Canadian Science Publishing | Canadian Journal of Plant Science

Using Soybean Pedigrees to Identify Genomic Selection Signatures Associated with Long-Term Breeding for Cultivar Improvement

Journal:	*Canadian Journal of Plant Science*
Manuscript ID	CJPS-2017-0339.R1
Manuscript Type:	Article
Date Submitted by the Author:	29-Jan-2018
Complete List of Authors:	Grainger, Christopher; University of Guelph, Dep. of Plant Agriculture Letarte, Jocelyne; University of Guelph, Plant Agriculture Rajcan, Istvan ; University of Guelph, Dep. of Plant Agriculture
Keywords:	Soybean, Selective sweep, QTL, Pedigree, Plant breeding
Is the invited manuscript for consideration in a Special Issue?:	

Using Soybean Pedigrees to Identify Genomic Selection Signatures Associated with Long-Term Breeding for Cultivar Improvement

Christopher M. Grainger, Jocelyne Letarte and Istvan Rajcan*

Department of Plant Agriculture, Crop Science Building, University of Guelph

50 Stone Road East, Guelph, Ontario, N1G 2W1, Canada

*Corresponding author: irajcan@uoguelph.ca

Abstract

Genetic hitchhiking uncovers selection signatures related to traits of agronomic importance in crops and has been primarily used at the level of domestication by comparing groups of wild germplasm to landraces or elite breeding lines. In this study, two groups of cultivars defined by an elite Canadian soybean cultivar, ‘OAC Bayfield’, were compared to identify selection signatures related to long-term breeding within a specific region. Cultivars were assigned to either a pre- or post-OAC Bayfield group. Of the 162 simple sequence repeat (SSR) markers used to genotype members of the pedigree, 14 were fixed and 19 exhibited a selective signature. An *in silico* analysis compared the results in this study to quantitative trait loci (QTL) reported in SoyBase and showed that 18 out of the 19 markers with a selective signature were associated with at least one QTL. From the 80 QTL associated with the 18 markers, half were related to plant architecture, yield or maturity. In addition, the number and type of QTL associated with the fixed versus selected loci differed particularly, for yield. Genomic regions exhibiting a selection signature may contain important loci that either need to be conserved for agronomic performance or be targeted for introgressive breeding and germplasm enrichment.

Keywords

Soybean, QTL, SSR, Selective Sweep, Genetic Bottleneck,, Pedigree, Plant Breeding

Introduction

Commercially successful soybean [*Glycine max* (L). Merr.] cultivars benefit a breeding program in two ways: first, by the performance of the cultivar *per se*, and second, as a crossing parent in the development of future cultivars. In fact, these highly utilized lines become known as founder genotypes that greatly influence the genetic architecture of a breeding program. Founder effects have been well documented in North American soybean breeding programs. For the northern growing regions, only 10 ancestral lines account for 80 % of the genetic base, with five of these ancestors accounting for more than 65 % (Gizlice et al. 1994). Moreover, there has been little research done on the development of a founder genotype and its subsequent impact within and among breeding programs. The question as to which genomic regions have been selectively maintained by breeders from founder genotypes over generations of breeding needs to be addressed.

In evolutionary terms, processes such as domestication and applied plant breeding can lead to events known as “selective sweeps” (Olsen et al. 2006; Shi and Lai 2015). In these events, genetic diversity is eroded due to natural or artificial selection of beneficial molecular variants for adaptation and agronomic performance (Vigouroux et al. 2002). As a consequence, QTL and association mapping methods may miss a promising class of genes due to a lack of genetic diversity attributable to previous selection whereas selective sweep studies have the ability to identify these genomic regions (Yamasaki et al. 2005; Shi and Lai 2015).

Human modification of crop genomes has largely been explored at the domestication level (Doebley et al. 2006). Effects of the domestication process on genetic diversity have been studied extensively for a number of crop species including maize (Wright et al. 2005), wheat (Haudry et al. 2007), rice (Zhu et al. 2007), sunflower (Liu and Burke 2006) and soybean (Lam et al. 2010). These studies have shown the effect of the domestication process to be one of genome-wide genetic diversity loss with particular regions exhibiting extreme diversity loss due in part to directional selection by humans. In some cases, specific genes have been associated with loss of genetic diversity such as *tb1* (Wright et al. 2005) and *y1* (Palaisa et al. 2004) in maize or *Dt1* in soybean (Tian et al. 2010). This is due to continuous selection pressure applied by plant breeders, where a favorable allele or haplotype of a gene can become fixed in a population or the germplasm. Usually this fixation occurs (in the context of cultivar development) when there is improved agronomics associated with the favourable haplotype (Vigouroux et al. 2002). In addition, the level at which fixation occurs can be associated with crop domestication (e.g. *tb1*) or cultivar improvement (e.g. *y1*) (Doebley et al. 2006).

Genetic bottlenecks are events that reduce genetic variation of a population due to a drastic reduction in population size (Barker et al. 2009).. Domestication of crops provides good examples of dramatic morphological and genetic modifications occurring on a short evolutionary time scale (Haudry et al. 2007). Small initial population sizes and intense human selection for agronomic traits may explain the decrease of genetic diversity in most crop plants (Tanksley and McCouch 1997). Thus, domestication can be seen as a population or genetic bottleneck in most crop species (Buckler et al. 2001). Hyten et al. (2006) investigated the impact of genetic bottlenecks on soybean genome diversity through various germplasm transition points (wild, landraces, North American ancestors and elite cultivars). They found that the domestication bottleneck from wild germplasm to landraces (i.e. *Glycine soja* to *Glycine max*) had the greatest impact on diversity loss, when the nucleotide diversity was reduced by 50 %. Interestingly, their study also found that modern cultivars have retained 72 % of the sequence diversity present in the Asian landraces and that modern soybean breeding has minimally affected the allelic structure of the genome compared with other historical bottlenecks (Hyten et al. 2006). This would suggest that soybean breeders have selected in specific genomic regions when they developed new cultivars for a particular growing region.

A population genetics approach used to detect molecular selection signatures is to apply a microsatellite genome scan between different groups or populations and identify outlier loci from the distribution of values (Soto-Cerda and Cloutier 2013). The genomic regions that “hitchhike” along with the marker (Harr et al. 2002) form candidate regions that may have been subjected to molecular selection (Casa et al. 2005). Genomic scans for selection have been applied to natural populations (Kauer et al. 2003; Nielsen 2005; Ihle et al. 2006) as well as artificial populations (Li et al. 2010; Ge et al. 2012) to elucidate genome changes associated with adaptation or improvement. In soybean, Jun et al. (2011) identified signatures of selection near QTL of agronomic importance. They compared the genetic diversity in three classes of SSR (no QTL, QTL for seed protein content and QTL for *Sclerotinia* resistance) between a group of *Glycine soja* and *Glycine max* lines. There were significant changes in the allele frequency for the SSR located in known QTL regions, whereas there was no change in diversity at the SSR in the QTL absent region.

Pedigree records of a breeding program detail both the historical and present use of various varieties/breeding lines and can identify key individuals that have had a major impact on the overall program itself. The overall genome composition changes in the elite Canadian soybean cultivar, OAC Bayfield was recently characterized by tracking chromosome inheritance through the pedigree (Grainger and Rajcan 2014). OAC Bayfield (Tanner et al. 1998) was chosen since it represents a landmark cultivar for soybean growers in Ontario. It was developed by the University of Guelph and commercially released in 1994. At its peak in 1998, it was grown on over 400,000 acres in Ontario. From 1994 to 2004, the estimated value of OAC Bayfield to the Ontario economy was in excess of $750 million [Agricultural Research Institute of Ontario (ARIO) 1998]. Moreover, the significance of this elite cultivar also extends to its important use as a founder genotype in other commercially successful cultivars (Grainger and Rajcan 2014). As a follow up to the genome characterization of OAC Bayfield, the present study aims at identifying signatures of selection related to long-term breeding within a specific growing region by comparing two groups of cultivars defined by the pedigree of OAC Bayfield.

