products were electrophoresed in 20x25 cm, 2.8 % (w/v) metaphor-agarose (Cambrex Bio Science Rockland Inc., ME, USA) gels in 1X TBE (Tris Boris Acid, EDTA) buffer. Gels were run for 2.5 h at 190 V, stained with ethidium bromide, illuminated by UV light, photographed, and manually scored for presence/absence of clear bands.

To determine if repeat length homoplasy was a consequence of high mutation rates (Slatkin 1993), we calculated correlations between F<sub>ST</sub> and allelic richness (O’Reilly et al. 2004), heterozygosity (O’Reilly et al. 2004), length and number of repeats (Brohede et al. 2002; Ellegren 2000; Petit et al. 2005). Our data showed low number of alleles for most SSR (2-7 alleles with a mean of 3.8968), and non-significant correlations between F<sub>ST</sub>
and allelic richness ($r = 0.066, p = 0.633$), heterozygosity ($r = 0.160, p = 0.247$), length of the repeats ($r = 0.178, p = 0.225$), and number of repeats ($r = 0.006, p = 0.966$).

**Field trials**

The selfed descendants of our initial sample of 280 wild barley genotypes were evaluated during the year 2005 at Colonia and Young in Uruguay. Colonia is in southwestern Uruguay (34.20° S, 57.10° W, and 81 m.), while Young is in northern Uruguay (32.41° S, 57.40° W, and 80 m.). A row-column (alpha lattice) planting design with 17 rows, 17 columns and 3 replications was used. Seed was treated with a fungicide (carbendazim and thiram) and then pre-germinated and vernalized for 20 days at 5°C in tan brown 38 lb towel paper. Seedlings were transplanted into the field on August 11th and 12th in Colonia and Young respectively. Seedlings were arranged in hill-plots with a spacing of 40 cm between rows and alternating 40 and 60 cm between columns. The field was fertilized at planting with 45 kg ha$^{-1}$ of urea to reach 40 mg g$^{-1}$ of N as NO$_3^-$ (the optimal N content for cultivated barley). A hydration gel was used to keep roots moist, and plants were watered until the root system was established. Depending on the number of seedlings available, three to 20 seedlings descendant from a single genotype were transplanted into separate hill-plots in the field. The number of plants alive was counted at the two-leaf stage and this number was used as a covariate in all statistical analyses used to estimate $Q_{ST}$. Because of the threat of disease to the experimental plants, weekly monitoring of disease was performed and fungicide applied when necessary. There was one systemic fungicide application (pyraclostrobin + epoxiconazole) on October 26th and 31st in Young and Colonia respectively. Full-plots were harvested on December 6th and 8th in Young and Colonia, respectively.

Several phenotypic traits were recorded either on plots (i.e., hill-plots containing up to 20 plants from a single genotype) or on individual plants. The phenotypic traits measured at the plot level were: total number of plants present at two-leaf stage (NPL), total number of tillers present at the end of tillering, leaf rust disease (scoring from 1-5, where 1 is low and 5 is high), total number of days between planting date and anthesis, total
number of days between planting and flowering, biomass weight at plant maturity (g),
total number of spikes at maturity, and weight of all spikes in the plot (g).

At flowering time, up to five plants from each hill-plot were chosen at random, color-
marked with plastic twist band, and their spikes covered with pollinating bags to avoid
shattering. Phenotypic traits measured at the individual level on marked plants were: flag
leaf length measured from the ligulae to the tip of the leaf (cm), flag leaf width measured
at 2.5 cm from the ligulae (cm), spike length measured from the base to the tip of the
spike (cm), awns length measured from the tip of the spike to the tip of the longest awn
(cm), number of grains per spike, and weight of grains produced by the 5 marked spikes
(g).

**Molecular marker analysis and $F_{ST}$ estimation**

Standard population genetic measures were estimated from the 56 locus SSR data set in
GDA (Lewis & Zaykin 2001). Population differentiation was measured in terms of $F_{ST}$ as
estimated by Weir and Cockerham’s $\theta$ (Weir & Cockerham 1984):

$$\theta = \frac{V_B}{V_B + V_W},$$

where $V_B$ is the among-population variance component, and $V_W$ is the within-population
variance component. Ninety-five percent confidence intervals about $\hat{\theta}$ were obtained
through parametric bootstrapping with 1,000 replications. Unless otherwise specified, a
threshold of 5% ($p < 0.05$) was used in all the statistical tests.

Two estimates of population differentiation were obtained: a single overall estimate
($F_{ST}$), and a full set of population pair-wise estimates ($F_{ST}^{ij}$). $F_{ST}$ was estimated as $\theta$
for all populations and all genotypes within populations. $F_{ST}^{ij}$ were estimated as $\theta_{ij}$
for all pairs of populations $i$ and $j$, resulting in a 23x23 matrix of values.

The presence of spatial patterns of gene migration and isolation by distance between
populations was tested with Slatkin’s (Slatkin and Maddison, 1990; Slatkin, 1991; 1993)
and Rousset’s method (Rousset, 1997). Slatkin’s method uses the pairwise effective migration rate: $\hat{M} = [(1/F_{ST}) - 1]/4$. The slope of the regression of $\log_{10}(\hat{M})$ against $\log_{10}(d)$ where $d$ is the geographical distance among populations, is indicative of the spatial dimension of gene flow. A slope of -1.0 is expected under a one-dimensional stepping stone model, -0.5 under a symmetric two-dimensional model, and 0.0 under an island model (Slatkin and Maddison, 1990; Slatkin, 1991). Rousset’s method uses the regression of $F_{ST}/(1-F_{ST})$ against $d$ or $\log_{10}(d)$ for one-dimension and two-dimension models respectively.

Quantitative trait analysis and $Q_{ST}$ estimation

Traits measured at the plot- and plant-levels were modeled according to the following linear models respectively:

$$Y_{ijklmn} = B_{ij} + R_{k(ij)} + C_{l(ij)} + P_m + G_{n(m)} + \beta^* NPL_{ijkl} + \epsilon_{ijklmn}, \quad (2)$$

$$Y_{ijklmno} = B_{ij} + R_{k(ij)} + C_{l(ij)} + P_m + G_{n(m)} + I_{o(ijk)} + \beta^* NPL_{ijkl} + \epsilon_{ijklmno}, \quad (3)$$

where $B_{ij}$ denotes the effect of $i^{th}$ block in the $i^{th}$ environment, $R_{k(ij)}$ the effect of the $k^{th}$ row in the $i^{th}$ block-environment combination, $C_{l(ij)}$ the effect of the $l^{th}$ column in the $i^{th}$ block-environment, $P_m$ the effect of the $m^{th}$ breeding program, $G_{n(m)}$ the effect of $n^{th}$ genotype in the $m^{th}$ breeding program, $I_{o(ijk)}$ the effect of the $o^{th}$ plot on the $i^{th}$ block-environment, $\beta$ the regression coefficient (slope) associated with the number of plants per hill-plot, $NPL_{ijkl}$ the number of plants at the two-leaf stage used as a covariate in the model, $\epsilon_{ijklmn}$ the residual error for the $n^{th}$ genotype in the $m^{th}$ breeding program and $i^{th}$ block-environment, and $\epsilon_{ijklmno}$ the residual error for the $o^{th}$ plant of the $n^{th}$ genotype in the $m^{th}$ breeding program and $i^{th}$ block-environment. Finally, because plants within a plot share similarities from belonging to the same plot, a plot effect ($I_{o(ijk)}$) was included in the plant-level linear model (equation 3).

Bayesian analysis were found to produce some of the most precise confidence intervals for $Q_{ST}$ (O’Hara and Merila, 2005), and because confidence intervals for the comparison of $Q_{ST}$ and $F_{ST}$ could also be obtained directly with this approach, a Bayesian model was used. We defined a hierarchical Bayesian model where we modeled both
means \((B_{ij}, R_{k(i)}, C_{l(i)}, P_m, G_{n(m)}, \text{and } I_{o(ijk)})\) and associated variances \((\sigma^2_R, \sigma^2_C, \sigma^2_P, \sigma^2_G, \sigma^2_I, \text{and } \sigma^2)\) in terms of explanatory variables (block, location, row, column, breeding program, genotype, and plot). At the first level of the Bayesian hierarchy, observations from plots and from plants were modeled as independent samples from a normal distribution, respectively, as:

\[
Y_{ijklmn} \mid B_{ij}, R_{k(i)}, C_{l(i)}, P_m, G_{n(m)}, \beta \ast NPL_{ijklmn}, \sigma^2 \\
\sim N\left(\frac{B_{ij} + R_{k(i)} + C_{l(i)} + P_m + G_{n(m)} + \beta \ast NPL_{ijklmn}}{\sigma^2}\right) 
\]

\[
Y_{ijklmno} \mid B_{ij}, R_{k(i)}, C_{l(i)}, P_m, G_{n(m)}, I_{o(ijk)}, \beta \ast NPL_{ijklmn}, \sigma^2 \\
\sim N\left(\frac{B_{ij} + R_{k(i)} + C_{l(i)} + P_m + G_{n(m)} + I_{o(ijk)} + \beta \ast NPL_{ijklmn}}{\sigma^2}\right) 
\]

The second level of the Bayesian hierarchy includes prior distributions for location parameters (i.e. means) \(B_{ij}, R_{k(i)}, C_{l(i)}, P_m, G_{n(m)}, I_{o(ijk)}, \beta, \text{and } NPL_{ijklmn} \), and observational variance \(\sigma^2\). Priors on all location parameters were normal with mean zero and variances defined to condition the desired level of information sharing among levels of the factor (e.g. block, row, column, population, genotype, or plot).

For block-environment means, \(B_{ij}\), the prior was defined with a very large variance to make the prior non-informative: \(B_{ij} \sim N(0, 10^{-7})\). This flat and independent prior is the Bayesian equivalent of defining the block effect as a fixed effect in classical linear models (Edwards and Jannink, 2006). We did not include location effects in our model so that all parameters were estimable. For the regression coefficient of the covariate, \(\beta\), the prior was also defined as a fixed effect: \(\beta \sim N(0, 10^{-7})\).

Row, column, population, genotype, plot and the covariate number of plants (NPL) were modeled with priors that treated them as equivalent to random effects in classical mixed linear models. Row, column, population, genotype, plot, and NPL effects were modeled as samples from a normal distribution with variance \(\sigma^2_R, \sigma^2_C, \sigma^2_P, \sigma^2_G, \sigma^2_I, \text{and } \sigma^2_{NPL}\), respectively: 

\(R_{k(i)} \mid \sigma^2_R \sim N(0, \sigma^2_R), C_{l(i)} \mid \sigma^2_C \sim N(0, \sigma^2_C), P_m \mid \sigma^2_P \sim N(0, \sigma^2_P), G_{n(m)} \mid \sigma^2_G \sim N(0, \sigma^2_G), I_{o(ijk)} \mid \sigma^2_I \sim N(0, \sigma^2_I), NPL_{ijklmn} \mid \sigma^2_{NPL} \sim N(0, \sigma^2_{NPL})\).

Priors on the variance of row, column, population, genotype, plot, and NPL were chosen to be non-informative with an inverse gamma (IG) distribution: 

\(\sigma^2_R, \sigma^2_C, \sigma^2_P, \sigma^2_G, \sigma^2_I, \sigma^2_{NPL} \sim IG(0.0001, 0.0001)\). Residual variance was also modeled as \(\sigma^2 \sim IG(0.0001, 0.0001)\).
All parameters were obtained via Markov chain Monte Carlo simulation using the Bayesian Gibbs Sampling software WINBUGS (Spiegelhalter et al., 2003). Two chains were run until convergence, and the next 10,000 iterations used for the analysis. Most traits converged after 15,000 iterations, however, biomass, number of spikes, and spike weight required 20,000 iterations to converge.

Among population differentiation for each quantitative trait (Q_{ST}) and associated 95% credible intervals were estimated in WINBUGS1.4 (Spiegelhalter et al., 2003), where the posterior distribution of Q_{ST} was obtained via a variance partition analysis following Bonnin et al (1996):

$$Q_{ST} = \frac{(1 + f) V_B}{(1 + f) V_B + 2 V_W}, \quad [6]$$

where \( f \) is the species-level within-population inbreeding coefficient (i.e. F_{IS}), \( V_B \) is the among-population component of variance (i.e. \( \sigma^2_P \)), and \( V_W \) is the additive genetic component of variance within-populations (i.e. \( \sigma^2_G \)). Since wild barley is a highly selfing species we used an \( f = 1 \).

Pair-wise Q_{ST}^{ij} estimates were obtained using a similar approach with the difference that row and column effects were not included in the linear models in order to make all parameters estimable.

Star-plots for all the populations and the variables under directional selection were constructed in SAS (SAS Institute 2004). Each radius of a star represents one phenotypic variable. These variables, shown clockwise, are days until anthesis, days until flowering, grain weight, spike length, awn length, and number of grains per spike. The length of each radius is proportional to the magnitude of the variable relative to the maximum magnitude of that variable across populations.

**Q_{ST}-F_{ST} comparisons**

Comparisons of population differentiation were conducted at two levels: overall and pair-wise. Overall-estimates of F_{ST} and Q_{ST} for each trait were compared to detect deviation from neutrality of each trait. The Bayesian posterior distribution of the
difference of $F_{ST}$ and $Q_{ST}$ ($\Delta = Q_{ST} - F_{ST}$) was fitted for each trait in WinBUGS 1.4 (Spiegelhalter et al. 2003) and mean values and 95% confidence intervals for $\Delta$ (CI$\Delta$) were obtained. For this purpose, we used the posterior distribution of $Q_{ST}$ as described above and the posterior distribution of $F_{ST}$ by assuming a normal distribution with parameters equal to those obtained from GDA estimation. Each trait was classified as either neutral (i.e. the CI$\Delta$ included zero, $Q_{ST} = F_{ST}$), under divergent selection (i.e. zero was not included in the CI$\Delta$, and $\Delta > 0$, $Q_{ST} > F_{ST}$), or under stabilizing selection (i.e. zero was not included in the CI$\Delta$, $\Delta < 0$, $Q_{ST} < F_{ST}$).

Traits that were classified as neutral by the overall statistics were re-analyzed using the pair-wise approach of Gutierrez et al. (2008) to test for deviations from neutrality. This approach consists of simulating phenotypes with the same among-population covariance structure as the molecular markers, scaled to the phenotypic variance observed within populations. $Q_{ST}^{ij}$ estimated for each pair of simulated populations thus reflects the inter-population covariance structure of the phenotypes expected under neutrality. From the simulated $Q_{ST}^{ij}$, the distribution of $\Delta_{ij} = Q_{ST}^{ij} - F_{ST}^{ij}$ under neutrality is obtained, providing a null model for comparison to the $\Delta_{ij}$ calculated from the actual data. If overall $Q_{ST} = F_{ST}$ and neutrality is in fact true then the variance in observed and simulated $\Delta_{ij}$ should be comparable. Alternatively, if different subsets of populations are subject to disruptive and to stabilizing selection, then dispersion in the observed $\Delta_{ij}$ should be significantly greater than simulated under neutrality. We thus simulated phenotypes and the variance of $\Delta_{ij}$ for all traits that had an overall $Q_{ST} = F_{ST}$, using 10,000 simulations to obtain critical values at the 5% level for the variance of $\Delta_{ij}$ expected under neutrality. If the observed variance in $\Delta_{ij}$ lies outside the critical values then we can conclude that the use of overall $Q_{ST}$ and $F_{ST}$ has led to the trait being misclassified as neutral. Furthermore, a significantly larger variance in $\Delta_{ij}$ is consistent with selection mosaics of disruptive and stabilizing selection. These tests and simulations were performed in GENALEX (Peakall & Smouse 2006) and R (Ihaka & Gentleman 1996).
For the traits that were misclassified as neutral, Bonferroni corrected p-values were used to test deviations from neutrality in each pair (Gutierrez et al., 2008). Color-coded matrices were constructed such that pairs of populations were colored according to the type of selection acting on it: white represents a pair under stabilizing selection; black represents a pair under divergent selection; and grey represent a pair for which neutrality was not rejected.

Traits under selection were further studied by Mantel test of correlations between matrices of $F_{ST}^{ij}$ and $Q_{ST}^{ij}$ estimates (Palo et al., 2003) using the program IBD (Bohonak, 2002), with significance based on 1,000 permutations of the data.

**Results**

$F_{ST}$ estimation

The average number of individuals successfully genotyped per population, averaged across markers, ranged from 6.2 to 46.9 (mean 11.2). The average number of alleles per polymorphic locus was high (3.7), with a range from two to seven alleles per locus. This number is slightly biased downward because of scoring technique. Due to the resolution of metaphor-agarose gels, and in order to avoid errors, some alleles that differed by only a few base-pairs were scored as one allele. Combining alleles reduced the information but also decreased the error. The average gene diversity (expected heterozygosity) within populations ranged from 0.163 to 0.502 (mean 0.369). The observed heterozygosity per locus was low ranging from 0 to 0.054 (mean 0.015). Therefore, the mean inbreeding coefficient ($F_{IS}$) was high (0.985). Among-population differentiation in marker frequencies was substantial and significantly greater than zero $F_{ST}$ (0.301). Pair-wise estimates of population differentiation ($F_{ST}^{ij}$) ranged from 0 to 0.586 with an average of 0.264.

No significant results were found for any of the models testing the relationship between $F_{ST}$ and geographical distance. Therefore, no isolation by distance model can be inferred.
**Q<sub>ST</sub> estimation**

Quantitative traits displayed significant among population variance (data not shown). Significant among-population differentiation was found for all traits (Q<sub>ST</sub> > 0), with mean Q<sub>ST</sub> ranging from 0.191 for flag leaf length to 0.669 for awn length (Table 2). Standard errors of Q<sub>ST</sub> were generally low (0.074-0.114, Table 2). Pair-wise Q<sub>ST</sub><sup>ij</sup> estimates for all traits ranged from 0.003 to 0.999, while the average (by trait) Q<sub>ST</sub><sup>ij</sup> ranged from 0.232 to 0.615.

**Comparison of Q<sub>ST</sub> and F<sub>ST</sub>**

Comparisons of overall Q<sub>ST</sub> and F<sub>ST</sub> estimates failed to reject the null hypothesis of neutrality in seven traits (Q<sub>ST</sub> = F<sub>ST</sub>: number of tillers, crown rust, biomass, number of spikes, spike weight, flag leaf length, and flag leaf width; Table 2), while six traits were found to be under divergent selection (Q<sub>ST</sub> > F<sub>ST</sub>: days until anthesis, days until flowering, grain weight, spike length, awn length, and number of grains). No trait was found to be under stabilizing selection.

When pair-wise estimates of Q<sub>ST</sub><sup>ij</sup> and F<sub>ST</sub><sup>ij</sup> were examined, the null hypothesis of neutrality was rejected for two of the traits classified as neutral by the contrast of overall statistics. Specifically, biomass and flag leaf width had an observed variance of Δ<sub>ij</sub> significantly larger than expected under neutrality (Table 2). Some pairs of populations were under stabilizing selection for biomass (e.g., P2, P7, P8, P11, P23 and P24 with each other, among others; Figure 2) and some for flag leaf width (e.g., P3 with P10, P12, P13, P15, P16, P18, 19, P20, and P21, among others). Additionally, some pairs were under divergent selection for biomass (e.g., P5 and P18 with several populations), and for flag leaf width (e.g., P3 with several populations). We did not find deviations from neutrality for other traits that were classified as neutral by the overall statistics.

