ISSN 1479-2621
Plant Genetic Resources: Characterization and Utilization (2011); 1–3
doi:10.1017/S1479262111000359

A novel genetic framework for studying response to artificial selection

Randall J. Wisser[1]*, Peter J. Balint-Kurti[2,3] and James B. Holland[3,4]

[1]*Department of Plant and Soil Sciences, University of Delaware, Newark, DE, USA,*
[2]*Department of Plant Pathology, North Carolina State University, Raleigh, NC, USA,*
[3]*United States Department of Agriculture-Agricultural Research Service, Plant Science Research Unit, Raleigh, NC, USA and* [4]*Department of Crop Science, North Carolina State University, Raleigh, NC, USA*

Abstract

Response to selection is fundamental to plant breeding. To gain insight into the genetic basis of response to selection, we propose a new experimental genetic framework allowing for the identification of trait-specific genomic loci underlying population improvement and the characterization of allelic frequency responses at those loci. This is achieved by employing a sampling scheme for recurrently selected populations that allows for the simultaneous application of genetic association mapping and analysis of allelic frequency change across generations of selection. The combined method unites advantages of the two approaches, permitting the estimation of trait-specific allelic effects by association mapping and the detection of rare favourable alleles by their significant enrichment over generations of selection. Our aim is to develop a framework applicable for many crop species in order to gain a broader and deeper understanding of the genetic architecture of response to artificial selection.

Keywords: adaptation; association mapping; plant breeding, quantitative trait; selection

Introduction

Insights into the genetic basis of response to artificial selection lag, despite substantial advances in our understanding of sequence diversity. Relative to the abundant literature on the genetic mapping of loci conditioning trait variation, few studies have investigated the genomic response or genetic architecture of generational improvement in crop species in which artificial selection is paramount. We are maintaining a collection of citations for studies that have investigated the genetic architecture of response to artificial selection, accessible at http://www2.udel.edu/wisserlab/RAS-citations or upon request. Herein, a novel genetic framework is proposed for examining the genetic architecture of response to artificial selection.

* Corresponding author. E-mail: rjw@udel.edu

Genetic association mapping

Genetic association mapping is now in widespread use in plants (for review, see Zhu *et al.*, 2008) and has led to the identification of quantitative trait loci and genes underlying standing variation (e.g. Brown *et al.*, 2008; Harjes *et al.*, 2008). It has also been applied to study the genetic architecture associated with long-term artificial selection for kernel oil composition in maize (Laurie *et al.*, 2004).

Association mapping works by identifying significant correlations between allelic and trait variation; however, methods are required to control for non-independence or relatedness among samples comprising many association mapping panels (e.g. Yu *et al.*, 2006). The primary advantages of association mapping are that pre-existing lines from within or across breeding programmes may be used without the need to develop new genetic stocks, multiple alleles can be examined at once and, in some cases, very high-resolution mapping can be achieved because of rapid decay in linkage disequilibrium (such

as in maize [Remington *et al.*, 2001]). The trade-off for high-resolution mapping panels is that, in the absence of good candidate genes, extremely large numbers of markers across the genome or whole genome sequences are required to identify associations. Other disadvantages to association mapping are that correction for relatedness, required to reduce false positive rates, can lead to a significant loss in power, that rare but potentially important alleles cannot be detected and that, for association panels including very diverse (some unadapted) germplasm, phenological variation can confound variation in the trait(s) of interest.

Analysis of allelic frequency change

The most common method used to study the genetic architecture of response to artificial selection has been through the analysis of shifts in allelic frequencies. Frequencies are estimated from samples taken from different generations of selection, and for each locus assayed, a statistic is used to test for significant departures from genetic drift expectations or directional change. Significant departures are taken as evidence of selection; it is commonly assumed that the influence of migration and mutation is minimal.

Initial molecular genetic studies seeking insight into artificial selection response were conducted in maize, using handfuls of allozyme loci (Brown, 1971; Brown and Allard, 1971; Stuber and Moll, 1972). These studies set the stage for the use of more efficient marker technologies and improved statistical methodologies that have allowed for genome-wide scale analyses and the identification of loci presumably underlying response to selection (e.g. Labate *et al.*, 1999; De Koeyer *et al.*, 2001; Coque and Gallais, 2006; Wisser *et al.*, 2008).

There are advantages and limitations to studying recurrently selected populations. A key advantage is that the direct products of a breeding population can be studied in which selection has sorted out the most favourable and unfavourable alleles. Relatively rare but important alleles are enriched by selection and detectable by the analysis of allelic frequency shifts. The primary limitation of this approach is that allelic effects on specific traits cannot often be estimated since selection is not typically conducted or achieved for a single trait. Development of the base population used to initiate selection and the intensity of artificial selection can impact the structure of linkage disequilibrium, which, in turn, can affect the mapping resolution for loci associated with selection. The statistical power of detecting loci through the analysis of allelic frequency shifts also needs to be addressed.