Methods and Materials

Plant material and SSR genotyping

Plant material and SSR genotyping methods are the same as those described in detail in Grainger and Rajcan (2014). Briefly, the pedigree of OAC Bayfield, which consisted of genotypes leading to and developed from OAC Bayfield, were genotyped with 162 SSR markers at a density of approximately 1 marker per 10 centimorgans (cM). The SSR markers were selected based on the composite linkage map (Song et al. 2004) in SoyBase (http://www.SoyBase.org). The genotypes derived from OAC Bayfield represented selections from six independent public and private soybean breeding programs based in Ontario and Quebec, Canada For ancestral genotypes (i.e., pre-OAC Bayfield), accessions were obtained from either the Plant Gene Resources of Canada (PGRC) in Saskatoon, SK, Canada or the United States Department of Agriculture (USDA) soybean germplasm collection in Urbana-Champaign, IL, USA. Seeds for the remaining genotypes were either collected directly from the University of Guelph soybean breeding program or received from various collaborating soybean breeding programs, which included: Pioneer Hi-Bred, Ridgetown College, La Coop fédérée, Semences Prograin and Agriculture and Agrifood Canada (AAFC).

Pedigree-based grouping of cultivars

By using pedigrees, cultivars can be identified as either contributors to the development of a founder genotype or as cultivars derived from this founder genotype. For this study, members of the pedigree were placed into two groups. First, all the cultivars that contributed to the development of OAC Bayfield were classified as the "*Ancestral Group*" (i.e. pre-OAC Bayfield), whereas the cultivars developed from OAC Bayfield were classified as the "*Current Group*" (i.e. post-OAC Bayfield). The reason for this grouping was to gain insights into the development of a founder genotype and identify selection signatures retained through it across a new group of germplasm developed from it (Figure 1).

SSR diversity between groups

SSR allele frequency was calculated using PowerMarker v3.0 (Liu and Muse 2005) for the full pedigree (i.e. no grouping of genotypes) to determine the number and distribution of loci that were fixed throughout the pedigree. Genotypes were then assigned to the *Ancestral Group* or *Current Group* and the program was used to calculate SSR gene diversity (marker diversity) values for each group (ancestors and current), which is a measure analogous to expected heterozygosity (Casa et al. 2005). A Mann-Whitney U test was used to determine if the gene diversity estimates were significantly different between the two groups.

Selection signature statistics

For identifying selective sweeps using SSR markers, a frequently used statistic is the ln RH test (Schlötterer 2002*a*; Schlötterer and Dieringer 2005) as it is a measure of the change in variability in expected heterozygosity or gene diversity between two groups or populations at each SSR locus and is derived using the following equation:

$$\ln[(RH)] = \ln\left[\frac{\left(\left(\frac{1}{1-He_{\mathrm{Pop1}}}\right)^2-1\right)}{\left(\left(\frac{1}{1-He_{\mathrm{Pop2}}}\right)^2-1\right)}\right]$$

With He_{Pop1} representing gene diversity (expected heterozygosity) in the *Current Group* (all individuals after OAC Bayfield), and He_{Pop2} representing gene diversity in the *Ancestral Group* (all individuals before OAC Bayfield). The ln RH statistic has been shown to be robust as it is largely independent of microsatellite mutation rates, mutational step-size and demographic events (Schlötterer 2002*a*). Also, the power of ln RH is highest immediately following the selective event (Casa et al. 2005), which is the case for this study. To test whether the observed ln RH values were normally distributed a Kolmogorov-Smirnov test was performed.

The other test statistic used was F_{st}, which is used to assess the level of genetic differentiation between groups or populations (Wright 1951). In selective sweep studies it is used as an indicator of selection as in many cases selection increases the degree of allelic differentiation between groups or populations (Nielsen 2005). To derive the F_{st} estimates, an analysis of molecular variation (AMOVA) was performed using GenAIEx ver. 6.5 (Peakall and Smouse 2006, 2012) to partition the genetic variation either within each group, ancestral or current or between the two groups as a measure of F_{st} at each microsatellite locus (Excoffier et al. 1992).

For each test statistic, significant loci were identified as outliers in the data distribution (Schlötterer 2002*b*). For the ln RH statistic, ln RH values were standardized across loci (Kauer 2003). Therefore, 95 % of loci were expected to have values that were between 1.96 and –1.96, with those loci falling outside this interval considered significant ($P < 0.05$) and interpreted as loci that may have been targets of selection (Casa et al. 2005). For the F_{st} values, the empirical approach of Casa et al. (2005) and Kayser et al. (2003) was used, where significant loci are those with the highest F_{st} values relative to the genome-wide average. For each test statistic monomorphic SSR loci (full pedigree) were excluded from the analysis, resulting in a final dataset of 148 SSR markers used for the selection tests.

Results

Group differentiation and genetic diversity loss

Results from the Mann Whitney test showed that there was a significant difference (P<0.001) between groups for the gene diversity (marker diversity) estimates. The mean gene (marker) diversity for the *Ancestral Group* was 0.56 compared to 0.45 for the *Current Group*. Therefore, it was deemed appropriate to consider the ancestors (pre-OAC Bayfield) as a separate group from the current (post-OAC Bayfield) group. Of the 148 markers, 121 showed reduced genetic diversity in the *Current Group*, with the remainder having equal or greater variability than the *Ancestral Group*. When considering the full pedigree with no group assignment, there were 14 microsatellite loci with a major allele frequency >0.95, which corresponds to approximately 9 % of the sampled genome that has remained fixed since the foundation ancestors of the pedigree. The fixed loci were not evenly distributed as half of them were located on chromosomes 3 and 8 (Table 1). SSR markers that had greater variability in the *Current Group* than the *Ancestral Group* were distributed among the chromosomes, except for chromosome 4, where there was a group of consecutive markers (Satt565, Soygpatr, Satt396 and Satt578) that all had higher levels of diversity in the *Current Group* than the *Ancestral Group* (Supplementary Table 1).

Genomic regions exhibiting signatures of selection

The Kolmogorov-Smirnov test confirmed that the observed ln RH values were normally distributed. Thus, the values could be standardized and the ln RH statistic was applied to scan for local selective sweeps (Kauer 2003). The mean F_{st} value between the two groups was 0.05 (Supplementary Table 2). As these groups are expected to have only subtle genetic subdivision due to the short time scale in evolutionary terms and degree of relatedness, the loci considered most likely to have been impacted by selection were those with F_{st} values that were at least three-fold higher (>0.15) than the genome-wide average. This threshold was derived in part from the study done by Casa et al. (2005). In their study the F_{st} values considered significant were values that were at least four-fold higher than the genome-wide average, which was 0.13 in the study. As the authors compared wild accessions to landraces, a greater degree of genetic differentiation was expected due to domestication. However, this degree of allelic differentiation would not be expected when comparing adapted ancestors to modern cultivars derived from them. Therefore, a three-fold higher F_{st} value was considered appropriate and not too restrictive for identifying potential regions of selection.

Out of the original 162 markers, 148 could be tested using both statistics (monomorphic loci were excluded). Ten loci were significant for the ln RH test (Figure 2), while 13 loci were identified with the F_{st} test (Figure 3), with four loci being detected by both tests (Satt357, Satt249, Sat104 and Satt578).

In silico QTL identification using selection signatures

Several studies have shown that selective sweeps appear to be frequent in genomic regions that contain quantitative trait loci or genes under directional selection (Vigouroux et al. 2002; Edelist et al. 2006; Zhang et al. 2007; Jun et al. 2011). Soybean has been extensively investigated at the molecular level in the form of QTL studies over the past 20 years. This has resulted in an abundant amount of marker-trait associations that has been deposited in databases, which provided the opportunity to investigate what QTL might be residing in these regions and for what traits. The results of this study were compared with the genetic maps (Song et al. 2004) in SoyBase (www.soybase.org 2017).