Mantel test correlations between Q<sub>ST</sub><sup>ij</sup> and F<sub>ST</sub><sup>ij</sup> indicate that selection and drift are acting in the same direction for awn length (r = 0.267, p = 0.041, b = 0.382, Table 2), indicating that for this trait, mechanisms that cause populations to differentiate due to
restricted gene flow and mechanisms that cause them to differentiate due to selection pressures tended to reinforce each other.

**Discussion**

Using information from both overall and pair-wise statistics provides more information in some cases, and a correct assessment of the neutrality in others. If a trait is under selection (i.e. $Q_{ST} > F_{ST}$, or $Q_{ST} < F_{ST}$), the use of pair-wise comparisons provides more information. The interpretations of the overall results do not change, and the use of both statistics provides a better understanding of the evolutionary process (i.e. selection is acting in the same general direction as gene flow or not). On the other hand, when a trait is classified as neutral, pair-wise comparisons allow to detect misclassification. If $Q_{ST} = F_{ST}$ and the variance of the observed $\Delta_{ij}$ is not larger than the variance of the expected $\Delta_{ij}$ under neutrality, then the trait is truly neutral. However, if the observed variance of $\Delta_{ij}$ is larger than the expected variance of $\Delta_{ij}$ under neutrality, then the trait is not neutral, and there might be pairs of populations under stabilizing and pairs under divergent selection. This means that the overall “neutrality” of the trait does not represent what is actually happening at the population level, and that on average all effects tend to cancel each other.

**Population Structure**

Significant structure was found for both molecular markers and quantitative traits. High population structure for markers ($F_{ST} = 0.3$) is consistent with other selfing species (Hamrick & Godt 1990; Schoen & Brown 1991), and with that found in other wild barley studies using SSRs (Baek et al. 2003; Turpeinen et al. 2001), AFLPs (Ozkan et al. 2005), RAPDs (Baum et al. 1997), and isozymes (Zhang et al. 1993). On the other hand, some studies found either slightly higher differentiation ($F_{ST} = 0.4$) with RAPDs (Dawson et al. 1993), isozymes (Chalmers et al. 1992; Nevo et al. 1986), and RFLPs (Zhang et al. 1993); or lower differentiation ($F_{ST} = 0.1$) with RAPDs (Volis et al. 2002).
While SSR markers are very informative and therefore provide high statistical power (Goudet et al. 1996; Hedrick 1999), some questions about their use for population studies have been raised (Estoup & Courten 1999; Estoup et al. 2002; O’Reilly et al. 2004). Microsatellites evolve according to a stepwise mutational model (Di Rienzo et al. 1994; Weber & Wong 1993) creating new alleles by the addition or deletion of only one or a few repeat units. If mutation rates are high, the creation of new alleles in the population can counteract the effect of drift that tends to eliminate rare alleles (Slatkin 1993). This will cause underestimation of $F_{ST}$ or $R_{ST}$ statistics (Estoup & Courten 1999). Some studies have shown a negative correlation between $F_{ST}$ and allelic richness (O’Reilly et al. 2004), heterozygosity (O’Reilly et al. 2004), and number of repeat units (Brohede et al. 2002; Ellegren 2000; Petit et al. 2005). We did not find correlations among any of those statistics in our study; therefore, we have no a priori reason to believe that $F_{ST}$ will not correctly estimate population differentiation at the molecular level.

High population structure was also found for quantitative traits with $Q_{ST}$ significantly greater than zero ($p < 0.05$) for all the traits ($Q_{ST} = 0.03-0.67$). However, a wide range of $Q_{ST}$ values was found for the different traits causing some traits to be neutral and some to be under divergent selection, consistent with other studies (Gomez-Mestre & Tejedo 2004; Hamrick 2004; Mc Kay & Latta 2002; Merilä & Crnokrak 2001; Morgan et al. 2001; Palo et al. 2003; Podolsky & Holtsford 1995; Spitze 1993; Steinger et al. 2002 among others).

$F_{ST}$ and $Q_{ST}$ comparison

Different traits might be under different selection pressures, and in consequence have different $Q_{ST}$-$F_{ST}$ relationships (see Mc Kay & Latta 2002 for a review). However, not all the traits are independent and quantitative genetics theory shows that functionally and/or developmentally related traits will evolve as a coordinated unit (Cheverud 1982; 1995; Lande 1979; 1980; 1984; Marroig & Cheverud 2001). For instance, life-history traits are predicted to be under strong directional selection, while morphological traits are expected to be under weak or stabilizing selection (Merilä & Sheldon 1999; Rieseberg et al. 2002).
We found seven traits that could be classified as neutral traits based on overall statistics, and six traits under divergent selection pressures. Many studies report either more traits under divergent selection or the average $Q_{ST}$ across traits larger than $F_{ST}$ (McKay & Latta 2002; Merilä & Crnokrak 2001; Morgan et al. 2001; Palo et al. 2003). Some studies have also shown cases of stabilizing selection (Edmans & Harrison 2003). The traits with significant $Q_{ST} > F_{ST}$, and therefore under divergent selection were days until anthesis, days until flowering, grain weight, spike length, awns length, and number of grains per spike. All of these traits other than awn length are considered life-history traits, and are expected to be under strong directional selection (Merilä & Sheldon 1999; Rieseberg et al. 2002).

Life history traits of two types, maturity (i.e. days until anthesis and days until flowering), and reproductive (i.e. grain weight, spike length, and number of grains) were under strong directional selection. Awn length was also under strong directional selection. It is not a life-history type of trait, but selection could be explained by the correlation between $Q_{ST}^{ij}$ and $F_{ST}^{ij}$. Selection and gene flow act in concert, having more phenotypic differentiation among populations that also experience more differentiation in their markers. In situations where there is a clear geographical pattern causing population structure at molecular markers, we would expect several traits with significant correlations between $Q_{ST}^{ij}$ and $F_{ST}^{ij}$. This situation would not imply a causality-relationship between the structures at the different levels but rather would probably represent a geographical structure at the quantitative traits caused by different environmental conditions that just happened to be associated with geographical location. However, molecular marker differentiation in our study is not explained by isolation by distance models. Therefore, we propose that other mechanisms despite clinal variation could also generate this correlation. High population differentiation among populations that are close was found in our study, and is consistent with other studies of *H. spontaneum* in Israel (Nevo et al. 1986; 1983; 1981). Therefore, it was surprising to find an association between $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ in this case with an idiosyncratic structure at the molecular marker. Further study of the association is necessary to reveal the mechanisms
that explain the correlation. Although, the Mantel test p-value of 0.041 becomes marginal with any multiple testing correction. Some studies use the correlation of $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ to explain the selection patterns (Palo et al., 2003). However, they do formalize the results in terms of hypothesis to be tested. We propose the use of Mantel test correlations among pair-wise $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ estimates to detect cases where gene flow and selection pressures are geographically correlated (Figure 3). For example, if the spatial dimension of gene flow and population differentiation is influenced by geographical distance (i.e., isolation by distance) and disruptive selection on phenotypes accentuates this geographical pattern of differentiation, then the effects of gene flow and selection on quantitative genetic structure can be considered as acting in concert. This is common in nature with clinal variation. On the other hand, there may be situations where selection pressures are not correlated with patterns of gene flow and therefore no relationship between $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ is found. Palo et al. (2001) used Mantel test correlations between $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ to understand the forces that drive selection. However, other studies use pair-wise $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ correlations with goals different from this study (Morgan et al., 2001; Steinger et al., 2002). The logic behind those studies was that if there is a significant correlation between $Q_{ST}^{ij}$ and $F_{ST}^{ij}$, there is no need to use both to study population structure. This is the same as proposed by Crnokrak and Merila (2002) and McKay and Latta (2002) but for inter-species comparisons. We are proposing the combined use of molecular markers and quantitative traits to understand the mechanisms that drive population differentiation. Using both $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ would allow us to detect situations where gene flow and selection act in concert and situations where there is no association between gene flow and selection pressures. Understanding the mechanisms that drive population structure is vital for devising conservation programs.

A combination of divergent and stabilizing selection was detected for the traits that were classified as neutral with overall $Q_{ST}$ and $F_{ST}$ and non-neutral with pair-wise $Q_{ST}^{ij}$ and $F_{ST}^{ij}$. To test which specific pairs were under divergent or stabilizing selection we used a two-step approach. First, we tested whether the trait was neutral or not with pair-wise comparisons of $Q_{ST}^{ij}$ and $F_{ST}^{ij}$. Traits where neutrality was rejected were further
studied. This would be similar to having a Fisher’s protected test. Second, we used Bonferroni corrected p-values so that the inference can be simultaneously valid in all pairs with at least a 95% confidence. This is a conservative test; a precise test should be constructed such that the non-independence of pair-wise comparisons is taken into account.

The following patterns were detected for pair-wise relationship in the traits that were classified as under divergent selection by overall statistics. First, selection and drift are acting in the same direction in awn length (Table 2). Mechanisms that cause populations to differentiate due to restricted gene flow and mechanisms that cause them to differentiate due to selection pressures tend to reinforce each other. Second, populations Mehola (P11), Rosh Pinna (P18) and Sede Boqer (P19) tend to be the populations that drive \( Q_{ST} > F_{ST} \) (11, 18, 19, Figure 4). Third, there are two sets of populations that tend to have \( Q_{ST}^{ij} < F_{ST}^{ij} \) for the comparisons within the group, and \( Q_{ST}^{ij} > F_{ST}^{ij} \) for the comparisons between the group populations and outside populations (G1 and G2, Figure 4).

A \( Q_{ST} > F_{ST} \) could be explained by two different mechanisms; all populations are diverging, or some populations are diverging and cause the overall \( Q_{ST} > F_{ST} \) trend. If the equal-divergence hypothesis is true, all populations contribute to \( Q_{ST}^{ij} > F_{ST}^{ij} \), and therefore we would find all the populations shifted around the expected \( Q_{ST}^{ij} - F_{ST}^{ij} \) line (i.e. the one-to-one relationship plus the difference in \( Q_{ST}^{ij} - F_{ST}^{ij} \)). However, if there are some driving populations, we would find some populations (i.e. the driving populations) consistently above the expected \( Q_{ST}^{ij} - F_{ST}^{ij} \) line. In the absence of a formal test along these lines, we used the following ad hoc approach. We identified driving populations by counting the number of times a population had a pair with \( Q_{ST}^{ij} > F_{ST}^{ij} \) in each trait. The populations with more than eleven of their pairs (more than half of the total pairs) with \( Q_{ST}^{ij} > F_{ST}^{ij} \) in at least four of the traits (more than half of the traits) under selection were considered driving populations. We found the following driving populations: Mehola (P11), Rosh Pinna (P18), and Sede Boqer (P19, Figure 4). Mehola is situated below sea level and has a steppic-marginal climate with low annual precipitation, high temperature
of the hottest month, day-night temperature difference, and evaporation, and is situated on alluvium soils (Nevo et al., 1979, 1984). The plants in this population tend to be early maturing with high grain weights and low spike length and number of grains. Rosh Pinna is an average population in terms of climatic characteristics (Nevo et al., 1979, 1984). However, their individuals have unique characteristics: very low grain weight, long spikes, and short awns. Sede Boqer is a desert population with very low rainfall, high evaporation, low humidity, and over loess soils (Nevo et al., 1979, 1984). Plants in this population are very early, have low grain yields and number of grains and have short spikes and awns. A formal test for detecting driving populations would need to be developed such that the null hypothesis would be equal-divergence and an alternative hypothesis would be driving-populations. The test could include a simulation of populations such that they preserve the observed $Q_{ST} - F_{ST}$ relationship while creating constant $Q_{ST}^{ij} - F_{ST}^{ij}$ relationships. Larger variances of $Q_{ST}^{ij} - F_{ST}^{ij}$ than expected under the null of equal-divergence would indicate that there are driving populations. The detection of the specific populations that drive the divergence could be accomplished by first rejecting the hypothesis of equal-divergence, and then using conservative multiple test comparison accounting for the number of times a population is significantly above the $Q_{ST}^{ij} - F_{ST}^{ij}$ relationship expected under the null.

Additionally, driving populations could be present in the form of groups of populations that tend to show stabilizing selection ($Q_{ST}^{ij} < F_{ST}^{ij}$) within groups, and divergent selection ($Q_{ST}^{ij} > F_{ST}^{ij}$) between groups. We used the same ad hoc procedure described above, but to identify specific pairs that consistently had a $Q_{ST}^{ij} < F_{ST}^{ij}$. The first group (G1) includes populations Afiq (P2), Caesarean (P7), Damon (P8), Herzliya (P9), Maalot (P10), Mt. Hermon (P12), Mt. Meron (P13), Nahal Oren (P14), and Neve Yaar (P15, Figure 4). All of these populations have large values for all the variables under selection (Figure 1). A hierarchical cluster algorithm grouping populations by their performance on the variables under selection clustered seven of these populations (P2, P7, P8, P10, P12, P13, and P15) into a single group (data not shown). The second group (G2) included populations Shechem (P20), Tabigha Niab (P21), Tabigha Transect (P22),
Talpiyyot (P23) and Wadi Qilt (P24, Figure 4). These populations did not have a clear pattern of consistent phenotypes and were not clustered in the same group when studying the phenotypic traits under selection with a hierarchical cluster algorithm (data not shown). While the ad hoc approach improved the description provided of inter-population evolutionary forces, a more formal, statistically justified test the hypothesis that a subset of populations drive the observation of $Q_{ST} > F_{ST}$ is still needed.

In summary, our results suggest that overall $F_{ST}$ and $Q_{ST}$ comparisons are useful for studying the evolutionary history of quantitative traits, however, they do not adequately describe the evolutionary forces acting at the inter-population levels. Pair-wise comparisons of $Q_{ST}^{ij}$ and $F_{ST}^{ij}$, together with the overall estimations, on the other hand, are better suited to explain these patterns. Furthermore, the combined studies of overall and pair-wise comparisons allow for a better understanding of the evolution of traits detecting diverging and stabilizing selection at the inter-population level where overall statistics would fail to find deviations from neutrality because of effects cancelling out, and detecting mechanisms that explain the population structure when overall divergent or stabilizing selection was found.

Acknowledgments

The authors wish to thank the research group at INIA La Estanzuela in Uruguay for support with field evaluations and the Monsanto Company for in-kind support for the genotyping of wild barley. L.G. was supported by a Natural Systems Agriculture Fellowship from The Land Institute during this study.

References


Gutterman Y, Gozlan S (1998) Amounts of winter or summer rain triggering germination and 'the point of no return' of seedling desiccation tolerance, of some Hordeum spontaneum local ecotypes in Israel. Plant and Soil 204, 223-234.


Kuittinen H, Mattila A, Savolainen O (1997) Genetic variation at marker loci and in quantitative traits in natural populations of Arabidopsis thaliana. Hereditas 2, 144-152.


**List of Figures**

**Figure 1.** Map of Israel with population location and relative performance for the variables under selection.

**Figure 2.** Pair-wise $\Delta_{ij}$ matrix for the variables that were wrongly classified as neutral: biomass (below diagonal) and flag leaf width (above diagonal). Each pair was color coded as either $Q_{ST}^{ij} > F_{ST}^{ij}$ (p < 0.0001, black), $Q_{ST}^{ij} < F_{ST}^{ij}$ (p < 0.0001, white) or $Q_{ST}^{ij} = F_{ST}^{ij}$ (not-significant, gray).

**Figure 3.** A theoretical model for the interpretations of overall and pair-wise $Q_{ST}^{ij}$-$F_{ST}^{ij}$ comparisons.

**Figure 4.** Pair-wise $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ for the variables under selection: a) days until anthesis, b) days until flowering, c) grain weight, d) spike length, e) awn length, and f) number of grains per spike. Populations with most of their pairs with $Q_{ST}^{ij} > F_{ST}^{ij}$ for all six variables are plotted as numbers where the center of the number represent the location of each data point. Groups of populations with $Q_{ST}^{ij} < F_{ST}^{ij}$ within group are plotted as solid circles or squares. The remaining pairs of populations are represented as dots.
Table 1. Description of populations sampled: Population (P), Latitude (LAT, decimals), Longitude (LON, decimals), Altitude (ALT, meters), Mean Annual Rainfall (MAR, mm.), Mean Annual Temperature (MAT, °C), Mean Hottest Month Temperature (MHMT, °C), Mean Coldest Month Temperature (MCMT, °C), Soil Type (SOIL), and Ecological Environment (CLIMATE).