Combining genetic association mapping with analysis of allelic frequency change

In an effort to understand response to artificial selection and the adaptation of exotic germplasm populations to new environments, we have designed and are implementing a framework (Fig. 1) that is expected to counterbalance the opposing limitations and exploit the complementary advantages of association mapping and analysis of allelic frequency change. Recurrently selected populations represent germplasm that has been selected and adapted to an environment of interest. Such populations thus offer the potential of identifying relevant alleles on which to exert selection in a breeding programme. In recurrently selected populations, initially, rare but favourable alleles increase in frequency from generation to generation while common but unfavourable alleles reduce in frequency. Thus, across generations, allelic frequencies at loci underlying selection response would be relatively balanced compared with allelic frequencies in any single generation. Although perfect balance is unlikely due to the transient nature of populations undergoing improvement, this multigenerational sampling strategy will generally overcome a major limitation of association mapping, which is the detection of low-frequency variation. In turn, the application of association mapping to genetically characterize artificial selection response overcomes the major limitation of analysis of allelic frequency change, i.e. the inability to resolve trait-specific locus associations. The ultimate outcome of combining these two approaches

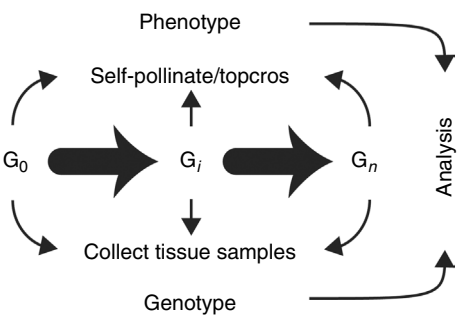

Fig. 1. A genetic framework for studying the genetic architecture of response to artificial selection. The figure depicts a sampling scheme in which random samples of individuals are taken from multiple generations of a recurrently selected population (G_0 representing the base population, G_i representing any intermediate generation and G_n representing the last generation). Tissue is collected from the sample of individuals that are also self-pollinated or topcrossed. Genotyping is conducted on the sample of individuals and phenotyping is conducted on the sample of families or topcrosses. In the analysis phase, association mapping is conducted using genotype data and family or topcross mean performance data. With the same genotype data, allelic frequency shifts can be examined at the quantitative trait loci identified by association mapping.

is a powerful approach to gain information about the map locations of genomic loci associated with trait-specific 'response' variation and insights into the population genetic basis for phenotypic change.

The application of association genetic mapping to any type of population requires consideration of genetic disequilibrium. For most recurrently selected populations, random mating is conducted among selection units. Therefore, gametic phase disequilibrium present in the initial generation or arising due to selection is expected to decay in each generation. This allows unlinked loci to be independently associated with specific traits, despite the fact that selection was exerted; however, family-level structure may need to be accounted for, if the sample of selection units is small or the chance of sampling siblings is high. The level of resolution of association mapping and allelic frequency tests depends on the extent of disequilibrium at linked loci. Recurrently selected populations are expected to have lower levels of linkage disequilibrium than biparental mapping populations and may be as low as diverse germplasm collections, if sufficient generations of random mating were involved in population development (Breseghello and Sorrells, 2006).

In the proposed framework, genotyping is conducted on random samples of individuals taken from different generations of selection. The same individuals are self-pollinated or crossed with a common tester to produce entries that are evaluated in replicated trials. Entry means are used to represent the phenotypes of the parental individuals from which genotype data were produced, and association mapping is conducted by fitting the genotype data to the phenotype data. With the same genotype data, estimates of allelic frequencies are obtained and examined across generations of selection, particularly at genomic loci identified via association mapping. The two approaches provide complementary tests to examine loci that are responsive to selection.

Acknowledgements

This research was supported by the USDA National Institute of Food and Agriculture project 2007-35301-18133/19859.

References

Breseghello F and Sorrells ME (2006) Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Science* 46: 1323–1330.

Brown AHD (1971) Isozyme variation under selection in *Zea mays*. *Nature* 232: 570–571.

Brown AHD and Allard RW (1971) Effect of reciprocal recurrent selection for yield on isozyme polymorphisms in maize (*Zea mays* L.). *Crop Science* 11: 888–893.

Brown PJ, Rooney WL, Franks C and Kresovich S (2008) Efficient mapping of plant height quantitative trait loci in a sorghum association population with introgressed dwarfing genes. *Genetics* 180: 629–637.

Coque M and Gallais A (2006) Genomic regions involved in response to grain yield selection at high and low nitrogen fertilization in maize. *Theoretical and Applied Genetics* 112: 1205–1220.

De Koeyer DL, Phillips RL and Stuthman DD (2001) Allelic shifts and quantitative trait loci in a recurrent selection population of oat. *Crop Science* 41: 1228–1234.

Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, Yan JB and Buckler ES (2008) Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* 319: 330–333.

Labate JA, Lamkey KR, Lee M and Woodman WL (1999) Temporal changes in allele frequencies in two reciprocally selected maize populations. *Theoretical and Applied Genetics* 99: 1166–1178.

Laurie CC, Chasalow SD, LeDeaux JR, McCarroll R, Bush D, Hauge B, Lai CQ, Clark D, Rocheford TR and Dudley JW (2004) The genetic architecture of response to long-term artificial selection for oil concentration in the maize kernel. *Genetics* 168: 2141–2155.

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doeblay J, Kresovich S, Goodman MM and Buckler ES (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences USA* 98: 11479–11484.

Stuber CW and Moll RH (1972) Frequency changes in isozyme alleles in a selection experiment for grain yield in maize (*Zea mays* L.). *Crop Science* 12: 337–340.

Wisser RJ, Murray SC, Kolkman JM, Ceballos H and Nelson RJ (2008) Selection mapping of loci for quantitative disease resistance in a diverse maize population. *Genetics* 180: 583–599.

Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S and Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38: 203–208.

Zhu C, Gore M, Buckler ES and Yu J (2008) Status and prospects of association mapping in plants. *The Plant Genome* 1: 5–20.