For the SSR markers that had allele frequencies greater than 0.95 (i.e. fixed throughout the pedigree), 9 out of the 14 had QTL that were associated with them, according to SoyBase (www.soybase.org 2017). There was a total of 39 QTL identified as being associated with the 9 markers (Table 1). The 39 QTL represented 24 different traits that belonged in the following categories: plant architecture, yield, maturity, oil, seed composition, seed quality, disease and other. Table 2 summarizes the grouping of individual QTL and the number of QTL within each category.

For the SSR markers identified by the selection tests, 18 out of 19 markers had at least one QTL associated with it. Only Satt636 had no reported QTL associated with it. In total there were 80 QTL that were associated with the 18 markers (Table 3). While there were 46 distinct trait QTL represented by the 80 total QTL, all of them could be classified into the following nine trait categories: plant architecture, yield, maturity, oil, protein, seed composition, seed quality, disease and other. Table 4 summarizes the grouping of individual QTL and the number of QTL within each category for the individual markers exhibiting a selective signature.

Discussion

Pedigree-based bottleneck versus selection effects

Genetic bottlenecks, drift and selection are all mechanisms that erode genetic diversity in a given species (Barker et al. 2009). The effects of genetic bottlenecks and directional selection differ in one key area; genetic bottlenecks affect the entire genome, whereas selection acts on specific regions (Li et al. 2010). The ln RH and F_{st} tests were employed to separate the loci that exhibited a change in diversity beyond that of the general bottleneck effect. In order to do this, the overall severity of the genetic bottleneck due to the single founder genotype, OAC Bayfield was first determined. The concern was that the bottleneck effect could skew the distribution of ln RH values toward a non-normal distribution of negative values, and a severe bottleneck would show a drastic reduction in the overall genome-wide level of genetic variability. This was not the case, as the values did form a normal distribution.

One factor that contributed to the normal distribution observed was the use of microsatellite markers themselves. Unlike single nucleotide polymorphisms (SNPs) that have an inherent low mutation rate and are more susceptible to confounding effects of demography, microsatellites have a much higher mutation rate, which provides sufficient variation in closely related populations. This increases the chance of detecting recent selective sweeps (Teschke et al. 2008). The use of SSR markers may be superior to SNPs for detecting local or subtle selective sweeps due to breeding for local adaptation in a specific environment or geographical location (Schlötterer 2002*b*). Notably, Schlötterer (2002*a*) has suggested that an approach to improve the power of variability ratio-based selection statistics is to compare two closely related populations, which was the case for this study.

Another key aspect in the ability to "re-coup" genetic diversity in the *Current Group* was due to the number of alternate parental genomes (n=16, for which we had information for) that were used in crossing with OAC Bayfield, and its progeny cultivars, 'OAC Champion' and 'OAC Kent'. Many of the alternate parents were from independent breeding programs (five in total) and represented additional genetic diversity (alleles from different gene pools) that may not have been captured by the University of Guelph's germplasm. Microsatellite diversity across multiple breeding programs is undoubtedly higher than in a single breeding program and is a contributing factor for the post-OAC Bayfield cultivars to have regained their differences quickly in non-selected areas. It is clear from the gene diversity estimates (Supplementary Table 1) that the overall bottleneck effect was a 17 % reduction of diversity from 0.56 in the *Ancestral Group* to 0.45 in the *Current Group*. This result is in agreement with the study by Hyten et al. (2006), which reported that "breeding bottlenecks" have not had the same impact on genetic diversity loss as other bottlenecks such as domestication.

Only *potential* regions that *may* have experienced selection were identified. Many selective sweep-based studies recommend caution when interpreting the results of these types of analysis due to confounding effects of demography or determining appropriate neutrality expectations (Nielsen 2001; Vigouroux et al. 2002; Casa et al. 2005; Teshima et al. 2006). Furthermore, the effects of genetic drift must be acknowledged as a factor in those regions that have remained fixed after OAC Bayfield, possibly resulting in some of the loci being false positives. However, it has been recommended that "agricultural scientists should moderate the usual concern about false positives so that all reasonable, even if marginal, candidates are advanced to the next level of testing" (Vigouroux et al. 2002). The use of multiple tests is one way to reduce false positives/negatives since loci detected by multiple tests are more likely to be chosen for further investigation. In this study four loci met the criteria (Satt578, Satt357, Satt249 and Sat104). Another method for validating a selective signature is to conduct genomic scans between several groups or populations and perform multiple pair-wise comparisons. Loci that are consistently identified across comparisons are more likely to be true selection signatures (Schlötterer 2002*a*) and this could be used within, and between breeding programs as many populations are routinely produced. Additionally, by performing similar studies with other "founder" or elite type varieties, consistent selection signatures could be identified since it has been shown that same selection signatures can be found in common between founder genotypes. In a study by Ge et al. 2012, they compared two sets of bred wheat varieties and their respective founder parents. It was shown that the alleles present at the loci identified as having undergone a selective sweep were often in common between the two founding parents. Thus, these genomic regions may be very important for cultivar development within a breeding program.

Trait QTL identified by in silico analysis

A limitation of selective sweep studies in certain species is the small amount of information about genotype-phenotype associations (Ihle et al. 2006). A luxury of economically important crop species is the wealth of information available in regards to genotype-phenotype associations through numerous QTL studies. An important aspect of this study was to empirically evaluate what QTL were present in the regions of the identified SSR loci. In other words, would the associated QTL correspond to traits that we would expect to be selected for, given the selection practice of the breeding program?

With the University of Guelph soybean breeding program, the three traits that would be expected to have the greatest amount of selection pressure applied would be agronomic characteristics, yield and maturity. Out of the 73 QTL associated with various markers, approximately 50 % (36 out of 73) belonged to these three categories. .

A search of QTL abundance based on trait type in SoyBase added support that the number of QTL in the defined categories could not be solely explained by QTL density for a given trait. There are 188 seed yield ("Sd yld") QTL in SoyBase, and taken together, Glycitein/Daidzein and Isoflavone have 194 QTL reported as of October 3, 2017. Six "Sd yld" QTL and seven QTL for Glycitein/Daidzein and Isoflavone were identified. Interestingly, no QTL associated with soybean cyst nematode (SCN) were found even though it is an abundant category with 207 different QTL reported. A possible explanation for this is that SCN has historically been a disease of southern growing environments and was not reported in Ontario until 1988 (OMAFRA 2017). Thus, it has not been a major selection parameter for northern growing environments until recently and it is therefore reflected in the pedigree of OAC Bayfield, with little representation of SCN resistant cultivars comprising the pedigree. Overall, the type and abundance of QTL that were identified are in the domain of "plant breeding" selection signatures and not simply QTL that are distributed randomly as an effect of genetic bottlenecks and drift. The 46 traits QTL associated with the 18 markers identified by the selection tests represent only 18 % of the possible types of QTL that are reported in SoyBase (251 different classes as of October 17, 2017).

There was additional support for the identified genomic regions being attributable to selection when comparing the types of QTL identified by the ln RH and F_{st} tests to those that were fixed from the founder ancestors of the pedigree (Tables 2 and 4). The number of SSR loci that did not have any QTL associated with them was greater for the fixed loci (5/14) than the selected loci (1/19). If these fixed regions represent historical fixation due to previous genetic bottlenecks or genetic drift, then they would be more likely to be found in random areas of the genome, which may or may not contain QTL. Conversely, genomic regions under directional selection would be more likely to contain QTL for trait improvement or regional adaptation (Jun et al. 2011). Furthermore, the type of QTL differed in particular ways between the fixed and selected loci. There were no protein QTL associated with the fixed loci and the only type of yield QTL was for seed weight ("Sd wt"). If one considers seed weight in the context of a domestication trait in relation to seed size differences between *G. soja* and *G. max* then it would be reasonable to expect fixation to have already occurred in some of these regions.