<table>
<thead>
<tr>
<th>P</th>
<th>Name</th>
<th>LAT</th>
<th>LON</th>
<th>ALT</th>
<th>MAR</th>
<th>MAT</th>
<th>MHMT</th>
<th>MCMT</th>
<th>SOIL</th>
<th>CLIMATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Afiq</td>
<td>32.78</td>
<td>35.70</td>
<td>325</td>
<td>455</td>
<td>21</td>
<td>27</td>
<td>11</td>
<td>Basalt</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Akhziv</td>
<td>33.05</td>
<td>35.10</td>
<td>10</td>
<td>620</td>
<td>20</td>
<td>26</td>
<td>12</td>
<td>Alluvium</td>
<td>Coastal plain</td>
</tr>
<tr>
<td>3</td>
<td>Ashqelon</td>
<td>31.63</td>
<td>34.60</td>
<td>50</td>
<td>420</td>
<td>20</td>
<td>27</td>
<td>14</td>
<td>Sandy Loam</td>
<td>Coastal plain</td>
</tr>
<tr>
<td>4</td>
<td>Bar Giyyora</td>
<td>31.72</td>
<td>35.08</td>
<td>760</td>
<td>540</td>
<td>17</td>
<td>26</td>
<td>10</td>
<td>Terra Rossa</td>
<td>Med. Mountain</td>
</tr>
<tr>
<td>5</td>
<td>Bet Shean</td>
<td>32.50</td>
<td>35.50</td>
<td>-120</td>
<td>290</td>
<td>23</td>
<td>30</td>
<td>13</td>
<td>Rendizia</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Caesarea</td>
<td>32.50</td>
<td>34.90</td>
<td>10</td>
<td>540</td>
<td>20</td>
<td>26</td>
<td>13</td>
<td>Sandy Loam</td>
<td>Coastal plain</td>
</tr>
<tr>
<td>7</td>
<td>Damon</td>
<td>32.73</td>
<td>35.00</td>
<td>425</td>
<td>686</td>
<td>19</td>
<td>24</td>
<td>11</td>
<td>Terra Rossa</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Herzliyya</td>
<td>32.17</td>
<td>34.80</td>
<td>25</td>
<td>530</td>
<td>20</td>
<td>26</td>
<td>13</td>
<td>Sandy Loam</td>
<td>Coastal plain</td>
</tr>
<tr>
<td>9</td>
<td>Ma'alot</td>
<td>33.00</td>
<td>35.27</td>
<td>500</td>
<td>790</td>
<td>17</td>
<td>23</td>
<td>8</td>
<td>Rendizia</td>
<td>Med. Mountain</td>
</tr>
<tr>
<td>10</td>
<td>Mehola</td>
<td>32.13</td>
<td>35.48</td>
<td>-150</td>
<td>270</td>
<td>22</td>
<td>30</td>
<td>13</td>
<td>Alluvium</td>
<td>Steppic/marginal</td>
</tr>
<tr>
<td>11</td>
<td>Mt. Hermon</td>
<td>33.28</td>
<td>35.75</td>
<td>1530</td>
<td>1600</td>
<td>11</td>
<td>20</td>
<td>1</td>
<td>Terra Rossa</td>
<td>Steppic/marginal</td>
</tr>
<tr>
<td>12</td>
<td>Mt. Meron</td>
<td>33.05</td>
<td>35.40</td>
<td>1150</td>
<td>1010</td>
<td>14</td>
<td>22</td>
<td>6</td>
<td>Terra Rossa</td>
<td>Med. Mountain</td>
</tr>
<tr>
<td>13</td>
<td>Nahal Oren (N and S)</td>
<td>32.43</td>
<td>35.02</td>
<td>75</td>
<td>690</td>
<td>19</td>
<td>24</td>
<td>11</td>
<td>-</td>
<td>Med. Mountain</td>
</tr>
<tr>
<td>14</td>
<td>Neve Yaar</td>
<td>32.44</td>
<td>38.10</td>
<td>100</td>
<td>600</td>
<td>20</td>
<td>27</td>
<td>11</td>
<td>Deep rendzina</td>
<td>Med. Mountain</td>
</tr>
<tr>
<td>15</td>
<td>Ovedat</td>
<td>30.80</td>
<td>34.75</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>Revivim</td>
<td>31.02</td>
<td>34.75</td>
<td>320</td>
<td>130</td>
<td>20</td>
<td>27</td>
<td>10</td>
<td>Alluvium</td>
<td>Dessert</td>
</tr>
<tr>
<td>17</td>
<td>Rosh Pinna</td>
<td>32.95</td>
<td>35.52</td>
<td>700</td>
<td>697</td>
<td>19</td>
<td>24</td>
<td>8</td>
<td>Terra Rossa</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Sede Boqer</td>
<td>30.87</td>
<td>34.78</td>
<td>450</td>
<td>90</td>
<td>19</td>
<td>26</td>
<td>9</td>
<td>Loess</td>
<td>Dessert</td>
</tr>
<tr>
<td>19</td>
<td>Shechem</td>
<td>32.23</td>
<td>35.23</td>
<td>400</td>
<td>620</td>
<td>18</td>
<td>24</td>
<td>9</td>
<td>Rendizia</td>
<td>Med. Mountain</td>
</tr>
<tr>
<td>20</td>
<td>Tabigha Niab (A126)</td>
<td>32.90</td>
<td>35.53</td>
<td>0</td>
<td>440</td>
<td>24</td>
<td>32</td>
<td>15</td>
<td>Basalt</td>
<td>Steppic/marginal</td>
</tr>
<tr>
<td>21</td>
<td>Tabigha Transect</td>
<td>32.90</td>
<td>35.53</td>
<td>0</td>
<td>436</td>
<td>0</td>
<td>32</td>
<td>15</td>
<td>Basalt</td>
<td>Steppic/marginal</td>
</tr>
<tr>
<td>22</td>
<td>Talpiyyot</td>
<td>31.75</td>
<td>35.23</td>
<td>800</td>
<td>486</td>
<td>18</td>
<td>24</td>
<td>9</td>
<td>Rendizia</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Wadi Qilt</td>
<td>31.83</td>
<td>35.38</td>
<td>50</td>
<td>170</td>
<td>23</td>
<td>30</td>
<td>14</td>
<td>Alluvium</td>
<td>Dessert</td>
</tr>
</tbody>
</table>
Table 2. Overall and pair-wise estimates of $Q_{ST}$ and $F_{ST}$. Overall estimates: mean of the full posterior distribution of $Q_{ST}$ (all significantly greater than zero, standard error in parenthesis), and the significance of the difference between $Q_{ST}$ and $F_{ST}$ ($Q_{ST} \neq F_{ST}$). Pair-wise estimates: variance of $\Delta_{ij} = Q_{ST}^{ij} - F_{ST}^{ij}$ (VD) for traits indicated by the overall contrast to be neutral, and Mantel test correlation ($r_{(F_{ST}, Q_{ST})}$) for traits indicated by the overall contrast to be under selection.

<table>
<thead>
<tr>
<th>Traits</th>
<th>$Q_{ST}$ (s.e.)</th>
<th>$Q_{ST} \neq F_{ST}$</th>
<th>Pair-wise</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Neutral Traits</strong></td>
<td></td>
<td></td>
<td>VD</td>
</tr>
<tr>
<td>1 Number of Tillers</td>
<td>0.285 (0.101)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>0.078</td>
</tr>
<tr>
<td>3 Crown Rust</td>
<td>0.277 (0.093)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>0.061</td>
</tr>
<tr>
<td>6 Biomass</td>
<td>0.323 (0.095)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>0.135 *</td>
</tr>
<tr>
<td>7 Number of Spikes</td>
<td>0.405 (0.114)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>0.083</td>
</tr>
<tr>
<td>9 Spike Weight</td>
<td>0.382 (0.096)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>0.058</td>
</tr>
<tr>
<td>10 Flag Leaf Length</td>
<td>0.191 (0.088)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>0.065</td>
</tr>
<tr>
<td>11 Flag Leaf Width</td>
<td>0.356 (0.097)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>0.134 *</td>
</tr>
<tr>
<td><strong>Traits under Selection</strong></td>
<td></td>
<td></td>
<td>$r_{(F_{ST}, Q_{ST})}$</td>
</tr>
<tr>
<td>4 Days until Anthesis</td>
<td>0.589 (0.088)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td>0.100</td>
</tr>
<tr>
<td>5 Days until Flowering</td>
<td>0.575 (0.100)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td>0.082</td>
</tr>
<tr>
<td>8 Grain Weight</td>
<td>0.503 (0.099)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td>0.053</td>
</tr>
<tr>
<td>12 Spike Length</td>
<td>0.592 (0.095)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td>0.144</td>
</tr>
<tr>
<td>13 Awn Length</td>
<td>0.669 (0.074)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td>0.267 *</td>
</tr>
<tr>
<td>14 Number of Grains</td>
<td>0.561 (0.093)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td>0.127</td>
</tr>
</tbody>
</table>

$(1) Q_{ST} = F_{ST}$ indicates that there is not a significant ($p = 0.05$) difference between $Q_{ST}$ and $F_{ST}$.

$Q_{ST} > F_{ST}$ indicates that $Q_{ST}$ is significantly greater than $F_{ST}$. $Q_{ST} < F_{ST}$ indicates $Q_{ST}$ is significantly lower than $F_{ST}$.

* Indicates significant ($p < 0.05$) values of VD or $r_{(F_{ST}, Q_{ST})}$. 
Figure 1. Populations map and performance of the traits under selection

Populations:
2. Afiq
3. Akhziv
4. Ashqelon
5. Bar Giyyora
6. Bet Shean
7. Caesarea
8. Damon
9. Herzliyya
10. Maalot
11. Mehola
12. Mt. Hermon
13. Mt. Meron
14. Nahal Oren (N and S)
15. Neve Yaar
16. Ovedat
17. Revivim
18. Rosh Pinna
19. Sede Boqer
20. Shechem
21. Tabigha Niab (A126)
22. Tabigha Transect
23. Talpiyyot
24. Wadi Qilt

Star-plots for each population represent the relative performance of the six variables that are under directional selection in the study. Each radius represents one variable (shown below). The length of each radius is proportional to the magnitude of the variable relative to the maximum magnitude across populations.
<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Figure 2.** Pair-wise $\Delta_{ij}$ matrix for non-neutral traits
Overall Q\textsubscript{ST}-F\textsubscript{ST} comparisons

<table>
<thead>
<tr>
<th>Overall Q\textsubscript{ST}-F\textsubscript{ST}</th>
<th>Significant relationship among pair-wise Q\textsubscript{ST}\textsuperscript{ij} and F\textsubscript{ST}\textsuperscript{ij}</th>
<th>No relationship among pair-wise Q\textsubscript{ST}\textsuperscript{ij} and F\textsubscript{ST}\textsuperscript{ij}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q\textsubscript{ST} &gt; F\textsubscript{ST} (A)</td>
<td>Overall divergent selection. Greater divergence of traits coincides with reduced gene flow (e.g. isolation by distance with clinal variation)</td>
<td>Overall divergent selection. No relationship between selection regime divergence and gene flow.</td>
</tr>
<tr>
<td>Q\textsubscript{ST} &lt; F\textsubscript{ST} (C)</td>
<td>Overall stabilizing selection. Less divergence of traits coincides with reduced gene flow.</td>
<td>Overall stabilizing selection. No relationship between selection regime divergence and gene flow.</td>
</tr>
</tbody>
</table>

**Figure 3.** Interpretation of overall and pair-wise Q\textsubscript{ST}\textsuperscript{ij}-F\textsubscript{ST}\textsuperscript{ij} comparisons
Figure 4. Pair-wise $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ for traits under selection
CHAPTER IV. MORPHOLOGICAL GENETIC DIVERSITY OF WORLDWIDE BARLEY AND MEGA-TARGETS OF SELECTION

A paper accepted in *Crop Science*

Lucía Gutiérrez,* John D. Nason, and Jean-Luc Jannink

**Abstract**

Germplasm exchange is essential for advancing genetic gain in a breeding program. Two aspects of breeding programs are relevant for germplasm exchange: the amount of genetic diversity within programs and the identification of breeding programs with similar breeding objectives and environments of selection (i.e., mega-targets of selection). The objective of this study was to quantify worldwide genotypic diversity of barley for morphological traits, and to use that information to aid in germplasm exchange between breeding programs. We evaluated 20 morphological traits in 353 genotypes of barley (*Hordeum vulgare* L.) from 23 private and public breeding programs distributed worldwide. We found significant amounts of genetic diversity for all traits, but differences in diversity among breeding programs for only seven traits. We identified breeding programs with high diversity (i.e., two-row Western Australia, Canada Saskatchewan, and Sweden), and low diversity (i.e., two-row Croatia spring, Germany, Busch Ag Res, and six-row Croatia winter, Idaho and Minnesota). We developed a methodology that produces groups of breeding programs with similar performance and response to the environments. We used the methodology to group the 23 breeding programs of barley into sets that might benefit most from germplasm exchange. The

---


Running Head: Diversity of Worldwide Barley and Mega-Targets of Selection

Keywords: Plant Genetic Resources, Germplasm Exchange, Characterization

Abbreviations: ME: Mega-Environments, MTS: Mega-Targets of Selection
identification of compatible breeding programs for germplasm exchange could be of significant relevance for improving genetic gains in breeding programs.

**Introduction**

Genetic diversity is essential in a breeding program for two reasons; as an insurance against unforeseeable changes in the environment (Gepts, 2006) and to maintain genetic progress (Gepts, 2006; Rasmusson, 2001). Incorporating every single allele just in case it may be needed in the future is neither possible nor desirable. Trying to broaden the diversity of a breeding program to account for the unknown will only slow genetic progress because the selection intensity on the traits of interest would be very small. Germplasm banks can maintain genetic diversity that is not immediately need in the breeding program. When new variation is needed because the environment has changed (e.g., a new disease appears), specific genetic variation can be brought into the breeding program. Diversity in a breeding context is also needed to maintain genetic progress (Gepts, 2006; Rasmusson, 2001). Since wide crosses usually do not recover the high performance of elite genotypes, narrower good by good crosses are needed. A desirable genotype would therefore be an elite line with new alleles at the loci of interest. The variation required can be created de novo (i.e., from mutation), by epistatic interactions, or brought in from new elite genepools (Rasmusson, 2001). For this purpose, assessment of diversity and performance of elite germplasm is needed.

Germplasm exchange of elite genotypes among breeding programs is an effective way to increase genetic gain. However, not all the elite genotypes will perform well in all the environments. Genotypes are adapted to the environment in which they were selected, and perform best under those conditions (Simmonds, 1991). Furthermore, breeding objectives and the environmental conditions of genotype evaluation shape those adaptations (Atlin et al., 2001; Ceccarelli, 1994). We call the combination of those factors targets of selection. However, it is not easy to identify targets of selection because of the multiple objectives and several environmental conditions of genotype evaluation.
Therefore, breeders need data-driven methods to identify compatible programs for germplasm exchange. Broadly speaking, programs will be compatible if they have the same targets of selection (i.e., if they belong to the same mega-target of selection – MTS). Two aspects of genotype evaluation are relevant in the identification of MTS, the first being genotypic performance. If genotypes are evaluated in the target environment, genotypic performance is an effective way of choosing compatible germplasm. However, it is not possible for a breeding program to evaluate every single genotype. Therefore, a method to aid in germplasm exchange when evaluations are conducted outside the target environment is needed. In the non-target environment, grouping genotypes by performance alone is not enough to identify compatible breeding programs. Genotypes could perform poorly for different reasons. For example, one set of genotypes could be limited because of disease pressure, while the other could be limited by photoperiod conditions. Germplasm exchange among those breeding programs would probably not provide an advantage. Therefore, a second key aspect in the identification of targets of selections is the response of a genotype to change in the environment. For example genotypes that produce similar yields under dry and non-dry conditions would be assigned to a group, while genotypes that perform well in non-dry conditions but poorly in dry conditions would be assigned to a different group.

Mega targets of selection are analogous to mega-environments (ME). ME were first defined as environments with similar ‘biotic and abiotic stresses, cropping system requirements, consumer preferences, and volume of production’ (Braun et al., 1996). The concept was later re-defined as environments that caused genotypes to perform similarly (Gauch and Zobel, 1996; 1997), and therefore little genotype by environment interaction is expected within ME (Yan et al., 2000). Furthermore, ME were defined in a multi-environment trial context as groups of environments that produce the same rank of genotypes, and where evaluation of genotypes in more than one environment of a ME would produce redundant information (Yan et al., 2000). Following the same principles, in MTS, groups of breeding programs are formed such that exchanging germplasm material within a group will produce genotypes that are well adapted and respond
similarly to the new environmental conditions. To belong to a MTS, breeding programs would have similar mean performance (i.e., the same ‘volume of production’ in Braun’s (1996) definition), and respond similarly to new environments (i.e. ‘no genotype by environment interaction’ in Gauch and Zobel’s (1996) definition).

Barley is a good model species for the combined study of diversity and targets of selection. It was one of the first crops to be domesticated 10,000 years ago (Harlan, 1971), and has undergone intensive breeding for more than one century (Hintum, 1994). Despite the long history of breeding, barley is still a highly diverse crop, and is adapted to a range of environmental conditions (Hayes et al., 2003), including tolerance to cold, draught, alkali, and salinity. Furthermore, breeding efforts have produced systematic genetic gain in barley for several traits (Gymer, 1981) despite the common use of elite parents that created narrow gene pools (Rasmusson and Phillips, 1997). Therefore, different targets of selection are expected, and breeders would be benefit from the identification of breeding programs between which germplasm exchange would be advantageous.

The aim of this study was to quantify worldwide genotypic diversity of barley for morphological traits and to use that information to aid in germplasm exchange between breeding programs. The three specific objectives were: first, to characterize genetic diversity at morphological traits of advanced inbred lines of barley; second, to describe the diversity of the breeding programs for those traits; and third, to develop a data-driven method identifying MTS in barley to group programs likely to be compatible for germplasm exchange. The MTS would help in germplasm exchange between breeders, where optimal exchange would be among breeding programs of the same MTS.

Materials and Methods

Materials

A total of 353 inbred lines of barley from 23 private and public breeding programs were evaluated. Each breeder responsible for a breeding program was asked to provide 20 advanced inbred lines or recently released cultivars of barley that represented current
diversity in their program. Two-row and six-row types were treated separately, and if a breeding program included both, separate samples were asked of each type. The programs that provided seed were from the United States (Washington State University; University of Minnesota; 2-row and 6-row programs of North Dakota State University; 2-row and 6-row programs of USDA-ARS-Aberdeen-Idaho; 2-row, 6-row and international programs of Busch Agricultural Resources), Canada (University of Saskatchewan; Alberta Agriculture, Food and Rural Development), Europe (Saatzucht Josef Breun in Germany; Svalöf Weibull Ab in Sweden; The Abed Foundation in Denmark; the spring and winter programs from Osijek Agricultural Institute of Croatia), Australia (University of Adelaide; Western Australia Department of Agriculture), and South America (National Agronomic Research Institutes of Chile and Uruguay).

Four testers repeated nine times and one tester repeated eleven times were also evaluated to complete the experimental design and to control for heading dates. Testers were Dayman, Perun, Ceibo, Clipper, and Quebracho. Data on testers is not shown.

**Field Trials**

All lines and testers were evaluated in a row-column (alpha lattice) design with 20 rows, 20 columns and 3 replications. Each line was sown in a hill-plot, where plots had a spacing of 40 cm between rows and alternating 40 and 60 cm between columns. Evaluations were conducted during the year 2005 at two locations in Uruguay. Colonia is in southwest Uruguay (34.20º S, 57.10º W, and 81 m.), and has fine, smectitic, thermic, Vertic Argiudol soils, while Young is in northern Uruguay (32.41 S, 57.40 W, and 80 m.), and has fine, smectitic, thermic, Typic Hapludert soils. Plots were fertilized with 45 kg ha\(^{-1}\) of urea to reach 40 mg kg\(^{-1}\) of N as NO\(_3^-\) at planting and nitrogen in plant was measured to adjust doses at the end of tillering but no N addition was needed.

Seeding date was 21 July and 30 July for Colonia and Young respectively. Seed emergence of the experiment was 30 July and 6 August. Given the disparity of origins of the materials, diseases were an important threat to plant survival. Therefore, weekly monitoring for disease was performed and a systemic fungicide was applied when such a
threat appeared and after disease scoring was completed. Each application consisted of 1 L ha\(^{-1}\) of the commercial fungicide Opera®: 133 g ha\(^{-1}\) of pyraclostrobin (Methyl N–[[1-(4 chlorophenyl)–1H– pyrazole –3–yl]oxymethyl]phenyl) N-methoxy carbamate) and 50 g ha\(^{-1}\) of epoxiconazole ((2RS,3RS)–1-[3-(2-chlorophenyl)–2,3-epoxy-2-(4-fluorophenyl)propyl]-1H-1,2,4-triazole). We applied fungicide twice in Colonia (7 Sept and 15 Oct), and once in Young (22 Sept). Full-plots were harvested on 7 Dec and 5 Dec for Colonia and Young respectively.

