Analysis of >1000 single nucleotide polymorphisms in geographically matched samples of landrace and wild barley indicates secondary contact and chromosome-level differences in diversity around domestication genes

Joanne Russell¹, Ian K. Dawson¹, Andrew J. Flavell², Brian Steffenson³, Eva Weltzien⁴, Allan Booth¹, Salvatore Ceccarelli⁵, Stefania Grando⁵ and Robbie Waugh¹

¹Scottish Crop Research Institute (SCRI), Invergowrie, Dundee DD2 5DA, UK; ²University of Dundee at SCRI, Invergowrie, Dundee DD2 5DA, UK; ³Department of Plant Pathology, University of Minnesota, St. Paul, MN 55108, USA; ⁴International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Mali Regional Office, B.P. 320, Bamako, Mali; ⁵International Center for Agricultural Research in the Dry Areas (ICARDA), P.O. Box 5466, Aleppo, Syria

Author for correspondence:
Joanne Russell
Tel: +44 1382 562 731
Email: Joanne.Russell@scri.ac.uk

Received: 8 December 2010
Accepted: 21 February 2011

New Phytologist (2011) 191: 564–578

Key words: crop domestication, Fertile Crescent, genetically matched sampling, human and natural selection, hybridization and introgression, landrace and wild barley, SNP analysis.

Summary

• Barley is a model species for the investigation of the evolution, adaptation and spread of the world’s important crops. In this article, we describe the first application of an oligonucleotide pool assay single nucleotide polymorphism (SNP) platform to assess the evolution of barley in a portion of the Fertile Crescent, a key region in the development of farming.
• A large collection of > 1000 genetically mapped, genome-wide SNPs was assayed in geographically matched landrace and wild barley accessions (N = 448) from Jordan and Syria.
• Landrace and wild barley categories were clearly genetically differentiated, but a limited degree of secondary contact was evident. Significant chromosome-level differences in diversity between barley types were observed around genes known to be involved in the evolution of cultivars. The region of Jordan and southern Syria, compared with the north of Syria, was supported by SNP data as a more likely domestication origin.
• Our data provide evidence for hybridization as a possible mechanism for the continued adaptation of landrace barley under cultivation, indicate regions of the genome that may be subject to selection processes and suggest limited origins for the development of the cultivated crop.

Introduction

Barley, one of a few key crops domesticated during the dawn of agriculture in the Fertile Crescent (Harlan, 1975), continues to play an important role in farming within the region and is the fourth most important cereal worldwide after maize, rice and wheat. As a result, in recent years, there has been considerable investment in the development of new methods for the study of the biology of the species, including powerful, novel approaches for molecular-level analysis. Genotyping techniques that provide large numbers of markers with wide genome coverage have become available, including high-density diversity arrays (DArT; Wenzl et al., 2006) and single nucleotide polymorphisms (SNPs) implemented on an Illumina oligonucleotide pool assay platform (Rostoks et al., 2005, 2006; Close et al., 2009; Waugh et al., 2009). Such SNPs are considered to be markers of choice in population biology studies, although examples of application outside human genetics remain rather limited to date (Seddon et al., 2005; Kijas et al., 2009; Li et al., 2010). These new marker resources provide great scope for high-resolution, chromosome-level delineation of the processes of crop evolution, which are only beginning to be explored.
In the current study, we investigate genome- and chromosome-level genetic structure within and between landrace (*Hordeum vulgare* ssp. *vulgare*) and wild (*H. vulgare* ssp. *spontaneum*) categories of barley in a portion of the Fertile Crescent. Our objective was to more fully understand the evolution of the crop, and thus help to guide strategies for future cultivar development and conservation based on resources in the region. Farmers change the genetic composition of crops in comparison with their wild progenitors through random sampling processes and by selection at targeted and linked loci. In addition, dynamic environmental conditions lead to natural selection continuously acting on both landrace and wild materials in ways that reinforce or contradict human-mediated selection (Burger et al., 2008; Mercer & Perales, 2010). To investigate diversity and structure, we describe the first application of an SNP oligonucleotide pool assay platform (Close et al., 2009) to undertake a comprehensive genome-wide investigation of the relationship between landrace and wild barley in an area covering over one-third of the Fertile Crescent region. In undertaking this research, we established the largest dataset yet available of mapped molecular markers for a cultivated–natural barley comparison. Our analysis was also more powerful than previous studies that have compared barley types because it was based on geo-referenced matched sampling at a country level, which eliminates some of the potentially confounding effects of geography on domestic vs wild plant comparisons (Dawson et al., 2008).