With genomic regions that showed a selective signature, there were specifically "Sd yld" QTL associated with certain markers. Similarly, if one considers yield as an improvement trait due to plant breeding selection pressure, then it should be reflected in the type of QTL found at the loci exhibiting a selective signature due to long-term plant breeding activities. The nature of the selection pressure (balancing vs. diversifying) however, is beyond the scope of this study.

Furthermore there were additional QTL in the mapping intervals surrounding the markers exhibiting a selective signature but QTL identifications were restricted to only those that were associated with markers based on formal QTL mapping studies (i.e. those with published studies as reported in SoyBase) to avoid excess speculation. (Palaisa et al. 2004) has shown that selective sweeps may impact regions well beyond the target gene. Therefore, a higher resolution scan across a wider sample of germplasm will be required to ultimately determine what specific genomic regions are exhibiting selective signatures and what functional genes may be in those regions and what phenotypic characteristics they may affect. Furthermore, it is acknowledged that the QTL identified in this study correspond to a list of typical QTL found at loci with exhibitive selective signature, which was the intended objective of this work. This does not imply that all of the QTL identified were being selected for, because the purpose of the *in silico* analysis was to determine if there was consistency in the types of QTL that were associated with the loci identified by the selection tests. Completely separating the effects of drift versus selection is challenging and would require a more intensive genomic scan of the material than what was possible in this study. Future studies involving next generation sequencing technology will aid in detecting selection signatures with greater confidence.

A number of the SSR markers were associated with multiple QTL for different traits suggesting that pleiotropic genomic regions could be identified and could contain genes important to multiple traits. An example of this is the number of isoflavone QTL identified in both the fixed and selected loci. The biological role of isoflavones in soybean is of a defence mechanism against insects and diseases, as well as environmental stresses (Murphy et al. 2009). Therefore, the identification of these QTL in the context of a selective sweep analysis could be due to beneficial adaptation of soybean varieties to the local environment, rather than breeding for increased concentration of isoflavone compounds *per se*. Further research will be required to address these types of questions.

Conclusion

For many years, bi-parental QTL mapping and association analysis have been the primary methods employed for uncovering genomic regions related to the phenotypic expression of traits in soybean. However, factors such as low recombination, high linkage disequilibrium and artificial selection can limit the types of genes that can be identified (Flint-Garcia et al. 2003). An alternative approach is to investigate the impact of selection itself by identifying selection signatures through the use of genetic hitchhiking methods (Zhang et al. 2007). The majority of selective sweep studies have focused on the comparison of wild ancestors with adapted landraces or elite breeding lines for determining the genomic changes associated with domestication. For soybean breeders, the determination of genomic changes associated with long term selection within a specific growing region provides valuable insights to the areas of the genome that are targeted for crop improvement.

With the capabilities and cost of current genomic technologies it is relatively simple and inexpensive to apply genome scans in soybean. Using selective sweep based tests to identify genomic regions associated with local adaptation or improvement could aid in breeding decisions. Genomic regions exhibiting a selection signature could be viewed as regions that should be conserved in a breeding program to maintain performance (Ge et al. 2012). Conversely, if genomic regions detected are considered to be of agronomic importance, it is conceivable that they may suffer from severe loss of genetic diversity, or mass fixation. This brings into question the long term ability to maintain sufficient genetic gain for a given trait (Hyten et al. 2006). This situation would allow for targeted introgressive breeding strategies for germplasm enrichment (Tanksley and McCouch 1997). Furthermore, by comparing regionally adapted germplasm to major germplasm collections, breeders would have the ability to identify location-or-trait specific selection signatures. This type of knowledge would be critically important to have if soybean breeders were to utilize and integrate information from large-scale genotyping of gene banks such as the current efforts to SNP genotype the approximately 19,000 soybean accessions of the USDA soybean germplasm collection (Song et al. 2013, 2015). Since a small proportion of a breeder's crosses ultimately lead to finished varieties, the choice of crossing parents is perhaps the most important decision in a breeding program (Zhuang 2003). The identification of selective signatures in key genotypes used in breeding programs could lead to new crossing strategies and address various breeding objectives. It would be of great value to identify genomic regions that are selectively maintained by soybean breeders and have retained selection signatures through the development of elite cultivars and the subsequent cultivars derived from them. In this study we have identified four SSR marker loci which are associated with 20 QTL of agronomic importance exhibiting a selective signature in each of the tests used. These loci may represent important regions for crop improvement or adaptation of the crop. In addition, it could also be a method to discover potentially more robust QTL for a breeding program as the regions identified are across multiple genetic backgrounds of germplasm that have become commercial cultivars. As there is an overabundance of QTL associated with the same traits on nearly every chromosome in soybean, a major challenge is to identify those that are the most beneficial to a particular breeding program. The selective sweep based approach offers a means to identify specific regions and analyze specific candidate QTL regions to determine their functional and historical importance in a breeding program.

Acknowledgements

We sincerely thank the following public and private breeding programs for supplying the various breeding lines for this study: Ridgetown College Campus of the University of Guelph, La Coop fédérée, Semences Prograin, Agriculture and Agri-Food Canada (AAFC) and Dupont Pioneer. We would also like to thank Angela Hollis and Jeff Gross for the genotyping that was done at the University of Guelph's Genomics Facility. Funding for this research was generously provided for by the Grain Farmers of Ontario (Grant # S2011GE02).

References

Agricultural Research Institute of Ontario (ARIO). 1998. Research accomplishments 1993–1998. 1998 ARIO led crops review.

Barker, J.S.F., Frydenberg, J., González J., Davies H.I., Ruiz, A., Sørensen J.G., and Loeschcke, V. 2009. Bottlenecks, population differentiation and apparent selection at microsatellite loci in Australian *Drosophila buzzatii*. Heredity, **102**: 389–401.

Buckler, E., Thornsberry, J.M., and Kresovich, S. 2001. Molecular diversity, structure and domestication of grasses. Genet Res. **77:** 213–218.

Casa, A.M., Mitchell, S.E., Hamblin, M.T., Sun, H., Bowers, J.E., Paterson, A.H., Aquadro, C.F., and Kresovich, S. 2005. Diversity and selection in sorghum: simultaneous analyses using simple sequence repeats. Theor. Appl. Genet. **111**: 23–30.

Doebley, J.F., Gaut, B.S., and Smith, B. 2006. The molecular genetics of crop domestication. Cell **127**:1309–1321.

Edelist, C., Lexer, C., Dillmann, C., Sicard, D., and Rieseberg, L. 2006. Microsatellite signature of ecological selection for salt tolerance in a wild sunflower hybrid species, *Helianthus paradoxus*. Mol. Ecol. **15**: 4623–4634.

Excoffier, L., Smouse, P.E., and Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics, **136**: 343–359.

Flint-Garcia, S.A., Thornsberry, J.M., and Buckler IV, E.S. 2003. Structure of linkage disequilibrium in plants. Annu. Rev. Plant. Biol. **54**: 357–374.

Ge, H., You, G., Wang, L., Hao, C., Dong, Y., Li, Z., and Zhang, X.Y. 2012. Genome selection sweep and association analysis shed light on future breeding by design in wheat. Crop Sci. **52**: 1218–1228.

Gizlice, Z., Carter, T.E. Jr., and Burton, J.W. 1994. Genetic base for North American public soybean cultivars released between 1947-1988. Crop Sci. **34**: 1143–1151.