Several morphological traits were recorded for each line either on plots or on individual plants. The phenotypic traits measured on a plot level were: total number of tillers (counted at the end of tillering, NTIL), spot blotch disease (scoring from 1-5, where 1 is low and 5 is high, SB), leaf rust disease (scoring from 1-5, where 1 is low and 5 is high, LR), powdery mildew disease (scoring from 1-5, where 1 is low and 5 is high, PM), total number of days between planting date and anthesis (DTA), total number of days between planting and flowering (DTF), plant height (measured from the ground to the average of the tips of the spikes, in cm, HTOT), biomass weight at plant maturity (g, BWT), grain yield (g m\(^{-2}\), YLD), total number of spikes (NSPK), test weight (measured on a volume of 6mL, kg m\(^{-3}\), TWT), weight of 100 kernels (g, W100G).

Five plants from each plot were chosen at random at flowering time, and were color-marked with plastic twist-bands. Measurement of traits in single plants instead of whole plots allowed to decrease the experimental error and to estimate within plot variation. The following phenotypic traits were measured on each individual plant (on a plant level): flag leaf length (measured from the ligulae to the tip of the leaf, in cm, FLL), flag leaf width (measured at 2.5 cm from the ligulae, in cm, FLW), spike length (measured from the base to the tip of the spike, without counting the awns, in cm, SLT), awn length (measured from the top of the spike to the tip of the longest awn, in cm, ALT), peduncle length (measured from the last node to the base of the spike, in cm, PEDL), flag leaf height (measured from the ground to the ligulae of the flag leaf, in cm, FLH), spike height (measured from the ground to the base of the spike, in cm, SHT), and number of grains per spike (NG).
Statistical Models

Traits measured at the plot- and plant-levels were modeled according to the following linear models respectively:

\[
Y_{ijklmn} = B_{ij} + R_{k(ij)} + C_{l(ij)} + P_m + G_{n(m)} + \varepsilon_{ijklmn},
\]

\[
Y_{ijklmno} = B_{ij} + R_{k(ij)} + C_{l(ij)} + P_m + G_{n(m)} + I_{o(ijk)} + \varepsilon_{ijklmno},
\]

where \(B_{ij}\) = effect of \(j^{th}\) block in environment \(i\), \(R_{k(ij)}\) = effect of \(k^{th}\) row on \(ij^{th}\) block-environment, \(C_{l(ij)}\) = effect of \(l^{th}\) column on \(ij^{th}\) block-environment, \(P_m\) = effect of \(m^{th}\) breeding program, \(G_{n(m)}\) = effect of \(n^{th}\) genotype on \(m^{th}\) breeding program, \(I_{o(ijk)}\) = effect of the \(o^{th}\) plot on the \(ij^{th}\) block-environment, \(\varepsilon_{ijklmn}\) = residual error for the \(n^{th}\) genotype on the \(m^{th}\) breeding program of the \(ij^{th}\) block-environment, and \(\varepsilon_{ijklmno}\) = residual error for the \(o^{th}\) plant of the \(n^{th}\) genotype on the \(m^{th}\) breeding program of the \(ij^{th}\) block-environment. Plants within a plot share similarities from belonging to the same plot, therefore, the plot effect (\(I_{o(ijk)}\)) was included in the linear model for plant-level variables (equation 2).

We defined a hierarchical Bayesian model following Edwards and Jannink (2006), which allowed for heterogeneous genotypic variance within populations. We modeled both means (\(B_{ij}\), \(R_{k(ij)}\), \(C_{l(ij)}\), \(P_m\), \(G_{n(m)}\), and \(I_{o(ijk)}\)) and associated variances (\(\sigma_R^2\), \(\sigma_C^2\), \(\sigma_P^2\), \(\sigma_G^2\), \(\sigma_I^2\), and \(\sigma^2\)) in terms of explanatory variables (block, location, row, column, breeding program, genotype, and plot).

At the first level of the Bayesian hierarchy, observations were modeled as independent samples from a normal distribution:

\[
Y_{ijkl} \sim B_{ij}, R_{k(ij)}, C_{l(ij)}, P_m, G_{n(m)}, \sigma^2 \sim N(B_{ij} + R_{k(ij)} + C_{l(ij)} + P_m + G_{n(m)}, \sigma^2)
\]

The second level of the Bayesian hierarchy includes prior distributions for location parameters (i.e. means) \(B_{ij}\), \(R_{k(ij)}\), \(C_{l(ij)}\), \(P_m\), \(G_{n(m)}\), and \(I_{o(ijk)}\), and observational variance \(\sigma^2\). Priors on all location parameters were normal with mean zero and variances defined to condition the desired level of information sharing among levels of the factor. For block-environment means, \(B_{ij}\), the prior was defined with a very large variance to make the prior non-informative: \(B_{ij} \sim N(0, 10^{-7})\). The flat and independent prior is a Bayesian
equivalent to defining the block effect as a fixed effect in classical linear models. We did not include location effects in our model so that all parameters were estimable.

Row, column, breeding program, genotype and plot were modeled with priors that treated them as equivalent to random effects in classical mixed linear models. Row, column, population and plot effects were modeled as samples from a normal distribution with variance $\sigma^2_R$, $\sigma^2_C$, $\sigma^2_P$, and $\sigma^2_I$, respectively: $R_{k(ij)} \mid \sigma^2_R \sim N(0, \sigma^2_R)$, $C_{l(ij)} \mid \sigma^2_C \sim N(0, \sigma^2_C)$, $P_m \mid \sigma^2_P \sim N(0, \sigma^2_P)$, $I_{o(ijk)} \mid \sigma^2_I \sim N(0, \sigma^2_I)$.

Genotype effects were modeled as samples from a normal distribution with variance of the genotype effect as a function of the breeding program: $G_{n(m)} \mid \sigma^2_{G(m)} \sim N(0, \sigma^2_{G(m)})$. The subscripted notation on the variance of genotype indicates that every genotype had a unique variance of the breeding program. Variance of the genotype, $\sigma^2_{G(m)}$, was modeled with a generalized linear model using a natural-log link function: $\ln(\sigma^2_{G(m)}) = a + ap_m$. Where $a$ is the average natural logarithm of the genotypic variance, and $ap_m$ is the breeding program effect on the natural logarithm of the genotypic variance. The parameter $a$ conditions an average variance across all genotypes. The parameter $ap_m$ describes the degree to which the genotypic variance tends to be higher for observations on some breeding programs (positive values of $ap_m$), and lower for observations on other breeding programs (negative $ap_m$ values). The parameters $a$ and $ap_m$ were specified as $a \sim N(0, 10^7)$, and $ap_m \mid \sigma^2_{GP} \sim N(0, \sigma^2_{GP})$. $\sigma^2_{GP}$ express the degree of heterogeneity of genotypic variance within breeding programs. Homogenous within-breeding-program genotypic variances would correspond to $\sigma^2_{GP} = 0$, while large $\sigma^2_{GP}$ would indicate heterogeneous within-breeding-program genotypic variances. Therefore, a test of $\sigma^2_{GP} = 0$ is a test of homogeneity of genetic variance within breeding programs.

$\sigma^2_{GP}$ was given non-informative priors: $\sigma^2_{GP} \sim IG(0.0001,0.0001)$. Priors on the variance of row, column, and breeding program effects were chosen to be non-informative as: $\sigma^2_R \sim IG(0.0001,0.0001)$, $\sigma^2_C \sim IG(0.0001,0.0001)$, $\sigma^2_P \sim IG(0.0001,0.0001)$. Residual variance was also modeled as $\sigma^2 \sim IG(0.0001,0.0001)$.

The model and prior distributions represent the full model for heterogeneous variances. It was designed so that no prior data is used to inform the prior distributions.
The prior distributions were defined using hyperparameters, which themselves had non-informative prior distributions. All parameters were obtained via Markov Chain Monte Carlo simulation using the Bayesian Gibbs Sampling software WINBUGS (Spiegelhater et al., 2004).

**Statistical Analysis**

All analyses were performed accounting for the number of rows (i.e., two-row or six-row barleys) because they represent the two most distinct germplasm pools in barley (Powell et al., 1990; Takahashi et al., 1975). Depending on the specific analysis, we either included row number in the linear model, or we performed the analysis for two and six-row populations separately.

Least-squares means of breeding programs for all traits were obtained using the linear model described above but including row number in the model. Means were subject to principal component and cluster analysis using SAS statistical software (SAS Institute, 2004). We used the Ward method (Ward, 1963) of clustering, to group breeding programs with similar performance (i.e., similar means) for all variables. It is a hierarchical method that groups breeding programs producing the least increase in the sum of squares within groups (Manly, 1988). Consequently, breeding programs with similar means for all the variables should be assigned to the same clusters. The cubic clustering criterion and pseudo-F were used to decide on the number of groups (SAS Institute, 2004; Franco et al., 2005).

Least-squares means of genotypes for all traits were obtained using the linear models described above for plot (equation 1) and plant level (equation 2) variables, with row number in the model and genotypes as fixed effects. A stepwise discriminant analysis on genotypic means in SAS (SAS Institute, 2004) was used to identify the traits that best discriminated among two-row and six-row types, and among breeding programs. A STEPWISE selection method and a 15% significance level (P < 0.15) were used (Gutierrez et al., 2003; Franco et al., 1998). Neither leaf rust disease (LR) nor peduncle length (PEDL) were included in the discriminant analysis. Leaf rust disease was fitted
using only one location (Colonia) because there was not a significant outbreak of the disease in the other location. Peduncle length, in contrast, had a high number of missing values for some genotypes. Genotypic means were also used to identify the best performing genotypes for each variable in two and six-row types.

We tested whether there were significant differences in the amount of genetic variance within breeding programs ($\sigma^2_{GP}$) in WINBUGS (Spiegelhalter et al., 2003) by using the highest posterior density intervals (Gelman et al., 2003) of $\sigma^2_{GP}$. Once a variable was identified as having a significant difference in the amount of genetic variance within breeding programs ($\sigma^2_{GP} > 0$), non-overlapping 95% credible intervals were used to identify the least and most diverse breeding programs.

We also grouped breeding programs by their response to a change in the environment. For this purpose, least-squares means of breeding programs by location were obtained. The difference in mean values between the two locations was subject to cluster analysis using the Ward method (Ward, 1963) in SAS (SAS Institute, 2004). This procedure grouped breeding programs that had similar responses across variables to the change in the environment, and therefore would produce the least increase in the sum of squares within groups. If we assume only one variable, for example grain yield, two breeding programs that had high yields in one environment and low yields in the second (i.e., they had high values for the response variable: yield in location 1 - yield in location 2), would be assigned to the same cluster. On the other hand, breeding programs with minimal difference in yield among locations (response variable close to zero) would be assigned to a different cluster. For this analysis, neither leaf rust disease (LR) nor peduncle length (PEDL) were used because they did not have complete data sets for both locations. Additionally, the Germany (GE) breeding program was excluded from this analysis because there was not enough seed to plant the genotypes in both locations. The cubic clustering criterion and pseudo-F were used to decide on the number of groups (SAS Institute, 2004; Franco et al., 2005).
Results

Differences between Breeding Programs

There were significant differences (p < 0.0001) among breeding programs for all the variables analyzed (data not shown). In the case of two-row barley, the best grain yielding (YLD) programs in our test locations were Croatia spring, Uruguay, and North Dakota, while the worst were Chile and Croatia winter programs (Table 1). The Uruguayan program had the best test weight (TWT), while North Dakota, Uruguay and Australia had the best weight of 100 grains (W100). Alberta Canada was the program with the highest number of grains (NG, Table 2). North Dakotan, Australian, and Uruguayan programs were the earliest maturing programs (DTA and DTF, Table 1). European, Australian and the international program of Busch Ag Res had the shortest plants on average, and Canada and North Dakota had the tallest plants (HTOT), while European programs had low incidence of powdery mildew (PM, Table 1).

In the case of six-row barley, the best grain yields (YLD) and test weight (TWT) were from the Minnesota and Idaho breeding programs, while the worst were from winter Croatia (Table 1). The smallest number of grains (NG) was obtained for winter Croatia, while the other programs had similar number of grains (Table 2). The earliest maturing (DTA and DTF) program was Minnesota and the winter program of Croatia had the lowest incidence of powdery mildew (PM) and leaf rust (LR, Table 1).

The first two principal component axis explained 69% of the total variation. The first principal component axis and the cluster dendogram largely separated two-row from six-row breeding programs (Figure 1 and 2). All variables except grain yield (YLD) and spike length (SLT) had high loadings for the first eigenvector (data not shown). The second axis explained differences between best and worst yielding programs, with high positive loadings on variables grain yield (YLD), test weight (TWT), and weight of 100 grains (W100G), and high negative values on maturity variables (DTA, and DTF) for the second eigenvector.

Cluster analysis indicated four groups of genotypic performance, formed mainly by number of rows (two-row or six-row), grain yield (YLD), and maturity traits (DTA and
DTF), and did not reflect origin of the genotypes (Figure 2). Group 1 included the high yielding with good kernel characteristics (high test weights and weight of 100 grains) and early maturing two-row programs (AU-AD, US-BSI, CA-SK, US-ND, CRO-SP, and UY). Additionally, these programs had low incidence of leaf rust. Group 2 included the two-row programs with low biomass and grain yield, and short plants (AU-WE, GE, and SW). Group 3 included all the remaining two-row programs (CA-AB, US-ID, US-BS, DE, and US-WS), except CL and CRO-WI. The breeding program of Chile had extreme mean values for most variables, did not group with other programs, and included only four genotypes. Therefore, it was considered an outlier and was not include it in the discriminant analysis. Both winter breeding programs of Croatia also had extreme mean values for most variables, and since they were the only winter programs, we also excluded them from the discriminant analysis. Group 4 included all the six-row programs (CA-AB, US-ID, US-ND, US-BS, US-MN, and CA-SK) except CRO-WI. Finally, the outliers, the Chile and winter programs of Croatia grouped together. The groups within row type were maintained when separate analyses were performed by row numbers (data not shown). Therefore, the groups reflect true differences among breeding programs within two-row and six-row programs.

The variables that most discriminated the two-row from six-row spring genotypes were number of grains, spike length, flag leaf width, biomass weight, number of tillers, flag leaf length, spot blotch, and weight of 100 grains (Table 3). Among two-row programs, the only variable that did not discriminate breeding programs was grain yield (Table 3). Spike characteristics (SHT, NSPK, SLT, and NG), as well as maturity, tillering and yield components (TWT, and NG) were the variables that most discriminated among six-row breeding programs.

**Differences between Genotypes**

The variance between genotypes within breeding programs was significantly different from zero (p < 0.05) for all variables except spot blotch disease (data not shown). The two-row genotypes that produced the best grain yields were from Uruguay and the spring
program of Croatia (Table 4). While the six-row genotypes that yielded the most were from Canada Alberta and Busch Ag Res breeding programs. The earliest maturing two-row and six-row genotypes were from the Australia, and Minnesota, respectively. The latest maturing genotypes were from the winter program of Croatia and Denmark for two-row, and Canada Saskatchewan and the winter program of Croatia for six-row.

Several genotypes excelled, being among the top ten for three or more traits (see Table 4). Some of the two-row genotypes that excelled were CA-SK-12 (HB329, SB, TWT, and NG), US-BSI-02 (Z010J016J, BWT, YLD, NSPK), and US-ND-02 (ND13299, SB, BWT, and W100G). There were more single six-row genotypes that excelled, including lines from Canada Alberta (10: M79108001013, 12: M79108001013A, 15: H83030002, and 17: H87020011), with high values for biomass and grain yield, and low leaf rust incidence; Busch Ag Res (07: 6B98-9339, 09: 6B99-6639, and 15: 6B00-1499) with high values for grain yield and several other variables; North Dakota (07: ND231, 09: NDB125, and 14: Barless) with high performance lines for disease incidence and plant height; and Minnesota (11: Sep2-33 and 14: M99-68) with top lines for weight of 100 grains and several other traits.

**Diversity within Breeding Programs**

Days until anthesis, days until flowering, test weight, weight of 100 grains, awns length, spike height, and number of grains were the only traits that showed a significant difference among breeding programs in the within-program genetic variance (Tables 1 and 2). The spring programs of Croatia and Germany, the two-row program of Busch Ag Res, the six-row winter program of Croatia, the six-row programs of Idaho and Minnesota were the least diverse breeding programs. Those programs were among the least diverse for all the variables that had a significant difference in the within program genetic variance. Several programs were among the most diverse for three traits: Western Australia (DTA, DTF, and W100G), Canada Saskatchewan two-row (TWT, W100G, and SHT), and Sweden (TWT, ALT, and NG). Denmark, Canada Saskatchewan six-row, and...
Canada Alberta six-row were the most diverse programs for two traits, and Washington and North Dakota six-row were the most diverse program for one trait.

**Response to a Change in the Environment**

Two and six-row barleys tended to respond differently to the differences between environments (data not shown). Two-row programs had a more stable response to different environments (i.e., they had similar yields, grain characteristics and maturity in both environments, Figure 3). Six-row programs on the other hand, had a more drastic response to the change in the environment.

Six groups were produced by clustering breeding programs according to their differential response to the change in environments (Figure 3). Group 1 was formed by two-row programs including AU-AD, AU-WE, US-BSI, CRO-SP, and US-ND. Group 2 included only CL, which behaved very differently than most programs. Group 3 was formed by two-row programs, including CA-AB, US-WS, CA-SK, US-ID, UY, and US-BS. Group 4 included both two and six-row programs including DE and SW for two-row, and US-BS and US-MN for six-row. Group 5 included only six-row programs, CA-AB, US-ID, and US-ND. Finally, Group 6 included CA-SK, and two and six-row programs of CRO-WI. The amount of difference in grain yield across locations was consistent with other variables (Figure 3), where breeding programs that were stable across environments for most variables were also stable for grain yield.

**Discussion**

Exchange of genetic material is essential for enhancing genetic gain in a breeding program. We evaluated two relevant aspects of breeding programs that will aid in germplasm exchange. First, we evaluated the amount of genetic diversity within breeding programs as there are some instances in which lack of genetic diversity can impede genetic gain (Gepts, 2006). We found significant genotypic variation within breeding programs for all 20 traits we measured. There were differences in the amount of genotypic variation within breeding programs for seven of these traits. We were able to
identify breeding programs that had systematically more genetic diversity and programs that had less genetic diversity. Second, we established, in a data-driven way, groups of breeding programs that would benefit from germplasm exchange (i.e., MTS). Using an analogue to ME (Braun et al., 1996; Gauch and Zobel, 1997), we were able to produce groups of breeding programs with similar performance and response to a change in the environment. We found significant variation among breeding programs for all quantitative traits studied. While grouping by performance was not related to the geographical location of the breeding programs (i.e., the two Australian and the two Canadian breeding programs were assigned to different groups and European and other American breeding programs were also found in several groups), grouping by response to a change in the environment was related to the geographical location of the breeding programs (i.e., both Australian breeding programs were assigned to the same group, and both Canadian programs and European programs were also assigned to the same groups). The combination of grouping by performance and by response to the change in the environment is relevant in the definition of Mega-Targets of Selection, and is discussed below.

The two most distinct germplasm pools in barley are the two-row and six-row barleys (Powell et al., 1990; Takahashi et al., 1975). Even though there are only two epistatic loci involved in the distinction between two and six-rows (Franckowiak and Lundqvist, 1997), due to historical patterns of breeding for usage and geographical distribution (i.e., two-row is used for malting in most of the world, except in the United States and Mexico), there are differences at other quantitative traits between the groups (Takahashi et al., 1975). We found differences between two-row and six-row for most of the traits studied. Six-row genotypes had more grains per spike, but shorter spikes, and lower weight of 100 grains. Additionally, six-row genotypes had fewer tillers and biomass weight, longer and wider leaves, and less spot blotch disease incidence. Other studies also found that two-row barleys have usually more tillers and larger, heavier seeds, while six-row barleys have more seeds per inflorescence (Marquez-Cedillo et al., 2001). Some studies explain this by pleiotropic effects (Allard, 1988), and others found linkage
between those loci and quantitative traits such as yield, kernel plumpness, test weight, heading date, and plant height (Marquez-Cedillo et al., 2001).

Genotype-by-environment interactions are common in nature (Allard and Bradshaw, 1964) and significant genotype-by-environment and breeding program-by-environment interactions were found in this study (data not shown). However, those interactions were mainly magnitude differences across environments, and not cross-over interactions. Furthermore, principal component and cluster analysis by environment produced the same groups (data not shown). Principal component and cluster analysis using variables in different environments as different traits did not change either the grouping of the breeding programs. Therefore, there are no hidden effects of genotype by location or breeding program by location.

**Genetic Diversity**

Narrowing of the gene pool in barley due to breeding efforts is still under debate, with some evidence in favor (Ordon et al., 2005; Rasmussen and Phillips, 1997; Russell et al., 1997) Russel et al 2000) and some against (Khlestkina et al., 2006; Koebner et al., 2003; Malysheva-Otto et al., 2007; Ordon et al., 2005). We found that some breeding programs preserved more genetic diversity than others. The Minnesota breeding program for instance was one of the programs with lowest diversity for all traits in our study. Low diversity has been documented for Minnesota using pedigree information (Rassmusson and Philips, 1997). The breeding program of Germany was also one of the programs with lowest diversity. However, Germany is among European two-row and six-row barleys identified to possess both a narrow gene pool (Russell et al., 1997) and no historical change in the genetic diversity (Malysheva-Otto et al., 2007) based on molecular markers. No previous information was found on the amount of diversity for the other programs with low diversity. Additionally, some of the most diverse breeding programs, such as those from Sweden and Denmark, were examined in studies where no narrowing of the gene pool was detected (Malysheva-Otto et al., 2007). The purpose of our research was to study the differential amount of diversity for quantitative traits of breeding interest
allocated in the breeding programs. It was out of the scope of our paper to study diversity at molecular markers. However, using the information from both quantitative traits and neutral molecular markers would provide a better understanding of the loss of diversity in the breeding programs.

Several mechanisms could explain high levels of diversity for the traits examined here. For instance, maturity traits (DTA and DTF) showed higher diversity in Western Australia than in all of the other programs. One mechanism that could explain this is that Western Australia breeds for a wide set of environmental conditions, including differences in photoperiod. In other traits like awn length and spike height, the high diversity could be explained by the lack of selection. Traits like test weight, weight of 100 grains, and number of grains could show high levels of diversity in programs that do not select explicitly for them, but only as a part of the yield component. This could happen if for instance there is no premium for kernel plumpness or if the breeding program includes both feed and malt varieties.

**Mega-Targets of Selection**

In order to identify data-driven groupings of breeding programs that would benefit from germplasm exchange, we conducted two distinct analysis. First, we grouped breeding programs by their performance in the environments studied. Four groups (and some outliers) were produced by genotypic performance: three groups of two-row barleys separated by grain yield and maturity, and one group of six-row barleys. Clusters based on genotypic performance do not necessarily group breeding programs that would benefit from germplasm exchange as genotypes from two breeding programs could both perform poorly in an environment due to different causes. For example, one set of genotypes might be limited because of disease pressure, the other by photoperiod conditions such that the programs are not adapted to the same environmental conditions. Therefore, in a second approach, we grouped breeding programs by their response to a change in the environment. Again, four groups (and some outliers) were produced by their response to the change in the environment: two groups of two-row barleys, one group of two and six-
row barleys, and one group of six-row barleys. These groups were related to the location of the breeding program.

Using both criteria, we identified three MTS. These are sets of breeding programs that belong to the same groups of genotypic performance and response to selection (Figure 2 and 3). The first MTS includes high-yielding two-row programs with good kernel properties (i.e., Group 1 of genotypic performance) that have a small response to the change in the environment (i.e., Group 1 of response to the change in the environment). The breeding programs included in this group are Australia Adelaide, the international program of Busch Ag Res, North Dakota, and the spring program of Croatia. The second MTS includes the two-row programs that have an average performance for all variables (i.e., Group 3 of genotypic performance) that have the largest response to a change in the environment (i.e., Group 3 of response to a change in the environment). The breeding programs included in this group are Canada Alberta, Idaho, Washington, and Busch Ag Res program. Finally, the third MTS comprises the six-row breeding programs that have a drastic response to the change in the environment. The programs included in this group are Canada Alberta, Idaho and North Dakota.

Exchanging germplasm within these groups should be beneficial because we expect genotypes to be adapted to similar conditions and perform similarly. Additional pairs of programs that could benefit from germplasm exchange are two-row Canada Saskatchewan and Uruguay, and six-row Busch Ag Res and Minnesota.

MTS are analogue to mega-environments (ME). In ME groups of environments that produce the same rank of the genotypes are formed (Yan et al., 2000). Genotypic evaluation in any of the environments within a ME is equivalent. In MTS, groups of breeding programs are formed such that exchanging germplasm material within a group will produce genotypes that are well adapted and respond similarly to the new environmental conditions. Similarly to ME, where using more genotypes and environments allow for broader generalization, in MTS, using more environments and breeding programs allow for broader generalization. Our study is limited by the number of environments used for genotypic evaluation. However, the main focus in our work is
in grouping breeding programs and not environments, and we evaluated 23 breeding programs. Additionally, there are numerous reports in which only a small number of genotypes and/or environments were used to study ME (Blanche and Myers, 2006; Robins et al., 2007; Samonte et al., 2005). Those studies are valuable within the genetic background of the genotypes and the environments evaluated. We used two environments with distinct soil types, mean temperature, precipitation, and disease pressure.

**Conclusion**

Ideally for germplasm exchange, we would require elite germplasm, well adapted, with different alleles at the loci of interest. We evaluated elite germplasm and provided a methodology to identify sets of breeding programs that are adapted to similar conditions, and therefore with whom germplasm exchange could be favorable. We also evaluated the performance of the genotypes to identify high yielding materials and we evaluated the amount of genetic diversity allocated in each program. Although outside the scope of this study, it would be useful to also identify genotypes with different alleles for the loci of interest.

**Acknowledgements**

We thank all breeders who provided seed: J. Eglinton (AU-AD), C. Li (AU-WE), J. Helm (CA-AB), B. Rossnagel (CA-SK), I. Matus (CL), J.E. Diaz (UY), B. Cooper (US-BS, US-BSI), D. Obert (US-ID), K. Smith (US-MN), J. Franckowiak (US-ND-2R), R. Horsley (US-ND-6R), S. Ullrich (US-WS), J. Lohde (SW), M.H. Poulsen (DE), B. Eriksen (DE), J. Breun (GE), A. Lalic (CRO-SP, CRO-WI). The authors would like to thank Dean Adams for his comments and suggestions. We would also like to thank the research group at INIA La Estanzuela in Uruguay for support with field evaluations. L.G. was supported by a Natural Systems Agriculture Fellowship from The Land Institute during this study.
References


List of Figures

**Figure 1.** Representation of means of breeding programs in the first two principal component axes. Breeding program abbreviations given in Table 1.

**Figure 2.** Cluster dendogram of mean genotypic performance of breeding programs (abbreviations given in Table 1). The line indicates the separation of groups based on the clustering criterion and the number of each group is indicated in the left.

**Figure 3.** Cluster dendogram of the difference of means across Colonia and Young locations for all variables showing the response to a change in the environment, and difference in grain yield among Colonia and Young locations. Breeding program abbreviations given in Table 1. The line indicates the separation of groups based on the clustering criterion and the number of each group is indicated in the left.
Table 1. Mean and diversity of breeding program for plot level traits: Mean breeding program values and standard errors (in parenthesis) and least (‡) and most (§) diverse programs for the traits measured at the plot level, number of tillers (NTIL), spot blotch disease (SB), leaf rust disease (LR), powdery mildew disease (PM), days until anthesis (DTA), days until flowering (DTF), plant height (HTOT), biomass (BWT), grain yield (YLD), number of spikes (NSPK), test weight (TWT), and weight of 100 grains (W100G).

<table>
<thead>
<tr>
<th>BP</th>
<th>N</th>
<th>NTIL</th>
<th>SB</th>
<th>LR</th>
<th>PM</th>
<th>DTA</th>
<th>DTF</th>
<th>HTOT</th>
<th>BWT</th>
<th>YLD</th>
<th>NSPK</th>
<th>TWT</th>
<th>W100G</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>two-row</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AU-AD</td>
<td>20</td>
<td>78.4</td>
<td>4.3</td>
<td>1.6</td>
<td>0.1</td>
<td>2.3</td>
<td>0.1</td>
<td>1.7</td>
<td>0.2</td>
<td>80.8</td>
<td>1.3</td>
<td>398.2</td>
<td>131.5</td>
</tr>
<tr>
<td>AU-WE</td>
<td>15</td>
<td>69.3</td>
<td>4.6</td>
<td>1.7</td>
<td>0.1</td>
<td>2.7</td>
<td>0.2</td>
<td>2.1</td>
<td>0.2</td>
<td>82.8</td>
<td>1.5</td>
<td>498.2</td>
<td>120.7</td>
</tr>
<tr>
<td>CA-AB</td>
<td>7</td>
<td>78.6</td>
<td>5.5</td>
<td>1.7</td>
<td>0.1</td>
<td>2.6</td>
<td>0.2</td>
<td>1.9</td>
<td>0.2</td>
<td>90.1</td>
<td>1.5</td>
<td>598.2</td>
<td>142.5</td>
</tr>
<tr>
<td>CA-SK</td>
<td>16</td>
<td>76.2</td>
<td>4.6</td>
<td>1.5</td>
<td>0.1</td>
<td>2.8</td>
<td>0.2</td>
<td>2.4</td>
<td>0.2</td>
<td>85.8</td>
<td>1.2</td>
<td>698.2</td>
<td>212.8</td>
</tr>
<tr>
<td>CL</td>
<td>4</td>
<td>60.7</td>
<td>6.9</td>
<td>1.6</td>
<td>0.2</td>
<td>2.9</td>
<td>0.2</td>
<td>1.7</td>
<td>0.3</td>
<td>92.1</td>
<td>1.7</td>
<td>798.2</td>
<td>106.9</td>
</tr>
<tr>
<td>CRO-SP</td>
<td>20</td>
<td>82.1</td>
<td>4.3</td>
<td>2.0</td>
<td>0.1</td>
<td>2.3</td>
<td>0.1</td>
<td>1.3</td>
<td>0.2</td>
<td>86.8</td>
<td>1.2</td>
<td>898.2</td>
<td>287.9</td>
</tr>
<tr>
<td>CRO-WI</td>
<td>15</td>
<td>94.9</td>
<td>4.7</td>
<td>1.4</td>
<td>0.1</td>
<td>2.8</td>
<td>0.2</td>
<td>2.0</td>
<td>0.2</td>
<td>95.0</td>
<td>1.3</td>
<td>998.2</td>
<td>154.0</td>
</tr>
<tr>
<td>DE</td>
<td>20</td>
<td>91.5</td>
<td>4.5</td>
<td>2.0</td>
<td>0.1</td>
<td>2.9</td>
<td>0.1</td>
<td>0.8</td>
<td>0.2</td>
<td>91.1</td>
<td>1.3</td>
<td>938.2</td>
<td>145.4</td>
</tr>
<tr>
<td>GE</td>
<td>20</td>
<td>77.1</td>
<td>5.8</td>
<td>1.5</td>
<td>0.2</td>
<td>2.3</td>
<td>0.2</td>
<td>1.8</td>
<td>0.2</td>
<td>91.7</td>
<td>1.5</td>
<td>918.2</td>
<td>130.2</td>
</tr>
<tr>
<td>SW</td>
<td>20</td>
<td>82.2</td>
<td>4.4</td>
<td>2.0</td>
<td>0.1</td>
<td>2.5</td>
<td>0.1</td>
<td>1.3</td>
<td>0.2</td>
<td>91.7</td>
<td>1.2</td>
<td>978.2</td>
<td>129.5</td>
</tr>
<tr>
<td>US-BS</td>
<td>18</td>
<td>80.0</td>
<td>4.3</td>
<td>1.6</td>
<td>0.1</td>
<td>3.0</td>
<td>0.1</td>
<td>2.4</td>
<td>0.2</td>
<td>89.0</td>
<td>1.2</td>
<td>888.2</td>
<td>153.1</td>
</tr>
<tr>
<td>US-BI</td>
<td>17</td>
<td>78.1</td>
<td>4.5</td>
<td>1.9</td>
<td>0.1</td>
<td>2.8</td>
<td>0.1</td>
<td>1.1</td>
<td>0.2</td>
<td>89.5</td>
<td>1.4</td>
<td>848.2</td>
<td>137.2</td>
</tr>
<tr>
<td>US-ID</td>
<td>13</td>
<td>76.5</td>
<td>5.0</td>
<td>1.6</td>
<td>0.1</td>
<td>3.3</td>
<td>0.2</td>
<td>2.4</td>
<td>0.2</td>
<td>88.4</td>
<td>1.3</td>
<td>758.2</td>
<td>135.5</td>
</tr>
<tr>
<td>US-ND</td>
<td>20</td>
<td>66.2</td>
<td>4.3</td>
<td>1.7</td>
<td>0.1</td>
<td>3.0</td>
<td>0.1</td>
<td>2.9</td>
<td>0.2</td>
<td>80.6</td>
<td>1.3</td>
<td>738.2</td>
<td>151.2</td>
</tr>
<tr>
<td>US-SW</td>
<td>19</td>
<td>82.1</td>
<td>4.4</td>
<td>1.8</td>
<td>0.1</td>
<td>2.9</td>
<td>0.1</td>
<td>2.1</td>
<td>0.2</td>
<td>88.5</td>
<td>1.2</td>
<td>718.2</td>
<td>146.5</td>
</tr>
<tr>
<td>UY</td>
<td>19</td>
<td>76.7</td>
<td>4.3</td>
<td>1.7</td>
<td>0.1</td>
<td>2.6</td>
<td>0.1</td>
<td>1.5</td>
<td>0.2</td>
<td>83.5</td>
<td>1.2</td>
<td>708.2</td>
<td>143.2</td>
</tr>
<tr>
<td><strong>six-row</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA-AB</td>
<td>14</td>
<td>57.9</td>
<td>4.8</td>
<td>1.3</td>
<td>0.1</td>
<td>3.3</td>
<td>0.2</td>
<td>2.4</td>
<td>0.2</td>
<td>84.8</td>
<td>1.4</td>
<td>398.2</td>
<td>127.1</td>
</tr>
<tr>
<td>CA-SK</td>
<td>5</td>
<td>49.5</td>
<td>6.2</td>
<td>1.3</td>
<td>0.2</td>
<td>3.2</td>
<td>0.2</td>
<td>2.1</td>
<td>0.3</td>
<td>87.5</td>
<td>1.9</td>
<td>398.2</td>
<td>109.3</td>
</tr>
<tr>
<td>CRO-WI</td>
<td>4</td>
<td>91.7</td>
<td>6.7</td>
<td>1.3</td>
<td>0.2</td>
<td>3.0</td>
<td>0.2</td>
<td>1.9</td>
<td>0.3</td>
<td>93.4</td>
<td>1.8</td>
<td>398.2</td>
<td>111.6</td>
</tr>
<tr>
<td>US-BS</td>
<td>14</td>
<td>45.2</td>
<td>4.4</td>
<td>1.3</td>
<td>0.1</td>
<td>3.4</td>
<td>0.1</td>
<td>2.8</td>
<td>0.2</td>
<td>83.2</td>
<td>1.2</td>
<td>398.2</td>
<td>122.9</td>
</tr>
<tr>
<td>US-ID</td>
<td>7</td>
<td>48.5</td>
<td>5.6</td>
<td>1.3</td>
<td>0.1</td>
<td>3.4</td>
<td>0.1</td>
<td>2.4</td>
<td>0.2</td>
<td>82.1</td>
<td>1.5</td>
<td>398.2</td>
<td>123.2</td>
</tr>
<tr>
<td>US-MN</td>
<td>18</td>
<td>43.7</td>
<td>4.3</td>
<td>1.4</td>
<td>0.1</td>
<td>3.8</td>
<td>0.1</td>
<td>2.8</td>
<td>0.2</td>
<td>80.8</td>
<td>1.2</td>
<td>398.2</td>
<td>121.6</td>
</tr>
<tr>
<td>US-ND</td>
<td>15</td>
<td>54.5</td>
<td>4.5</td>
<td>1.2</td>
<td>0.1</td>
<td>3.2</td>
<td>0.2</td>
<td>2.8</td>
<td>0.2</td>
<td>82.9</td>
<td>1.2</td>
<td>398.2</td>
<td>125.0</td>
</tr>
</tbody>
</table>

‡ Breeding Programs (BP): Adelaide Australia (AU-AD), Western Australia (AU-WE), Alberta Canada (CA-AB), Saskatchewan Canada (CA-SK), INIA Chile (CL), spring Croatia (CRO-WI), winter Croatia (CRO-SP), winter Croatia (CRO-WI), Abed Denmark (DE), Breun Germany (GE), Svalof Sweden (SW), Busch USA (US-BS), international Busch USA (US-BSI), Idaho USA (US-ID), Minnesota USA (US-MN), North Dakota USA (US-ND), Washington USA (US-WS), and INIA Uruguay (UY).

§ GVW = variance in the genetic variance within breeding programs.

* significant at the 5% level.

**USA** (US-MN), **North Dakota USA** (US-ND), **Washington USA** (US-WS), and INIA Uruguay (UY).
Table 2. Mean and diversity of breeding program for plant level traits: Mean breeding program values and standard errors (in parenthesis) and least (‡) and most ($) diverse programs for the traits measured at the plant level. flag leaf length (FLL), flag leaf width (FLW), spike length (SLT), awns length (ALT), flag leaf height (FLH), spike height (SHT), peduncle length (PEDL), and number of grains per spike (NG).