Grainger, C.M., and Rajcan, I. 2014. Characterization of the genetic changes in a multi-generational pedigree of an elite Canadian soybean cultivar. Theor. Appl. Genet. **127**: 211–229.

Harr, B., Kauer, M., and Schlötterer, C. 2002. Hitchhiking mapping: A population-based fine-mapping strategy for adaptive mutations in *Drosophila melanogaster*. PNAS, **99**: 12949–12954.

Haudry, A., Cenci, A., Ravel, C., Bataillon, T., Brunel, D., Poncet, C., Hochu, I., Poirier, S., Santoni, S., Glémin, S., and David, J. 2007. Grinding up wheat: A massive loss of nucleotide diversity since domestication. Mol. Biol. Evol. **24**: 1506–1517.

Hyten, D.L., Song, Q., Zhu, Y., Choi, I.-Y., Nelson, R.L, Costa, J.M., Specht, J.E., Shoemaker, R.C., and Cregan, P.B. 2006. Impacts of genetic bottlenecks on soybean genome diversity. PNAS, **103**: 16666–16671.

Ihle, S., Ravaoarimanana, I., Thomas, M., and Tautz, D. 2006. An analysis of signatures of selective sweeps in natural populations of the house mouse. Mol. Biol. Evol. **23**: 790–797.

Jun, T.-H., Van, K., Kim, M.Y., Kwak, M., and Lee, S.-H. 2011. Uncovering signatures of selection in the soybean genome using SSR diversity near QTLs of agronomic importance. Genes & Genomics, **33**: 391–397.

Kauer, M.O., Dieringer, D., and Schlötterer, C. 2003. A microsatellite variability screen for positive selection associated with the "out of Africa" habitat expansion of *Drosophila melanogaster*. Genetics **165**: 1137–1148.

Kayser, M., Brauer, S., and Stoneking, M. 2003. A genome scan to detect candidate regions influenced by local natural selection in human populations. Mol. Biol. Evol. **20**: 893–900.

Lam, H.-M., Xu, X., Liu, X., Chen, W., Yang, G., Wong, F.-L., Li, M.W., He, W., Qin, N., Wang, B., Li, J., Jian M., Wang, J., Shao, G., Wang, J., Sun, S., and Zhang, G. 2010. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nature Genet. **42**: 1053–1059.

Li, M.-H., Iso-Touru, T., Laurén, H., and Kantanen, J. 2010. A microsatellite-based analysis for the detection of selection on BTA1 and BTA20 in northern Eurasian cattle (*Bos taurus*) populations. Genet. Sel. Evol. **42**: 32.

Liu, A., and Burke, J.M. 2006. Patterns of nucleotide diversity in wild and cultivated sunflower. Genetics, **173**: 321–330.

Lui, K., and Muse, S. 2005. PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics, **21**:2128–2129.

Murphy, S.E., Lee, E.A., Woodrow, L., Seguin, P., Kumar, J., Rajcan, I., and Ablett, G.R. 2009. Genotype x environment interaction and stability for isoflavone content in soybean. Crop Sci. **49**: 1313–1321.

Nielsen, R. 2001. Statistical tests of selective neutrality in the age of genomics. Heredity, **86**: 641–647.

Nielsen, R. 2005. Molecular signatures of natural selection. Annu. Rev. Genet. **39**: 197–218.

Olsen, K.M., Caicedo, A.L., Polato, N., McClung, A., McCouch, S., and Purugganan, M.D. 2006. Selection under domestication: Evidence for a sweep in the rice *waxy* genomic region. Genetics, **173**: 975–983.

OMAFRA, Ontario Ministry of Agriculture, Food and Rural Affairs. 2017. Diseases of field crops: Soybean diseases. Accessed October 2017 from www.omafra.gov.on.ca.

Palaisa, K., Morgante, M., Tingey, S., and Rafalski, A. 2004. Long-range patterns of diversity and linkage disequilibrium surrounding the maize Y1 gene are indicative of an asymmetric selective sweep. PNAS, **101**: 9885–9890.

Peakall, R., and Smouse, P.E. 2006. GenAIEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes **6**: 288–295.

Peakall, R., and Smouse, P.E. 2012. GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics. **28**: 2537–2539.

Schlötterer, C. 2002*a*. A microsatellite-based multilocus screen for the identification of local selective sweeps. Genetics, **160**: 753–763.

Schlötterer, C. 2002*b*. Towards a molecular characterization of adaptation in local populations. Curr. Opin. Genet. Dev. **12**: 683–687.

Schlötterer, C., and Dieringer, D. 2005. A novel test statistic for the identification of local selective sweeps based on microsatellite gene diversity. Chapter 5 (55-64) *Selective Sweep*. Eurekah.com and Klüver Academic/Plenum Publishers. Georgetown, TX.

Shi, J., and Lai, J. 2015. Patterns of genomic changes with crop domestication and breeding. Curr. Opin. Plant Biol. **24:** 47–53.

Song, Q., Hyten, D.L., Jia, G., Quigley, C.V., Fickus, E.W., Nelson, R.L., and Cregan, P.B. 2013. Development and Evaluation of SoySNP50K, a High-Density Genotyping Array for Soybean. PLoS ONE, **8**: e54985. doi:10.1371/journal.pone.0054985.

Song, Q., Hyten, D.L., Jia, G., Quigley, C.V., Fickus, E.W., Nelson, R.L., and Cregan PB. 2015. Fingerprinting soybean germplasm and its utility in genomic research. G3: Genes, Genomes, Genetics **5**: 1999–2006.

Song, Q.J., Marek, L.F., Shoemaker, R.R., Lark, K.G., Concibido, V.C., Delannay, X., Specht, J.E., and Cregan, P.B. 2004. A new integrated genetic linkage map of the soybean. Theor. Appl. Genet. **109**: 122–128.

Soto-Cerda, B.J., and Cloutier, S. 2013. Outlier loci and selection signatures of simple sequence repeats (SSRs) in flax *(Linum usitatissimum* L.). Plant Mol. Biol. Rep. **31**: 978–990.

SoyBase. 2017. Soybean breeders toolbox. Accessed October 2017 from http://soybase.org.

Tanksley, S.D., and McCouch, S.R. 1997. Seed banks and molecular maps: unlocking genetic potential from the wild. Science, **277**: 1063–1066.

Teschke, M., Mukabayire, O., Wiehe, T., and Tautz, D. 2008. Identification of selective sweeps in closely related populations of the house mouse based on microsatellite scans. Genetics, **180**: 1537–1545.

Teshima, K.M., Coop, G., and Przeworski, M. 2006. How reliable are empirical genomic scan for selective sweeps? Genome Res. **16**: 702–712.

Tian, Z., Wang, X., Lee, R., Li, Y., Specht, J.E., Nelson, R.L., McClean, P.E., Qiu, L., and Ma, J. 2010. Artificial selection for determinate growth habit in soybean. PNAS, **107**: 8563–8568.

Varshney, R.K., Terauchi, R. and McCouch, S.R. (2014) Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding. PLoS Biol. 12, e1001883.

Vigouroux, Y., McMullen, M., Hittinger, C.T., Houchins, K., Schulz, L., Kresovich, S., Matsuoka, Y., Doebley, J. 2002. Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. PNAS, **99**: 9650–9655.

Wright, S. 1951. The genetical structure of populations. Ann Eugenics **15**: 323–354.

Wright, S.I., Bi, I.V., Schroeder, S.G., Yamasaki, M., Doebley, J.F., McMullen, M.D., and Gaut, B.S. 2005. The effect of artificial selection on the maize genome. Science, **308**: 1310–1314.

Yamasaki, M., Tenaillon, M.I., Bi, I.V., Schroeder, S.G., Sanchez-Villeda, H., Doebley, J.F., Gaut, B.S., and McMullen, M.D. 2005. A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. Plant Cell, **17**: 2859–2872.

Zhang, X.Y., Tong, Y.P., You, G., Hao, C., Ge, H., Wang, Y., Dong, Y., Li, Z. 2007. Hitchhiking effect mapping; A new approach for discovering agronomic important genes. Agri. Sci. China, **6**: 255–264.

Zhuang, Q.S. 2003. Chinese wheat improvement and pedigree analysis. Agricultural Press, Beijing.

Zhu, Q., Zheng, X., Luo, J., Gaut, B.S., and Ge, S. 2007. Multilocus analysis of nucleotide variation of Oryza sativa and its wild relatives: Severe bottleneck during domestication of rice. Mol. Biol. Evol. **24**: 875–888.

Table 1. Trait QTL associated with fixed allele frequencies (>0.95) throughout the pedigree of OAC Bayfield.

Chromosome	Marker	Allele Frequency	Marker Map Position (cM)	Trait QTL Association Reported
1	Satt 370	0.98	60.99	Pod wall wt 1-1, Pod wall wt 1-2, Sd wt 37-5,
3	Sat 379	0.98	4.33	None reported
3	Satt234	1.00	84.59	None reported
3	Satt 22	1.00	102.05	Lf wdth 9-7, Lf wdth 8-15, Oil 24-34, Oil 28-3, Sd length 1-4 Pl ht 26-17 Stear 6-5, Ole 6-1, Linole 6-1, Lflt chlorophyll 1-3, Oil 36-9, Pod wall wt 1-4
5	Satt 174	0.98	88.58	Oil 8-1, Sd wt 7-3, CWP 1-1, Ara/Gal 1-1, Pectin 1-1
7	Satt 494	1.00	71.7	Sd length 1-7
8	Satt315	1.00	45.29	Sd length 1-3, Glycitein 9-2, Row spacing response 1-4, Genistein 9-2, Daidzein 9-2
8	Satt470	1.00	116.73	Sd wt 10-2, Rag 2-1
8	Satt538	0.95	159.63	Daidzein 2-5, Sd wt 36-1, Sd wt 34-1
8	Satt378	0.97	165.72	R3 1-3, R7 1-3
13	Satt395	0.97	146.41	None reported
15	Satt213	1.00	3.72	None reported
17	Satt328	1.00	16.76	None reported
19	Satt238	1.00	19.93	Lf lgth 9-8, Lf lgth 8-7, Lf wdth 8-8, Lflt shape 6-14, Pod dehis 3-7, Linolen 8-3

Table 2. Trait QTL classification and abundance for QTL associated with fixed allele frequencies (> 0.95) throughout the pedigree of OAC Bayfield.

Trait	Plant Architecture	Yield	Maturity	Oil	Seed Composition	Seed Quality	Disease	Other
QTL Name	Pod wall wt 1-1	Sd wt 37-5	R3 1-3	Oil 24-34	Stear 6-5	Sd length 1-4	Rag 2-1	Lflt chlorophyll 1-3
	Pod wall wt 1-2	Sd wt 7-3	R7 1-3	Oil 28-3	Ole 6-1	Sd length 1-3		CWP 1-1
	Lf wdth 9-7	Sd wt 10-2		Oil 36-9	Linole 6-1	Sd length 1-7		Row spacing response 1-4
	Lf wdth 8-15	Sd wt 36-1		Oil 8-1	Pectin 1-1			Pod dehis 3-7
	Pl ht 26-17	Sd wt 34-1			Glycitein 9-2			
	Pod wall wt 1-4				Genistein 9-2			
	Lf lgth 9-8				Daidzein 9-2			
	Lf lgth 8-7				Daidzein 2-5			
	Lf wdth 8-8				Linolen 8-3			
	Lflt shape 6-14				Ara/Gal 1-1			
QTL Total	**10**	**5**	**2**	**4**	**10**	**3**	**1**	**4**

Table 3. Trait QTL associated with SSR markers significant ($P<0.05$) for the ln RH and F_{st} tests.

Chromosome	SSR Marker	Marker Map Position (cM)	Trait QTL Association Reported
1	Satt 531	40.86	Sd hrd 1-2, Fflr 16-1
2	Satt 274	116.34	NitR5 1-4, Oil 19-1, sd-Gly 1-3, sd-Thr 1-3, sd-Glu 1-3, sd-Tyr 1-2 sd-Phe 1-2 sd-Leu 1-2, sd-Arg 1-1, Oil 24-3, Sd crack 3-3, Sd wt 37-10, Phytoph 10-3
3	Satt 387	53.25	Sclero 3-16, Pl ht 17-5, Pod mat 16-4, Sd yld 15-12, Shoot wt-dry 1-3, Shoot wt-fresh 1-4
4	Satt 578	65.08	Pod mat 8-5, Prot 7-2, Sd set 1-5, Prot 19-1, Stm str 1-4, Sd height 1-12, Sd volume 1-9, Sd length 1-13, Linolen 6-2, Oil 29-1
5	Satt 050	46.45	Sd abrt 1-6, Lf wdth 9-1, Lf wdth 8-1, Pod dehis 3-1, Ole 6-5
6	Satt 319	113.40	Fflr 12-2, Pl ht 21-2, Ldge 18-2, Node num 2-2, Sd yld 19-1, Sd yld 22-11, Glycitein 6-2
6	Satt 643	94.65	Stm str 1-5
6	Satt 357	151.91	Pl ht 27-1, Rt lgth 2-1,
7	Satt 636	5.00	None reported
9	Satt 242	14.35	Glycitein 6-4
11	Satt 426	28.30	Pl ht 24-4, Branching 1-1, Branching 1-2, Node num 3-3, Node num 3-4, Pod num 3-1
11	Satt 597	80.90	Pod mat 18-1
15	Satt 369	56.30	Lflt shape 9-6, Lflt shape 8-10, Sd volume 1-2, Sd width 1-2
16	Satt 249	11.74	Isoflv 1-4, Daidzein 2-7, sd-Glu 1-4, sd-Phe 1-3, Drought index 1-5, Stear 4-2, Isoflv 4-1
17	Satt 186	105.40	Sd yld 15-7, Pod mat 19-2, Glycitein 7-6, Oil 30-6
17	Satt 386	125.00	Isoflv 4-4
19	Satt 561	71.40	Sd yld 8-1, Reprod 5-6, Phytate 2-2, Pod mat 24-4, Sd wt 34-7, Sd wt 36-7, cqPhytate-002
20	Sat 104	65.60	Fe effic 13-3
20	Sat 419	98.10	Sd yld 22-5, Ole 5-3

Table 4. Trait QTL classification and abundance for QTL associated with SSR markers significant (P<0.05) for the In RH and F_{st} tests.

	Plant					Seed	Seed		
Trait	**Architecture**	**Yield**	**Maturity**	**Oil**	**Protein**	**Composition**	**Quality**	**Disease**	**Other**
QTL Name	Lf wdth 9-1	Sd yld 22-5	Pod mat 24-4	Oil 19-1	sd-Gly 1-3*^{z}	Glycitein 6-2	Sd volume 1-2	Phytoph 10-3	NitR5 1-4
	Lf wdth 8-1	Sd yld 15-12	Pod mat 16-4	Oil 24-3	sd-Thr 1-3*	Glycitein 6-4	Sd volume 1-9	Sclero 3-16	Shoot wt-dry 1-3
	Lflt shape 9-6	Sd yld 15-7	Pod mat 8-5	Oil 29-1	sd-Glu 1-3*	Glycitein 7-6	Sd width 1-2		Shoot wt-fresh 1-4
	Lflt shape 8-10	Sd yld 8-1	Pod mat 18-1	Oil 30-6	sd-Tyr 1-2*	Isoflv 4-1	Sd hrd 1-2		Pod dehis 3-1
	Pl ht 21-2	Sd yld 19-1	Pod mat 19-2		sd-Phe 1-2*	Isoflv 4-4	Sd crack 3-3		Rt lgth 2-1
	Pl ht 17-5	Sd yld 22-11	Fflr 16-1		sd-Leu 1-2	Isoflv 1-4	Sd height 1-12		Drought index 1-5
	Pl ht 27-1	Sd wt 37-10	Fflr 12-2		sd-Arg 1-1*	Daidzein 2-7	Sd length 1-13		Phytate 2-2
	Pl ht 24-4	Sd wt 34-7	Reprod 5-6		sd-Glu 1-4**	Ole 5-3			cqPhytate-002
	Stm str 1-5	Sd wt 36-7			sd-Phe 1-3**	Ole 6-5			Fe effic 13-3
	Stm str 1-4	Sd set 1-5			Prot 7-2	Linolen 6-2			
	Branching 1-1	Sd abrt 1-6			Prot 19-1	Stear 4-2			
	Branching 1-2	Pod num 3-1							
	Node num 3-3								
	Node num 3-4								
	Node num 2-2								
	Ldge 18-2								
QTL Total	**16**	**12**	**8**	**4**	**4**	**11**	**7**	**2**	**9**