<table>
<thead>
<tr>
<th>BP†</th>
<th>FLL</th>
<th>FLW</th>
<th>SLT</th>
<th>ALT</th>
<th>FLL</th>
<th>SHT</th>
<th>PEDL</th>
<th>NG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cm</td>
<td>cm</td>
<td>cm</td>
<td>cm</td>
<td>cm</td>
<td>cm</td>
<td>cm</td>
<td></td>
</tr>
<tr>
<td>two-row</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AU-AD</td>
<td>11.4 (0.6)</td>
<td>0.75 (0.1)</td>
<td>8.2 (0.2)</td>
<td>12.1 (0.3)</td>
<td>65.6 (1.2)</td>
<td>66.6 (1.8) ‡</td>
<td>5.6 (1.0)</td>
<td>24.5 (3.2) †</td>
</tr>
<tr>
<td>AU-WE</td>
<td>12.2 (0.7)</td>
<td>0.75 (0.1)</td>
<td>9.0 (0.2)</td>
<td>10.9 (0.2) ‡</td>
<td>67.9 (1.2)</td>
<td>69.6 (1.8) ‡</td>
<td>5.1 (1.0)</td>
<td>25.4 (3.2) †</td>
</tr>
<tr>
<td>CA-AB</td>
<td>14.8 (0.8)</td>
<td>0.97 (0.1)</td>
<td>10.0 (0.3)</td>
<td>10.5 (0.3) ‡</td>
<td>72.6 (1.5)</td>
<td>76.8 (2.0) ‡</td>
<td>6.1 (1.2)</td>
<td>33.1 (3.6) †</td>
</tr>
<tr>
<td>CA-SK</td>
<td>13.4 (0.6)</td>
<td>0.84 (0.1)</td>
<td>9.7 (0.2)</td>
<td>11.3 (0.3) ‡</td>
<td>72.1 (1.3)</td>
<td>75.3 (2.1) ‡</td>
<td>6.3 (1.0)</td>
<td>28.8 (3.2) †</td>
</tr>
<tr>
<td>CL</td>
<td>14.1 (1.0)</td>
<td>0.81 (0.1)</td>
<td>9.6 (0.3)</td>
<td>12.1 (0.4) ‡</td>
<td>70.4 (1.8)</td>
<td>72.1 (2.5)</td>
<td>5.9 (1.5)</td>
<td>28.7 (3.9) †</td>
</tr>
<tr>
<td>CRO-SP</td>
<td>12.5 (0.7)</td>
<td>0.72 (0.1)</td>
<td>8.7 (0.2)</td>
<td>11.3 (0.2) ‡</td>
<td>71.2 (1.1)</td>
<td>72.9 (1.6) ‡</td>
<td>4.5 (1.0)</td>
<td>27.1 (3.3) †</td>
</tr>
<tr>
<td>CRO-WI</td>
<td>11.3 (0.7)</td>
<td>0.69 (0.1)</td>
<td>9.4 (0.2)</td>
<td>10.8 (0.3) ‡</td>
<td>70.6 (1.3)</td>
<td>72.2 (1.7) ‡</td>
<td>1.8 (1.1)</td>
<td>27.3 (3.2) †</td>
</tr>
<tr>
<td>DE</td>
<td>11.7 (0.6)</td>
<td>0.66 (0.1)</td>
<td>9.1 (0.2)</td>
<td>11.7 (0.3) ‡</td>
<td>67.9 (1.1)</td>
<td>65.0 (1.5) ‡</td>
<td>0.6 (0.9)</td>
<td>26.7 (3.2) †</td>
</tr>
<tr>
<td>GE</td>
<td>12.2 (0.7)</td>
<td>0.72 (0.1)</td>
<td>9.6 (0.2)</td>
<td>11.7 (0.3) ‡</td>
<td>71.9 (1.3)</td>
<td>71.6 (1.7) ‡</td>
<td>5.6 (3.2)</td>
<td>27.3 (3.4) †</td>
</tr>
<tr>
<td>SW</td>
<td>13.3 (0.7)</td>
<td>0.77 (0.1)</td>
<td>9.4 (0.2)</td>
<td>11.0 (0.3) ‡</td>
<td>69.8 (1.1)</td>
<td>69.0 (1.6) ‡</td>
<td>3.6 (1.0)</td>
<td>27.9 (3.3) §</td>
</tr>
<tr>
<td>US-BS</td>
<td>12.5 (0.7)</td>
<td>0.88 (0.1)</td>
<td>9.4 (0.2)</td>
<td>11.3 (0.2) ‡</td>
<td>71.6 (1.1)</td>
<td>75.1 (1.5) ‡</td>
<td>5.7 (1.0)</td>
<td>29.2 (3.4) †</td>
</tr>
<tr>
<td>US-BSI</td>
<td>12.2 (0.7)</td>
<td>0.72 (0.1)</td>
<td>9.2 (0.2)</td>
<td>11.6 (0.2) ‡</td>
<td>69.7 (1.2)</td>
<td>69.1 (1.5) ‡</td>
<td>2.2 (1.0)</td>
<td>26.4 (3.2) †</td>
</tr>
<tr>
<td>US-ID</td>
<td>13.1 (0.7)</td>
<td>0.79 (0.1)</td>
<td>9.4 (0.2)</td>
<td>11.3 (0.3)</td>
<td>72.5 (1.3)</td>
<td>75.7 (1.8) ‡</td>
<td>5.8 (1.1)</td>
<td>28.3 (3.2) †</td>
</tr>
<tr>
<td>US-ND</td>
<td>15.0 (0.7)</td>
<td>1.02 (0.1)</td>
<td>9.2 (0.2)</td>
<td>10.8 (0.2) ‡</td>
<td>73.3 (1.1)</td>
<td>79.8 (1.5) ‡</td>
<td>9.8 (0.9)</td>
<td>26.8 (3.3) †</td>
</tr>
<tr>
<td>US-WS</td>
<td>12.4 (0.7)</td>
<td>0.73 (0.1)</td>
<td>9.1 (0.2)</td>
<td>11.6 (0.3)</td>
<td>71.7 (1.1)</td>
<td>72.9 (1.6) ‡</td>
<td>3.9 (0.9)</td>
<td>29.2 (3.4) §</td>
</tr>
<tr>
<td>UY</td>
<td>12.5 (0.7)</td>
<td>0.76 (0.1)</td>
<td>8.6 (0.2)</td>
<td>11.5 (0.2) ‡</td>
<td>71.2 (1.2)</td>
<td>72.1 (1.7) ‡</td>
<td>0.6 (1.0)</td>
<td>25.7 (3.2) †</td>
</tr>
<tr>
<td>six-row</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA-AB</td>
<td>17.3 (0.7)</td>
<td>1.63 (0.1)</td>
<td>9.1 (0.2)</td>
<td>10.3 (0.3) ‡</td>
<td>67.6 (1.4)</td>
<td>72.3 (2.0) §</td>
<td>7.6 (1.1)</td>
<td>63.5 (3.3) †</td>
</tr>
<tr>
<td>CA-SK</td>
<td>15.4 (0.9)</td>
<td>1.48 (0.1)</td>
<td>8.8 (0.3)</td>
<td>10.4 (0.4) ‡</td>
<td>71.9 (1.9)</td>
<td>76.0 (3.0) §</td>
<td>5.9 (1.4)</td>
<td>59.4 (4.3) †</td>
</tr>
<tr>
<td>CRO-WI</td>
<td>10.6 (1.0)</td>
<td>0.72 (0.1)</td>
<td>8.2 (0.3)</td>
<td>11.6 (0.4) ‡</td>
<td>71.5 (1.8)</td>
<td>72.0 (2.4) ‡</td>
<td>0.8 (1.6)</td>
<td>52.0 (4.4) †</td>
</tr>
<tr>
<td>US-BS</td>
<td>17.9 (0.7)</td>
<td>1.67 (0.1)</td>
<td>8.8 (0.2)</td>
<td>10.7 (0.2) ‡</td>
<td>73.1 (1.2)</td>
<td>80.3 (1.5) ‡</td>
<td>8.6 (1.0)</td>
<td>62.4 (3.3) †</td>
</tr>
<tr>
<td>US-ID</td>
<td>16.6 (0.9)</td>
<td>1.52 (0.1)</td>
<td>8.7 (0.3)</td>
<td>10.5 (0.3) ‡</td>
<td>71.3 (1.5)</td>
<td>78.4 (1.9) ‡</td>
<td>10.2 (1.2)</td>
<td>63.4 (3.3) †</td>
</tr>
<tr>
<td>US-MN</td>
<td>18.7 (0.7)</td>
<td>1.67 (0.1)</td>
<td>8.8 (0.2)</td>
<td>10.7 (0.2) ‡</td>
<td>70.6 (1.1)</td>
<td>77.4 (1.6) ‡</td>
<td>9.9 (1.0)</td>
<td>61.2 (3.3) †</td>
</tr>
<tr>
<td>US-ND</td>
<td>17.1 (0.7)</td>
<td>1.66 (0.1)</td>
<td>8.6 (0.2)</td>
<td>10.2 (0.1) ‡</td>
<td>77.4 (1.4)</td>
<td>86.3 (1.8) ‡</td>
<td>10.2 (1.0)</td>
<td>61.0 (3.5) §</td>
</tr>
</tbody>
</table>

GVW¶ 0.508 (0.45) 0.801 (0.49) 1.166 (0.77) 0.292 (0.27)* 0.801 (0.49) 1.166 (0.77)* 0.292 (0.27) 0.882 (0.27)*

† Breeding Programs (BP): Adelaide Australia (AU-AD), Western Australia (AU-WE), Alberta Canada (CA-AB), Saskatchewan Canada (CA-SK), INIA Chile (CL), spring Croatia (CRO-SP), winter Croatia (CRO-WI), Abed Denmark (DE), Bremen Germany (GE), Svalöf Sweden (SW), Busch USA (US-BS), international Busch USA (US-BSI), Idaho USA (US-ID), Minnesota USA (US-MN), North Dakota USA (US-ND), Washington USA (US-WS), and INIA Uruguay (UY).

¶ GVW = variance in the genetic variance within breeding programs.

* significant at the 5% level.
Table 3. Discriminant traits among two- and six-row and breeding programs. Traits are: number of tillers (NTIL), spot blotch disease (SB), leaf rust disease (LR), powdery mildew disease (PM), days until anthesis (DTA), days until flowering (DTF), plant height (HTOT), biomass weight (BWT), grain yield (YLD), number of spikes (NSPK), test weight (TWT), and weight of 100 grains (W100G), flag leaf length (FLL), flag leaf width (FLW), spike length (SLT), awns length (ALT), flag leaf height (FLH), spike height (SHT), peduncle length (PEDL), and number of grains per spike (NG).

<table>
<thead>
<tr>
<th>Trait</th>
<th>F-value</th>
<th>P</th>
<th>Trait</th>
<th>F-value</th>
<th>P</th>
<th>Trait</th>
<th>F-value</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTIL</td>
<td>5.7</td>
<td>0.0178</td>
<td>DTA</td>
<td>16.0</td>
<td>&lt;.0001</td>
<td>SHT</td>
<td>8.7</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>SLT</td>
<td>63.8</td>
<td>&lt;.0001</td>
<td>HTOT</td>
<td>13.3</td>
<td>&lt;.0001</td>
<td>NSPK</td>
<td>6.1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>BWT</td>
<td>22.5</td>
<td>&lt;.0001</td>
<td>NTL</td>
<td>6.1</td>
<td>&lt;.0001</td>
<td>SLT</td>
<td>4.7</td>
<td>0.0009</td>
</tr>
<tr>
<td>W100G</td>
<td>2.5</td>
<td>0.1130</td>
<td>TWT</td>
<td>5.5</td>
<td>&lt;.0001</td>
<td>DTA</td>
<td>4.2</td>
<td>0.002</td>
</tr>
<tr>
<td>W100G†</td>
<td>0.7</td>
<td>0.6346</td>
<td>W100G</td>
<td>3.8</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† trait removed during the STEPDSC procedure.
Table 4. Ten best genotypes for some variables of breeding interest: spot blotch disease (SB), leaf rust disease (LR), powdery mildew disease (PM), days until anthesis (DTA), plant height (HTOT), biomass weight (BWT), grain yield (YLD), number of spikes (NSPK), test weight (TWT), weight of 100 grains (W100G), and number of grains (NG). Breeding program abbreviations given in Table 1.

<table>
<thead>
<tr>
<th>SB</th>
<th>LR</th>
<th>PM</th>
<th>DTA</th>
<th>DTA</th>
<th>HTOT</th>
<th>BWT</th>
<th>YLD</th>
<th>NSPK</th>
<th>TWT</th>
<th>W100G</th>
<th>NG</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA-SK-12</td>
<td>CA-SK-12</td>
<td>DE-03</td>
<td>AU-AD-09</td>
<td>DE-20</td>
<td>DE-06</td>
<td>CA-SK-05</td>
<td>UY-14</td>
<td>CRO-SP-07</td>
<td>UY-18</td>
<td>UY-18</td>
<td>CA-SK-03</td>
</tr>
<tr>
<td>US-ND-14</td>
<td>US-ND-14</td>
<td>CRO-WI-01</td>
<td>USA-MN-19</td>
<td>CA-SK-21</td>
<td>CA-SK-14</td>
<td>CA-AB-17</td>
<td>USA-MN-07</td>
<td>CA-AB-17</td>
<td>USA-BS-10</td>
<td>USA-MN-03</td>
<td>USA-MN-12</td>
</tr>
</tbody>
</table>
Figure 1. Principal component analysis of means of breeding programs
Figure 2. Cluster dendogram of performance of breeding programs
Figure 3. Cluster dendogram of difference of means and grain yield difference
CHAPTER V: EVOLUTIONARY HISTORY OF AGRONOMICALLY RELEVANT TRAITS IN BARLEY

A paper to be submitted to *Crop Science*

Lucia Gutierrez, John D Nason, Jean-Luc Jannink

**Abstract**

It has been proposed that understanding both the evolutionary history of agriculturally relevant traits and the structure of the diversity allocated among and within populations is crucial for genetic resource utilization and conservation. We propose the use of a methodology widely used in evolutionary studies that accomplishes both objectives by comparing the population structure of quantitative traits ($Q_{st}$) to the population structure of neutral molecular markers ($F_{st}$). The main objective of this study was to understand the evolutionary history of agriculturally relevant traits in cultivated barley. Specifically, we first tested the hypothesis of neutral evolution of quantitative traits. Second, we described the patterns of divergence across barley breeding programs distributed worldwide. Third, we compared the patterns of selection in the traits with those of wild barley. Finally, we discussed the relevance of this methodology for breeding purposes, genetic resources conservation and utilization purposes, and for understanding the domestication process.

We used 66 polymorphic SSR and 19 quantitative traits in 353 genotypes of barley from 19 breeding programs of barley distributed worldwide. We found nine traits under overall disruptive selection, three traits under disruptive and stabilizing selection in different pairs of breeding programs, and seven neutral traits. By simultaneously estimating population structure at morphological traits and neutral molecular markers, we are able to detect compatible breeding programs for germplasm exchange, and breeding programs with unique characteristics worth preserving. Additionally, we found some traits under divergent selection in both cultivated and wild species. However, caution is advised in the interpretation of results for fitness-related traits, traits with important GxE, and for strong artificially-imposed genetic structure, especially in the artificial selection context of crop improvement breeding programs.
Introduction

Genetic diversity is essential for breeding purposes and for developing more sustainable agricultural systems (Brummer, 1998; Duvick et al., 2004; Stuthman, 2002). Breeding programs require genetic diversity as both, an insurance against unforeseeable changes in the environment (i.e. diseases, attacks by pest, and changing climate), and to maintain genetic progress (Brown-Guedira et al., 2000). However, finding a meaningful measure of genetic diversity is challenging (Kim and Ward, 2000; Purvis and Hector, 2000). Different types of data have been used to attempt genetic characterization, including morphology (Ortiz et al., 2002), molecular markers (Donini et al., 2000; Kim and Ward, 2000; Koebner et al., 2003; Malysheva-Otto et al., 2007; Ordon et al., 2005; Russell et al., 1997), and pedigree information (Cox et al., 1985; Delannay et al., 1983; Rasmusson and Phillips, 1997; Smith et al., 2004). Gepts (2006) states that genetic resources conservation and utilization will be facilitated by the understanding of the genotypic basis of agriculturally important traits, and specifically by understanding the evolutionary, ecological and anthropic mechanisms that led to the current characteristics of the traits. He even goes further to say that understanding the distribution of genetic diversity within and among populations (i.e. the population structure) has become the ‘holy grail’ of the science of genetic resources.

The form of selection of quantitative traits can be studied by the comparison of the population structure at quantitative traits against that of presumably neutral molecular markers (McKay and Latta, 2002; Merilä and Crnokrak, 2001). Population structure at neutral molecular markers has been traditionally studied with the statistic $F_{ST}$ (Wright, 1951). $F_{ST}$ estimates the amount of genetic diversity allocated among populations compared to the total amount of genetic diversity present in the populations. Several estimators of $F_{ST}$ have been developed to be used in slightly different systems including $G_{ST}$ (Nei, 1972; 1973; 1978), $\theta_{ST}$ (Cockerham and Weir, 1983), $\Phi_{ST}$ (Excoffier et al., 1992), and $R_{ST}$ (Slatkin, 1993), among others. More recently, $Q_{ST}$ was developed as an estimator of population structure using quantitative trait measurements (Spitze, 1993). Quantitative genetics theory shows that for neutral additive traits, $Q_{ST} = F_{ST}$ (McKay and
Furthermore, $Q_{ST} > F_{ST}$ is expected in traits that are under divergent or disruptive selection, and $Q_{ST} < F_{ST}$ is expected for traits under stabilizing selection. This methodology has been widely used in natural populations (see review by Leinonen et al., 2008). We propose the use of the comparison of $Q_{ST}$ and $F_{ST}$ in cultivated species to understand the evolutionary mechanisms that drove agriculturally relevant traits. The main objective of this research was to apply this methodology to barley as a case study, to understand the evolutionary history of agriculturally relevant traits in barley. First, we tested the hypothesis of neutral evolution of 19 traits. Our main hypothesis was that some of the agriculturally important traits would be under strong divergent selection (e.g., life-history traits), while some would be under stabilizing selection (e.g., some morphological traits) and others would be neutral (e.g., other morphological traits). Second, we described the patterns of divergence across 23 breeding programs distributed worldwide. Third, we compared the patterns of selection in the traits with those of wild barley. Finally, we discuss the relevance of this methodology for breeding purposes, genetic resources conservation and utilization purposes, and for understanding the domestication process.

**Materials and Methods**

**Materials**

A total of 353 inbred lines of barley from 23 private and public breeding programs were evaluated. Each breeder responsible for a breeding program was asked to provide 20 advanced inbred lines or recently released cultivars of barley that represented current diversity in their program. Two-row and six-row types were treated separately, and if a breeding program included both, separate samples were requested of each type. The programs that provided seed were from the United States (Washington State University-2-row; University of Minnesota-6-row; 2-row and 6-row programs of North Dakota State University; USDA-ARS-Aberdeen-Idaho-2-row; 2-row, 6-row and international programs of Busch Agricultural Resources), Canada (University of Saskatchewan-2-row; Alberta Agriculture, Food and Rural Development-6-row), Europe (Saatzucht Josef Breun in Germany-2-row; Svalöf Weibull Ab in Sweden-2-row; The Abed Foundation in
Denmark-2-row; the spring and winter programs from Osijek Agricultural Institute of Croatia-2-row), Australia (University of Adelaide-2-row; Western Australia Department of Agriculture-2-row), and South America (National Agronomic Research Institutes of Chile-2-row and Uruguay-2-row). Additionally, five checks (Dayman, Perun, Ceibo, Clipper, and Quebracho) were evaluated to complete the experimental design and to control for heading dates. Data on checks is not shown.