^{z} * indicates separate QTL regions; ** indicated that there were two distinct QTL regions where groups of individual amino acids QTL were found on the same chromosome.

List of Figures Legends:

Figure 1. Group assignment of genotypes based on the pedigree of OAC Bayfield. Members of the pedigree were divided into two groups, the *Ancestral Group* and *Current Group* (hatched boxes). The genotypes in the *Current Group* are from various public and private soybean breeding programs as identified by the coloured names. Cultivars with an asterisk were not available for study. Adapted from Grainger and Rajcan (2014).

Figure 2. Plot of standardized ln RH values for the 142 SSR markers. Significant values (P <0.05) are indicated by arrows along with the marker name.

Figure 3. Plot of F_{st} values for the 142 SSR markers. Significant values (P <0.05) are indicated by arrows along with the marker name.

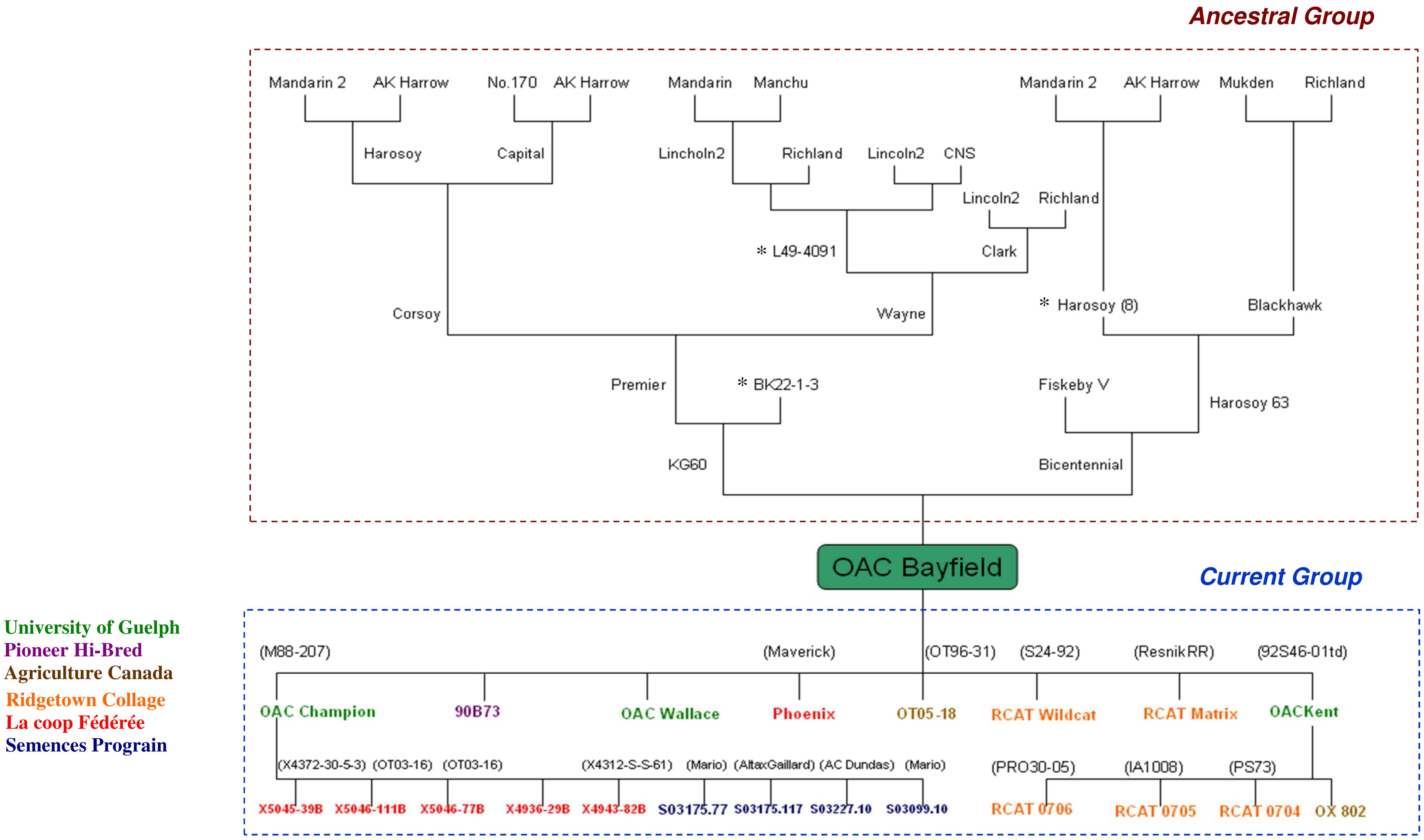
Ancestral Group
Mandarin 2
AK Harrow
No.170
AK Harrow
Mandarin
Manchu
Mandarin 2
AK Harrow
Mukden
Richland
Harosoy
Capital
Lincholn2
Richland
Lincoln2
CNS
Lincoln2
Richland
* L49-4091
Clark
Corsoy
Wayne
* Harosoy (8)
Blackhawk
Premier
* BK22-1-3
Fiskeby V
Harosoy 63
KG60
Bicentennial
OAC Bayfield
Current Group
University of Guelph
Pioneer Hi-Bred
Agriculture Canada
Ridgetown Collage
La coop Fédérée
Semences Prograin
(M88-207)
(Maverick)
(OT96-31)
(S24-92)
(Resnik RR)
(92S46-01td)
OAC Champion
90B73
OAC Wallace
Phoenix
OT05-18
RCAT Wildcat
RCAT Matrix
OACKent
(X4372-30-5-3)
(OT03-16)
(OT03-16)
(X4312-S-S-61)
(Mario)
(AltaxGaillard)
(AC Dundas)
(Mario)
(PRO30-05)
(IA1008)
(PS73)
X5045-39B
X5046-111B
X5046-77B
X4936-29B
X4943-82B
S03175.77
S03175.117
S03227.10
S03099.10
RCAT 0706
RCAT 0705
RCAT 0704
OX 802