**Molecular Marker Analysis**

Eighty SSR markers of cultivated barley (Becker and Heun, 1995; and Smith K. pers. com.; Liu et al., 1996; Ramsay et al., 2000; Saghai Maroof et al., 1994; Struss and Plieske, 1998) were screened in all genotypes. After discarding monomorphic and inconsistently amplifying loci, data on 66 polymorphic markers is reported.Markers were chosen to have good genome coverage. Plant tissue collection, DNA extraction, PCR amplifications, and SSR scoring was conducted according to (Gutierrez et al., 2008).

**Field Trials**

All lines and checks were evaluated in a row-column design with 3 replications at 2 locations in Uruguay as reported elsewhere (Gutierrez et al., 2008). Several morphological traits were recorded for each line either on plots or on individual plants. The phenotypic traits measured on a plot level were: total number of tillers (counted at the end of tillering, NTIL), spot blotch disease (scoring from 1-5, where 1 is low and 5 is high, SB), leaf rust disease (scoring from 1-5, where 1 is low and 5 is high, LR), powdery mildew disease (scoring from 1-5, where 1 is low and 5 is high, PM), total number of days between planting date and anthesis (DTA), total number of days between planting and flowering (DTF), plant height (measured from the ground to the average of the tips of the spikes, in cm, HTOT), biomass weight at plant maturity (g, BWT), grain yield (g m$^{-2}$, YLD), total number of spikes (NSPK), test weight (measured on a volume of 6mL, kg m$^{-3}$, TWT), weight of 100 kernels (g, W100G).
Five plants from each plot were chosen at random at flowering time, and were color-marked with plastic twist-bands. Measurement of traits on single plants instead of whole plots decreased the experimental error and allowed estimation of within plot variation. The following phenotypic traits were measured on each individual plant: flag leaf length (measured from the ligulae to the tip of the leaf, in cm, FLL), flag leaf width (measured at 2.5 cm from the ligulae, in cm, FLW), spike length (measured from the base to the tip of the spike, without counting the awns, in cm, SLT), awn length (measured from the top of the spike to the tip of the longest awn, in cm, ALT), flag leaf height (measured from the ground to the ligulae of the flag leaf, in cm, FLH), spike height (measured from the ground to the base of the spike, in cm, SHT), and number of grains per spike (NG).

### Statistical Models

Traits measured at the plot- and plant-levels were modeled according to the following linear models respectively:

\[
Y_{ijklmn} = B_{ij} + R_{k(ij)} + C_{l(ij)} + P_{m} + G_{n(m)} + I_{o(ijk)} + \varepsilon_{ijklmn},
\]  

\[
Y_{ijklmno} = B_{ij} + R_{k(ij)} + C_{l(ij)} + P_{m} + G_{n(m)} + I_{o(ijk)} + \varepsilon_{ijklmno},
\]

where \(B_{ij} = \) effect of \(j^{th}\) block in environment \(i\), \(R_{k(ij)} = \) effect of \(k^{th}\) row on \(ij^{th}\) block-environment, \(C_{l(ij)} = \) effect of \(l^{th}\) column on \(ij^{th}\) block-environment, \(P_{m} = \) effect of \(m^{th}\) breeding program, \(G_{n(m)} = \) effect of \(n^{th}\) genotype on \(m^{th}\) breeding program, \(I_{o(ijk)} = \) effect of the \(o^{th}\) plot on the \(ij^{th}\) block-environment, \(\varepsilon_{ijklmn} = \) residual error for the \(n^{th}\) genotype on the \(m^{th}\) breeding program of the \(ij^{th}\) block-environment, and \(\varepsilon_{ijklmno} = \) residual error for the \(o^{th}\) plant of the \(n^{th}\) genotype on the \(m^{th}\) breeding program of the \(ij^{th}\) block-environment. Plants within a plot share similarities from belonging to the same plot, therefore, the plot effect \((I_{o(ijk)})\) was included in the linear model for plant-level variables (equation 2).

We defined a hierarchical Bayesian model where we modeled both means \((B_{ij}, R_{k(ij)}, C_{l(ij)}, P_{m}, G_{n(m)}, \text{ and } I_{o(ijk)})\) and associated variances \((\sigma_{R}^2, \sigma_{C}^2, \sigma_{P}^2, \sigma_{G(m)}^2, \sigma_{I}^2, \text{ and } \sigma^2)\) in terms of explanatory variables (block, location, row, column, breeding program,
At the first level of the Bayesian hierarchy, observations were modeled as independent samples from a normal distribution:

\[
Y_{ijklmn} \mid B_{ij}, R_{k(ij)}, C_{l(ij)}, P_m, G_{n(m)}, \sigma^2 \sim N\left(B_{ij} + R_{k(ij)} + C_{l(ij)} + P_m + G_{n(m)}, \sigma^2 \right)
\]

\[
Y_{ijklmno} \mid B_{ij}, R_{k(ij)}, C_{l(ij)}, P_m, G_{n(m)}, I_{o(ijk)}, \sigma^2 \sim N\left(B_{ij} + R_{k(ij)} + C_{l(ij)} + P_m + G_{n(m)} + I_{o(ijk)}, \sigma^2 \right).
\]

The second level of the Bayesian hierarchy includes prior distributions for location parameters (i.e. means) \(B_{ij}, R_{k(ij)}, C_{l(ij)}, P_m, G_{n(m)}\), and observational variance \(\sigma^2\). Priors on all location parameters were normal with mean zero and variances defined to condition the desired level of information sharing among levels of the factor.

For block-environment means, \(B_{ij}\), the prior was defined with a very large variance to make the prior non-informative: \(B_{ij} \sim N(0, 10^{-7})\). The flat and independent prior is a Bayesian equivalent to defining the block effect as a fixed effect in classical linear models. We did not include location effects in our model so that all parameters were estimable.

Row, column, breeding program, genotype and plot were modeled with priors that treated them as equivalent to random effects in classical mixed linear models. Row, column, population, genotype and plot effects were modeled as samples from a normal distribution with variance \(\sigma^2_R, \sigma^2_C, \sigma^2_P, \) and \(\sigma^2_I\), respectively: \(R_{k(ij)} \mid \sigma^2_R \sim N(0, \sigma^2_R), C_{l(ij)} \mid \sigma^2_C \sim N(0, \sigma^2_C), P_m \mid \sigma^2_P \sim N(0, \sigma^2_P), G_{n(m)} \mid \sigma^2_G \sim N(0, \sigma^2_G), I_{o(ijk)} \mid \sigma^2_I \sim N(0, \sigma^2_I)\). Priors on the variance of row, column, breeding program and genotype effects were chosen to be non-informative as: \(\sigma^2_R, \sigma^2_C, \sigma^2_P, \) and \(\sigma^2_G \sim IG(0.0001,0.0001)\). Residual variance was also modeled as \(\sigma^2 \sim IG(0.0001,0.0001)\).

All parameters were obtained via Markov Chain Monte Carlo simulation using the Bayesian Gibbs Sampling software WINBUGS (Spiegelhalter et al., 2003).

**FST Estimation**

Population differentiation was measured in terms of \(F_{ST}\) as estimated by Weir and Cockerham’s (Weir and Cockerham, 1984) \(\theta\) in GDA (Lewis and Zaykin, 2002):
\[ F_{ST} = \frac{V_B}{V_B + V_W}, \]  

where \( V_B \) is the among-population variance component of the genetic marker, and \( V_W \) is the within population variance component. Confidence intervals were obtained through parametric bootstrapping with 1,000 replications.

We distinguished two types of estimates of population differentiation: overall estimates (\( F_{ST} \)), and pair-wise estimates (\( F_{ST}^{ij} \)). \( F_{ST} \) was estimated for all populations and all genotypes within populations; therefore, a single \( F_{ST} \) estimate was obtained. Whereas \( F_{ST}^{ij} \) were estimated for all pairs of populations and therefore, a 23x23 matrix of \( F_{ST}^{ij} \) estimates was obtained.

**QST Estimation**

Among population differentiation (\( Q_{ST} \)) was estimated in WINBUGS1.4 (Spiegelhalter et al., 2003) through a variance partition analysis following Bonnin et al (1996):

\[ Q_{ST} = \frac{(1+f) V_B}{(1+f) V_B + 2V_W}, \]  

where \( f \) is the inbreeding coefficient, \( V_B \) is the among population component of variance for each trait (i.e. \( \sigma^2_P \)), and \( V_W \) is the additive genetic variance component within populations (i.e. \( \sigma^2_G \)). Since barley is a selfing species we used an \( f=1 \).

\( Q_{ST} \) estimates and their 95% confidence intervals were obtained from the posterior distribution of \( Q_{ST} \) using breeding program (\( \sigma^2_P \)) and genotype (\( \sigma^2_G \)) variance as estimates of \( V_B \) and \( V_W \) respectively, obtained from the linear models described in equation 1 and 2 where \( \sigma^2_P \) is the variance associated with breeding program effect (\( P_m \)) and \( \sigma^2_G \) the variance associated with genotype effect (\( G_{n(m)} \)). The model was fitted by Markov chain Monte Carlo, using WinBUGS1.4 (Spiegelhalter et al., 2003). Two chains were run, and after the chains converged (5,000 iterations), the next 10,000 iterations were taken from each chain.
$Q_{ST}^{ij}$ estimates were obtained using a similar approach. One data set for each pair of populations was created and then analyzed with the same models described previously (equations 1 and 2) but excluding row and column effects to make all parameters estimable.

**$Q_{ST}$-$F_{ST}$ Comparisons**

Comparisons of population differentiation were conducted at two levels: overall and pair-wise. Overall-estimates of $F_{ST}$ and $Q_{ST}$ for each trait were compared in WinBUGS1.4 (Spiegelhalter et al., 2003) to detect deviation from neutrality of each trait. The full posterior distribution of $Q_{ST}$ was obtained as described above. The distribution of $F_{ST}$ was obtained in WinBUGS1.4 by assigning $F_{ST}$ a normal distribution with parameters equal to those obtained from GDA estimation. The posterior distribution of the difference of $F_{ST}$ and $Q_{ST}$ ($\Delta = Q_{ST} - F_{ST}$) was sampled for each trait in WinBUGS1.4 (Spiegelhalter et al., 2003) and mean values and 95% confidence intervals for $\Delta$ were obtained. Each trait was classified as either neutral ($\Delta = 0$, $Q_{ST} = F_{ST}$), under divergent selection ($\Delta < 0$, $Q_{ST} > F_{ST}$), or under stabilizing selection ($\Delta > 0$, $Q_{ST} < F_{ST}$).

We also used the methodology proposed by Gutierrez et al. (2008) to detect deviations from neutrality when overall $Q_{ST}$-$F_{ST}$ comparison failed to reject the hypothesis of neutrality. Analysis with overall estimates of $Q_{ST}$ and $F_{ST}$ may fail to reject the hypothesis of neutrality when some pairs of populations are under stabilizing and other pairs are under divergent selection such that these effects cancel out. The approach consisted of the studying pair-wise comparisons by simulating phenotypes with the same structure as the molecular markers, and therefore that represent the expected phenotypic structure under a neutral expectation. The observed structure of the phenotypes is then compared to the expected structure obtained through the simulation approach. Properties of the observed distribution of the relationship between $Q_{ST}^{ij}$ and $F_{ST}^{ij}$ are used to find significant deviations from the simulated neutral expectation. We used this approach with 10,000 simulations in all the traits that had an overall $Q_{ST} = F_{ST}$. Analyses were
performed in GENALEX (Peakall and Smouse, 2006) and R (Ihaka and Gentleman, 1996).

Matrices representing $\Delta_{ij}$, $(Q_{ST}^{ij} - F_{ST}^{ij})$ were created for each of the traits where selection was detected. For the traits that were misclassified as neutral, Bonferroni corrected p-values were used to test deviations from neutrality in each pair (Gutiérrez et al., 2008). Pairs of breeding programs that are significantly under divergent or stabilizing selection ($p < 0.05$) were represented in color-coded matrices. For the traits under divergent selection, the value of $\Delta_{ij}$ is represented as an indication of the strength of the selection in the color-coded matrices.

We first describe all $Q_{ST}-F_{ST}$ comparisons in a classical way, using the interpretations proposed elsewhere where $Q_{ST} > F_{ST}$ implies divergent selection, $Q_{ST} < F_{ST}$ implies stabilizing selection, and $Q_{ST} = F_{ST}$ implies neutral evolution (McKay and Latta, 2002; Merilä and Crnokrak, 2001). These interpretations were developed for wild populations undergoing natural selection and do not necessarily reflect the true mechanisms affecting quantitative traits when artificial selection is conducted. Therefore, even though we first describe the results naively in a literal way, in the discussion section we address some of the limitations of those interpretations and we attempt explanations for the patterns found in cultivated species under artificial selection.

**Results**

Overall $F_{ST}$ was significantly greater than zero ($p < 0.05$) 0.368. While $F_{ST}^{ij}$ ranged from 0 to 0.691 with an average of 0.268. Overall $Q_{ST}$ was significantly greater than zero ($p < 0.05$) for all the traits. It ranged from 0.207 in spike length to 0.774 in flag leaf width (Table 1).

Ten traits were classified as neutral by overall statistics: number of tillers, spot blotch disease, leaf rust disease, weight of 100 grains, spike length, awn length, flag leaf height, biomass, grain yield, and test weight (Table 1).

Three traits that were classified as neutral by overall statistics were not neutral when pair-wise relationships among breeding programs were considered: biomass, grain yield,
and test weight (Table 1). Chile, Croatia-spring, Croatia-winter, Denmark, Busch-two-row, North Dakota-two-row, Washington and Uruguay breeding programs were diverging from all the other programs for biomass (Figure 1). On the other hand, some pairs of breeding programs were converging (subject to stabilizing selection): Busch-six-row with Germany, Sweden, and Busch-international; and North Dakota-six-row with Sweden and Busch-international. For grain yield we found Chile, Croatia-spring, Croatia-winter and Idaho diverging from other programs (Figure 1), while other breeding programs were converging: Busch-six-row with Western Australia, Canada Saskatchewan, Denmark, Germany, Busch-two-row, North Dakota-two-row, and Washington; and Minnesota with Canada Alberta and Busch-international. Canada Alberta, Chile, North Dakota-two-row, and Uruguay were diverging for different optima in test weight (Figure 1) while Busch-six-row was converging with Canada Saskatchewan and Denmark, and Denmark was also converging with Minnesota.

Nine traits were under divergent selection: powdery mildew disease, days until anthesis, days until flowering, plant height, number of spikes, flag leaf length, flag leaf width, spike height, and number of grains (Table 1). For some traits, all breeding programs seemed to diverge equally from each other (i.e. each breeding programs had their optimum): flag leaf length and width, spike height, and number of grains (Figure 2). For other traits a few of the breeding programs were similar but diverged from the rest: Australia Adelaide and the Croatia-winter for days until anthesis and flowering; Denmark for plant height; Croatia-spring and Germany for number of spikes; and Chile, Denmark, and Busch-international for powdery mildew. We only found very few programs that experienced stabilizing selection for these traits: Busch-six-row and Minnesota for days until anthesis and flowering were examples of such selection.

Discussion

We found several traits under overall disruptive selection, some traits under disruptive and stabilizing selection in different pairs of breeding programs, and some neutral traits. However, we did not find any trait under overall stabilizing selection.
Different traits might be under different selection pressures, and in consequence have different $Q_{ST}-F_{ST}$ relationships (see McKay and Latta, 2002 for a review). For instance, life-history traits are predicted to be under strong directional selection, while morphological traits are expected to be under weak or stabilizing selection (Merilä and Sheldon, 1999; Rieseberg et al., 2002). However, very few empirical studies have been able to detect stabilizing traits (Edmands and Harrison, 2003); most studies report more traits under divergent selection (McKay and Latta, 2002; Merilä and Crnokrak, 2001; Morgan et al., 2001; Palo et al., 2003).

Some breeding programs were found to be diverging from the other programs for many traits: Chile, Croatia-spring, Croatia-winter, and Denmark. Chile, Croatia-spring and Croatia-winter were identified as breeding programs with unusual morphological characteristics in Gutierrez et al (2008). This overall pattern of divergence is relevant as a general characteristic of the breeding programs. For example, breeding programs that are diverging for most of their traits would probably have unique germplasm that is worth preserving in germplasm banks. Additionally, pairs of breeding programs that are under stabilizing selection could potentially benefit from germplasm exchange because they will be phenotypically similar, meaning that they might be adapted to similar environmental conditions (but see discussion below) while at the same time being genotypically different. This combination raises the possibility that they might have different alleles at the loci of interest. This type of analysis could be useful for non-fitness traits that are relevant for breeding purposes because the sustainability of the system is considered (Brummer, 1998; Duvick et al., 2004; Stuthman, 2002). For example it can be used for disease resistance (Stuthman, 2002). However, careful consideration should be made in the study of fitness-related traits, germplasm with strong artificially-imposed genetic structure, and genotype-by-environment (GxE) interaction.

We found the following pairs of breeding programs under ‘stabilizing selection’ for grain yield (i.e. $Q_{ST} < F_{ST}$): Busch-six-row with Western Australia, Canada Saskatchewan, Denmark, Germany, Busch-two-row, North Dakota-two-row, and Washington; and Minnesota with Canada Alberta and Busch-international. However,
given that grain yield and all traits closely related to fitness are expected to be under strong directional selection (Merilä and Sheldon, 1999), we question whether the result that $Q_{ST} < F_{ST}$ is truly due to what is understood as stabilizing selection. Two circumstances are relevant here, first, that the traits observed can be presumed to be under directional selection, and second, that in the crop improvement context, strong genetic structure can be maintained between lineages that nevertheless coexist in similar environments. For completeness, we first define stabilizing selection. We then propose three alternative possible mechanisms that could generate observations mimicking stabilizing selection (i.e., causing $Q_{ST} < F_{ST}$) in the crop improvement context. These explanations are strong artificially-imposed genetic structure, genotype by environment interaction (GxE), and a sub-category of GxE that we call the *Anna Karenina syndrome*.

By stabilizing selection, we mean that both that the trait displays similar values in the two populations and that the trait has maintained a steady optimum through the genetic divergence of the populations. We contrast this image of stabilizing selection to one where a trait has similar values in populations that are separated by very great genetic distance, as exemplified by two- and six-row barley. In the crop improvement context, gene exchange is strongly imposed because there is a specific reason to avoid gene exchange. In particular two- and six-row barleys have different uses in the world. Most breeding programs that breed for malting quality use two-row barleys (except USA and Mexico) and avoid gene flow from six-row germplasm because of poorer malting quality resulting from its smaller grains. Therefore, two- and six-row types are rarely intermated. This barrier to gene flow has created a strong genetic structure such that $F_{ST}$ is quite high. At the same time, one can argue that a trait such as yield could be considered a different trait in the two-row than in the six-row context. If we allow a loose definition, comparable yields across two- and six-row populations could be considered as a form of homoplasy (Futuyma, 1986), or perhaps more accurately, as parallel evolution that occurs when similar traits evolved independently in closely related lineages (Doolittle, 1994). Additionally, the strong resemblance among grain yields in the two distinct groups could be a consequence of both reaching similar yield plateaus: gain may continue for both but
only in small increments. This would produce a seemingly similar yield in all the
breeding programs regardless of whether it is truly “optimal” or not. Most pairs of
breeding programs that were found to be under stabilizing selection were pairs consisting
of a two-row and a six-row population. We therefore think that this conjunction of strong
genetic divergence with phenotypic resemblance due to reasons other than selection for
an optimal value may explain some of our observations of $Q_{ST} < F_{ST}$.

A second mechanism causing $Q_{ST} < F_{ST}$ in yield could be GxE. Studies involving
overall $Q_{ST}$ estimations are robust to GxE, however, pair-wise estimates are more
sensitive to GxE (Palo et al., 2003). Genotype-by-environment interaction is a necessary
but not sufficient condition to cause $Q_{ST}$ values to diverge when traits are evaluated in
different environments. That is, even if large rank changes occur in measurements in
different environments, the overall between-population variance may be stable. Such
stability will generally not be the case for pair-wise differences, such that pair-wise $Q_{ST}$
across environments are uninterpretable in the presence of GxE. This situation can be
illustrated simplistically by assuming two populations whose performance is opposite in
two environments (e.g., P1-E1 = 10, P1-E2 = 20, P2-E1 = 20, P2-E2 = 10). When $Q_{ST}$ is
computed for the combination of the environments, the populations appear to be under
stabilizing selection, even though they are diverging in each environment. Poor estimates
of pair-wise $Q_{ST}$ are therefore obtained in the presence of GxE. Viewed from the
perspective of genetic correlation between traits, we can also see that strong GxE could
lead to the appearance of neutral evolution, even when a trait is under strong selection.
Such GxE is equivalent to a low genetic correlation between trait values as measured in
the common garden environment where the research takes place versus the home
environment in which selection is taking place. That low genetic correlation in turn
means that the trait, as measured in the common garden, is evolving neutrally because
nowhere is it under selection.

Finally, the third explanation for $Q_{ST} < F_{ST}$ in yield is a specific case of GxE that we
call the Anna Karenina syndrome. In the non-target environment, genotypes could
perform poorly for different reasons (Gutierrez et al., 2008). For example, one set of
genotypes could be limited because of disease pressure, while the other could be limited by photoperiod conditions. We called this mechanism the *Anna Karenina syndrome* because it exemplifies the famous phrase that opens Tolstoy’s novel *Anna Karenina* (Tolstoy, 1998) ‘all happy families resemble one another; each unhappy family is unhappy in its own way’. High yielding genotypes are similar; each poorly performing genotype performs poorly in its own way. Even though the two sets of poorly performing genotypes seem to be converging because they produce similar yields in the common garden environment, they are clearly not under stabilizing selection for yield. If the genotypes were evaluated in an environment without disease pressure but with photoperiod limitations, it would be obvious that they perform differently. Therefore, finding \( Q_{ST} < F_{ST} \) could be a consequence of genotypes performing poorly but for different reasons.

Comparing the evolutionary history of these traits with their selection history in the ancestor of barley, wild barley (*Hordeum spontaneum* K Koch) could bring us an understanding of the patterns toward domestication of the species. Domestication of barley started 10,000 years ago causing the divergence of cultivated barley from wild barley (Harlan and Zohary, 1966). The single most important trait during the domestication of barley was shattering of the seed (Zohary and Hopf, 2000), where alleles reducing shattering were selected when humans started to harvest the seed. However, several other traits were presumably also involved in the divergence between the species. By comparing the selection history of barley with that of its ancestor, we may be able to identify traits that contributed to the domestication process. We compared results presented here on cultivated barley to those obtained by Gutierrez et al (2008) in wild barley. The most biologically interesting and statistically feasible comparisons between species are the corners of the grid formed by the three possible types of selection affecting each trait in each species (i.e. stabilizing, neutral, and divergent selection, Figure 3). We named each cell according to its characteristics. *Adaptation traits* are traits under divergent selection in both the wild and the cultivated species (\( Q_{ST} > F_{ST}, Q_{ST} > F_{ST} \)). They may confer local adaptation advantages to the individuals both in wild species
and cultivated species, and are relevant for breeding purposes. We found days until anthesis, days until flowering, grain weight, and number of grains per spike in this category. Although we did not find any trait in the remaining categories, we describe them for completeness. Domestication pre-adaptation traits are traits that were under stabilizing selection in the wild species and are under stabilizing selection in the domesticated species ($Q_{ST} < F_{ST}$, $Q_{ST} < F_{ST}$). If they have the same optimum, selection for the same phenotypes occurred throughout the populations. These traits could be considered pre-adapted to selection because, as considered by Bock (1959), in a phyletic line, post-adapted changes in the ancestral species will be the pre-adaptive changes in the descendant species. Domesticity traits are traits that were under directional selection in the wild (i.e. conferring local adaptations), and are now under stabilizing selection ($Q_{ST} > F_{ST}$, $Q_{ST} < F_{ST}$). Domestic niche traits are traits that were under stabilizing selection in the wild species and are now under directional selection ($Q_{ST} < F_{ST}$, $Q_{ST} > F_{ST}$). This could occur if local adaptations provide advantages for breeding characteristics (e.g. more spikes are favored in one environment while longer spikes in another). Additionally, we found some traits to be neutral in one or both species. Number of tillers and leaf rust disease were neutral in both species. Number of spikes and flag leaf length were neutral in the wild ancestor, but were under directional selection in cultivated barley. Finally, weight of 100 grains, spike and awn length were neutral in the cultivated barley but were under divergent selection in wild barley. We failed to detect any trait under overall stabilizing selection.

This methodology could be used to facilitate genetic resource conservation and utilization and to facilitate germplasm exchange among breeding programs. More efficient methods for maintaining and evaluating genetic diversity in germplasm banks are needed (Holbrock and Dong, 2005). Traditionally, germplasm banks used phenotypic data for documenting their diversity (Crossa et al., 1994; Malosetti and Abadie, 2001; Taba et al., 1998; Upadhyaya and Ortiz, 2001; Vaughan, 1991). In the last decades, the use of molecular markers has become extremely common, as a result of technological improvements that caused them to be cheaper and faster. In consequence, molecular
markers substituted in some cases (Chavarriaga-Aguirre et al., 1998; Struss and Plieske, 1998), and complemented in others (Elias et al., 2000; Ghislain et al., 2004; Hokanson et al., 1998; Skroch et al., 1998; Staub et al., 2002; Ude et al., 2004) the use of morphological information to construct core collections and genetic resource classification. Given the widespread use of molecular markers, and the already available information on phenotypic traits in germplasm banks, the comparison of molecular markers to phenotypic traits methodology could be readily used in germplasm banks to understand the evolutionary history of agriculturally relevant traits. Furthermore, understanding the structure of the diversity in the populations at both levels will improve the allocation of conservation efforts, and will facilitate the utilization by identifying compatible accessions or populations (Gepts, 2006) and distinct sources of germplasm at both levels. Additionally, the comparison of population structure at different breeding programs can facilitate germplasm exchange among breeders. Breeders need elite materials with new alleles at the loci of interest (Gutierrez et al., 2008). By simultaneously estimating population structure at morphological traits and neutral molecular markers, we are able to detect compatible breeding programs for germplasm exchange such that they share local adaptations but have different alleles at the neutral loci, and therefore could potentially have new alleles at the specific loci of interest. However, caution is advised in the use of this methodology for fitness-related traits, traits with important GxE, and with strong artificially-imposed genetic structure, especially in artificial selection context.

References


List of Figures

Figure 1. Pair-wise relationships among breeding programs for the traits that were classified as neutral by overall analysis and as under selection by pair-wise analysis: a) grain yield (above diagonal) and biomass (below diagonal), and b) test weight (below diagonal). Significant deviations from neutrality (p < 0.05) are represented in black (divergent selection) and white (stabilizing selection).

Figure 2. Pair-wise relationship among breeding programs for the traits under selection: a) Days until anthesis (below diagonal) and days until flowering (above diagonal), b) flag leaf length (below diagonal) and flag leaf width (above diagonal), c) plant height (below diagonal) and spike height (above diagonal), d) number of spikes (below diagonal) and...
number of grains per spike (above diagonal), and e) powdery mildew disease (below diagonal).

**Figure 3.** Chart showing the possible categories of traits that we could find for the comparison of selection pressures across the domesticated and cultivated species.
Table 1. Overall and pair-wise comparison of $Q_{ST}$ and $F_{ST}$: $Q_{ST}$ estimate and its standard error in parenthesis (s.e.), significance of the difference between $Q_{ST}$ and $F_{ST}$ ($Q_{ST} ≠ F_{ST}$) and $Q_{STij}$ and $F_{STij}$ ($Q_{STij} ≠ F_{STij}$).

<table>
<thead>
<tr>
<th>Neutral traits</th>
<th>$Q_{ST}$ (s.e.)</th>
<th>$Q_{ST} ≠ F_{ST}$</th>
<th>$Q_{STij} ≠ F_{STij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Number of Tillers</td>
<td>0.549 (0.089)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>$Q_{STij} = F_{STij}$</td>
</tr>
<tr>
<td>3 Spot blotch disease</td>
<td>0.258 (0.078)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>$Q_{STij} = F_{STij}$</td>
</tr>
<tr>
<td>4 Leaf rust disease</td>
<td>0.370 (0.088)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>$Q_{STij} = F_{STij}$</td>
</tr>
<tr>
<td>13 Weight of 100 grains</td>
<td>0.383 (0.089)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>$Q_{STij} = F_{STij}$</td>
</tr>
<tr>
<td>16 Spike length</td>
<td>0.207 (0.075)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>$Q_{STij} = F_{STij}$</td>
</tr>
<tr>
<td>17 Awn length</td>
<td>0.288 (0.086)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>$Q_{STij} = F_{STij}$</td>
</tr>
<tr>
<td>18 Flag leaf height</td>
<td>0.355 (0.113)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>$Q_{STij} = F_{STij}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Traits under stabilizing and disruptive selection</th>
<th>$Q_{ST}$ (s.e.)</th>
<th>$Q_{ST} ≠ F_{ST}$</th>
<th>$Q_{STij} ≠ F_{STij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 Biomass weight</td>
<td>0.315 (0.090)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>$Q_{STij} ≠ F_{STij}$</td>
</tr>
<tr>
<td>10 Grain yield</td>
<td>0.292 (0.086)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>$Q_{STij} ≠ F_{STij}$</td>
</tr>
<tr>
<td>12 Test weight</td>
<td>0.254 (0.078)</td>
<td>$Q_{ST} = F_{ST}$</td>
<td>$Q_{STij} ≠ F_{STij}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Traits under disruptive selection</th>
<th>$Q_{ST}$ (s.e.)</th>
<th>$Q_{ST} &gt; F_{ST}$</th>
<th>$Q_{STij} ≠ F_{STij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Powdery mildew disease</td>
<td>0.586 (0.085)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td></td>
</tr>
<tr>
<td>6 Days until anthesis</td>
<td>0.609 (0.084)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td></td>
</tr>
<tr>
<td>7 Days until flowering</td>
<td>0.604 (0.084)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td></td>
</tr>
<tr>
<td>8 Plant height</td>
<td>0.587 (0.086)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td></td>
</tr>
<tr>
<td>11 Number of spikes</td>
<td>0.621 (0.083)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td></td>
</tr>
<tr>
<td>14 Flag leaf length</td>
<td>0.611 (0.083)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td></td>
</tr>
<tr>
<td>15 Flag leaf width</td>
<td>0.774 (0.061)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td></td>
</tr>
<tr>
<td>19 Spike height</td>
<td>0.575 (0.088)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td></td>
</tr>
<tr>
<td>21 Number of grains per spike</td>
<td>0.754 (0.064)</td>
<td>$Q_{ST} &gt; F_{ST}$</td>
<td></td>
</tr>
</tbody>
</table>

$^\dagger$ $Q_{ST} = F_{ST}$ indicates that there is not a significant difference (p = 0.05) among $Q_{ST}$ and $F_{ST}$; $Q_{ST} > F_{ST}$ indicates that $Q_{ST}$ is significantly greater than $F_{ST}$ (p = 0.05).

$^\dagger$ $Q_{STij} = F_{STij}$ indicates that there is not a significant deviation from the neutral expectation (p < 0.05) when pairs of breeding programs are studied; $Q_{STij} ≠ F_{STij}$ indicates that there is a significant deviation from neutrality (p < 0.05) when pairs of breeding programs are studied.
Figure 1. Pair-wise \( Q^{ST}_{FST} \) for non-neutral traits

- AU-AD (2R)
- AU-WE (2R)
- CA-AB (6R)
- CA-SK (2R)
- CL (2R)
- CR-SP (2R)
- CR-WI (2R)
- DE (2R)
- GE (2R)
- SW (2R)
- US-BS (2R)
- US-BS (6R)
- US-BSI (2R)
- US-ID (2R)
- US-MN (6R)
- US-ND (2R)
- US-ND (6R)
- US-WS (2R)
- UY (2R)
Figure 2. Pair-wise QSTij-FSTij for divergent traits

AU-AD (2R) - AU-WE (2R) - CA-AB (6R) - CA-SK (2R) - CL (2R) - CR-SP (2R) - CR-WI (2R) - DE (2R) - GE (2R) - SW (2R) - US-BS (2R) - US-BS (6R) - US-BSI (2R) - US-ID (2R) - US-MN (6R) - US-ND (2R) - US-ND (6R) - US-WS (2R) - UY (2R)
<table>
<thead>
<tr>
<th>cultivated barley</th>
<th>$Q_{ST} &gt; F_{ST}$</th>
<th>$Q_{ST} = F_{ST}$</th>
<th>$Q_{ST} &lt; F_{ST}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adaptation Traits</td>
<td></td>
<td>Domestic-niche Traits</td>
</tr>
<tr>
<td></td>
<td>$Q_{ST} = F_{ST}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Domesticity Traits</td>
<td></td>
<td>Domestic Pre-adaptation Traits</td>
</tr>
</tbody>
</table>

**Figure 3.** Classification of traits for study of domestication
CHAPTER VI: GENERAL CONCLUSIONS

We demonstrated that overall statistics such as $Q_{ST}$ and $F_{ST}$ can fail to detect selection when types of selection are heterogeneous across populations. Indeed, selection mosaics with subsets of populations under disruptive selection ($Q_{ST} > F_{ST}$) and subsets of populations under stabilizing selection ($Q_{ST} < F_{ST}$) can lead to a trait being misclassified as neutral. We developed a methodology using a combination of overall and pair-wise analyses that is capable of detecting those selection mosaics in both simulated and empirical data sets. Our results also demonstrated that the approach is capable of detecting such selection mosaics under a range of population sizes, levels of inbreeding, and number of populations. Furthermore, we provided ample evidence of the occurrence of this phenomenon in both cultivated and wild barley. Therefore, reanalyzing the many studies that failed to reject the hypothesis of neutrality could bring new insights into the true evolutionary patterns that shape populations.

The combined use of overall and pair-wise statistics could also be used to gain a better understanding of the selection process. Pair-wise comparisons can be used in traits under selection to identify: 1) cases where selection and gene flow patterns reinforced each other, such as those found in clinal variation; 2) populations responsible for the $Q_{ST} > F_{ST}$ pattern (i.e., driving populations); and 3) groups of populations that tend to have $Q_{ST} < F_{ST}$ for comparisons within groups, and $Q_{ST} > F_{ST}$ for comparisons between groups (i.e., homogenizing groups). While the ad hoc approach we used improved the description of inter-population evolutionary forces, a more formal, statistically justified test is still needed.

Germplasm exchange among breeding programs could be facilitated by the data-driven methodology we developed. We produced groups of breeding programs with similar performance and response to the environment that we call Mega-Targets of Selection. The identification of compatible breeding programs for germplasm exchange could be of significance for improving genetic gains in breeding programs. However, the application of this methodology in our study was limited by the number of environments we used to produce the classification.
Comparisons of $Q_{ST}$ with $F_{ST}$ in cultivated species under artificial selection provide understanding of the structure of the diversity within and among populations which has been identified as one of the key components of genetic resources conservation and utilization. Furthermore, using the structure at both molecular markers and morphological traits provides better results because neutral diversity and the specific adaptations could be maintained. Moreover, by simultaneously estimating population structure at morphological traits and neutral molecular markers, compatible breeding programs for germplasm exchange, and breeding programs with unique characteristics worth preserving can be identified. However, caution is advised in the use of this methodology for fitness-related traits, traits with considerable GxE, and in germplasm with substantial artificially imposed structure, such as two-row versus six-row barley.

In wild barley, five traits were found to be neutral, two traits experienced both disruptive and stabilizing selection, and six traits were under divergent selection. The following patterns were detected for pair-wise relationship in the traits that were classified as under divergent selection by overall statistics. First, selection and drift were acting in the same direction on awn length. Second, three populations drove $Q_{ST} > F_{ST}$. Third, there are two groups of populations that tended to have $Q_{ST}^{ij} < F_{ST}^{ij}$ for comparisons within groups, and $Q_{ST}^{ij} > F_{ST}^{ij}$ for the comparisons between groups. In cultivated barley, seven traits were found to be neutral, three traits experienced both under disruptive and stabilizing selection, and nine traits were under overall disruptive selection. Additionally, some breeding programs evidenced more diversity in morphological traits than others.

**Final thought**

In the process of getting a Ph.D, you start out as a young and *naïve* person, full of desire to solve *the* main problem in whatever your field is, plant genetic resources conservation was my crusade. You are then faced with the reductionist approaches to science, and struggle to find something really meaningful out of your necessarily very limited experiment because you have to write papers, and ‘all in all [they are] just another
brick in the wall’. However, it is not until you step back to think of the whole process that you realize how profoundly revolutionizing your experience was. You were able to not only think of a problem (a real and meaningful one), design an experiment to test it, conduct the experiment (yes, extensive field testing), extract meaningful data patterns, summarize the new knowledge, and discuss the limitations of your approach such that future work would address those limitations. And your papers are bricks, but really strong bricks. However, the most illuminating experience is the process and not the end results. In the process I had the chance to learn from seemingly opposed mind views. I had the opportunity to shake hands with the father of the Green Revolution, Norman Borlaug, and the Mother of the Seed Saving movement, Vandana Shiva. I also had the great opportunity to work with one of the largest transnational seed companies, the Monsanto Company, and from one of the most radical agricultural NGOs, The Land Institute. I also had the chance to utilize the sometimes seemingly opposed genotype and phenotype methods to conservation through approaches in Plant Breeding and Evolutionary Biology. There is so much more to science than *just* the collection of papers, and that is what makes it worthwhile.