Fig. 1

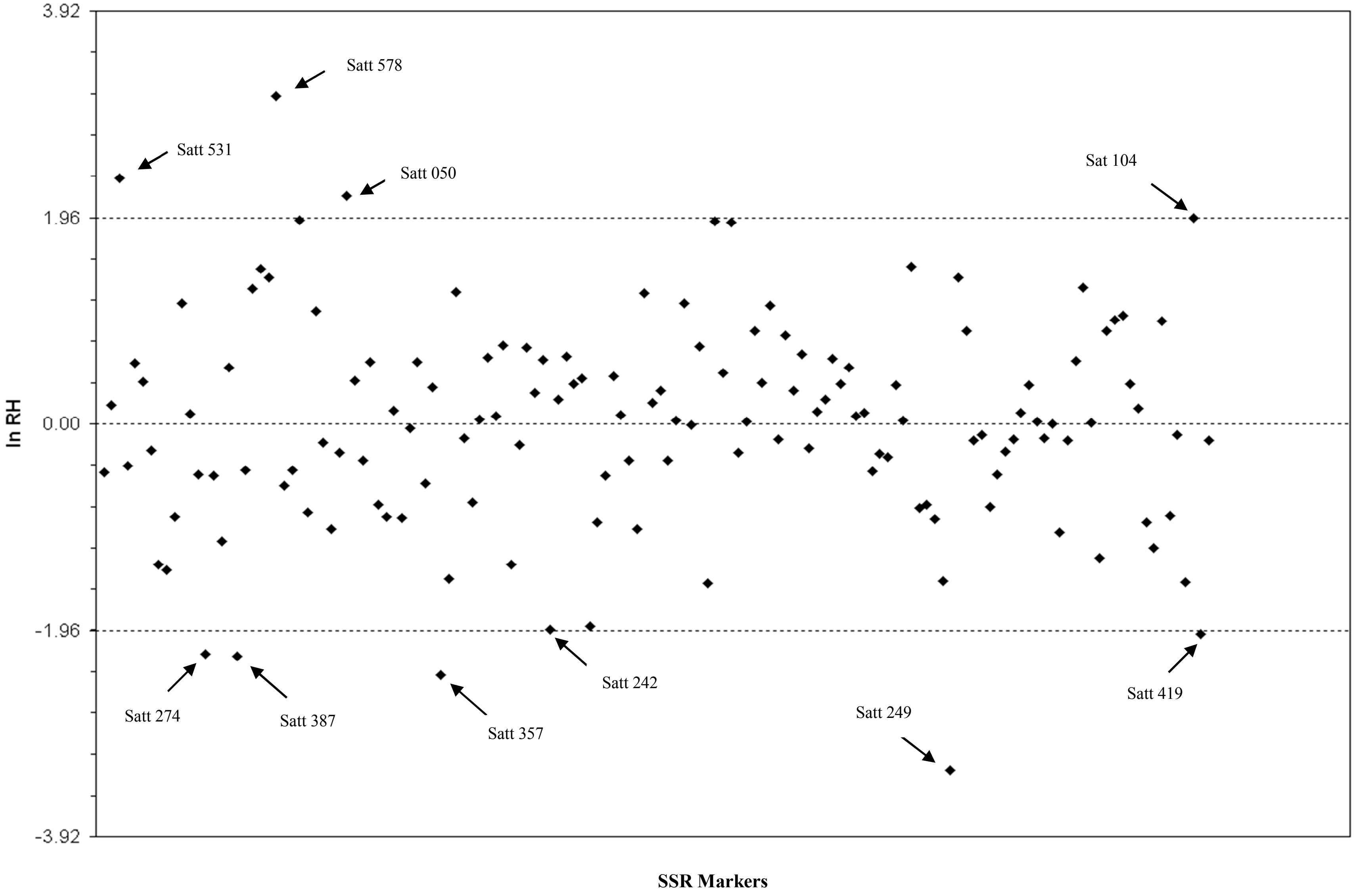
3.92
1.96
0.00
-1.96
-3.92
ln RH
SSR Markers
Satt 578
Satt 531
Satt 050
Sat 104
Satt 274
Satt 387
Satt 357
Satt 242
Satt 249
Satt 419

Fig. 2.

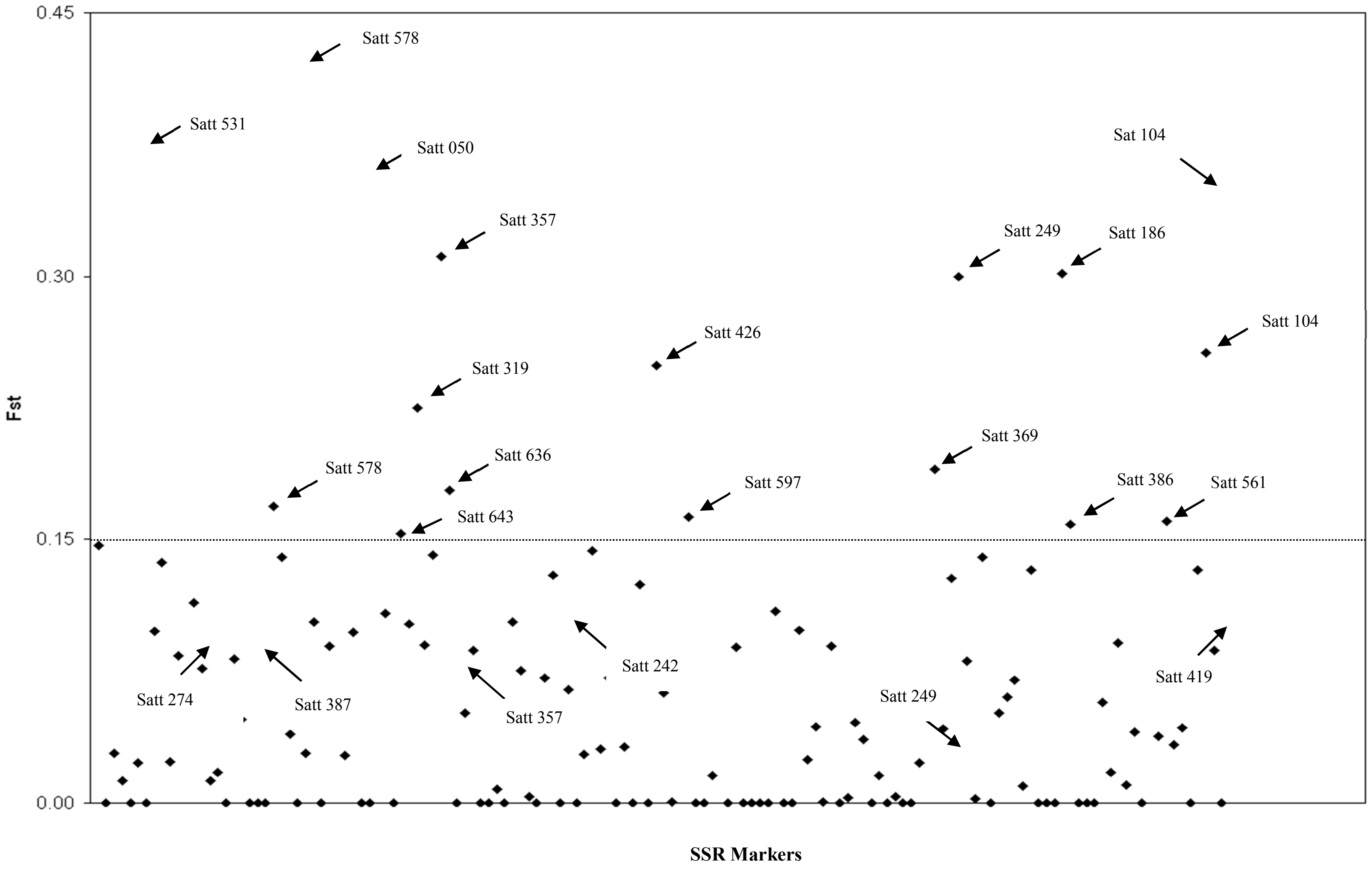
0.45
0.30
0.15
0.00
Fst
SSR Markers
Satt 578
Satt 531
Satt 050
Sat 104
Satt 357
Satt 249
Satt 186
Satt 104
Satt 426
Satt 319
Satt 369
Satt 636
Satt 578
Satt 597
Satt 386
Satt 561
Satt 643
Satt 242
Satt 274
Satt 387
Satt 357
Satt 249
Satt 419

Fig. 3.