In total, the present study involved 448 accessions, 317 of landrace material and 131 of wild barley, which were collected from across Jordan and Syria. The landrace populations genotyped here were chosen for analysis because they constitute a key resource for the study of environmental adaptation in the cultivated gene pool (Weltzien, 1988). In addition, they may be important sources of new traits for introduction into advanced cultivars (Ceccarelli et al., 1987; van Leur et al., 1989). The studied landraces are also threatened within the Fertile Crescent because of anthropogenic climate change and other challenges to traditional farming, meaning that genetic data are all the more important for the design of conservation strategies (Mercer & Perales, 2010). The same landrace stands have been the subject of past nuclear and chloroplast diversity research, but based on much more limited numbers of molecular marker loci; previous studies have demonstrated geographically based genetic differentiation in these landraces and have suggested the potential of higher resolution genotyping – as carried out in the research presented here – to provide more detailed information on the processes of barley evolution, especially if landrace and wild stands are compared directly with each other (Russell et al., 2003). The SNP data collected here illustrate how such markers can be used to begin to further explore evolutionary questions in other crops of historical and current importance.

**Materials and Methods**

**Barley collections**

A concern during current sampling was to minimize the confounding effects of large-scale geography that are often not fully accounted for in comparisons of landrace and wild barley stands (e.g. Badr et al., 2000; Kilian et al., 2006). For the purpose of effective stratification of genetic variation between landrace and wild barley, therefore, we sampled both categories as closely as possible from within the same geographic areas in Jordan and Syria (Fig. 1; Table S1). In these countries, wild barley is sometimes found in mixed populations with landraces, around culti-

---

**Fig. 1** Geographic locations of landrace and wild barley sample sites from the same areas of Jordan and Syria. Twenty-four landrace populations, comprising 317 accessions in total (an average of c. 13 individuals per population), and 131 individual wild barley accessions, were included in the analysis (see text for further details on the different sampling approaches applied to barley categories). For the purposes of comparison in certain analyses (see Materials and Methods), accessions were divided into three adjacent geographic regions of approximately equal area, each containing a minimum of 100 individuals in total: Region 1, northeastern Syria; Region 2, northwestern Syria; Region 3, Jordan–southern Syria. Eight, seven and nine landrace populations were sampled, respectively, from these regions, with 112 (14.0), 84 (12.0) and 121 (13.4) individuals tested, respectively (mean population sample sizes given in parentheses). Individual wild accessions were assigned to regions based on geographic proximity to landrace stands; 14, 29 and 88 individuals were collected from Regions 1, 2 and 3, respectively. Greater sampling of wild barley from Jordan–southern Syria reflects higher current-day prevalence in this region. More information on the sampled accessions is given in Supporting Information Table S1.
vated barley fields, or on its own in discrete stands. Sampling was subdivided into three adjacent regions of approximately equal dimensions (each of c. 6 squared degrees) that each contained a minimum of 100 individuals in total for testing. Stratification was determined first by the natural geography of Jordan and Syria, and second (and most importantly) by the sampling strategy adopted during original landrace collections. Twenty-four two-row landrace populations (each originally representing a single cultivated field) were chosen as a representative geographic subset from an extensive systematic collection of landraces from across Jordan and Syria, undertaken by one of the authors (EW) in 1980 (Weltzien, 1988). This collection is currently maintained by the International Center for Agricultural Research in the Dry Areas (ICARDA) in Syria. According to farmers interviewed by EW at the time of sampling, these collections represent old indigenous landraces. In total, one individual from each of 317 lines (each from a single spike) derived from this collection was included in the present study (i.e. a mean of c. 13 individuals from each of the 24 populations represented in the analysis). For comparison, 131 wild barley individuals were taken from geo-referenced collections made in the same geographic areas. Wild material was a subset of the World Barley Diversity Collection (WBDC), which consists of over 300 accessions sampled from the Fertile Crescent and more widely; our sampling involved the use of accessions that had been sampled from Jordan and Syria (Steffenson et al., 2007; Roy et al., 2010). Most wild accessions also originated from ICARDA and were assembled by Dr Jan Valkoun, the former curator of the gene bank there.

The material analyzed here represents the most comprehensive collection of landrace and wild barley accessions that is currently available for a comparison of categories from across Jordan and Syria. There is, however, a difference in sampling approach, reflecting the methods adopted historically during collection: landraces were collected as populations and wild accessions as individuals at sites. There is therefore a compromise in our decision to adopt a geographically matched collection approach at a ‘macro’ scale and the material available for testing locally. It is also the case, however, that there is no perfect strategy for strictly equivalent sampling because of the alternative histories of categories. We consider that the different approaches adopted here do not detract significantly from our ability to provide an overall understanding of the relationship between landrace and wild barley in the two countries. The involvement of multiple populations of landrace barley, and the use of mapped markers that provide for chromosome-level comparisons, should help limit the effects of sampling differences. Furthermore, although there is nesting into landrace populations, the average distance between pairs of accessions for both categories of material (based on individual-by-individual geographic distance matrices) is very similar, at 3.0 and 2.7 degrees for landraces and wild accessions, respectively: therefore, nesting of the former category has not led to lower geographic distances on average; distances were in fact slightly greater in cultivated than in wild germplasm.

SNP analysis

SNP analysis involved the use of the barley oligonucleotide pool assay 1 platform (BOPA1, composed of 1536 SNPs). The development of this platform and the generation of a consensus linkage map of polymorphisms along barley chromosomes have been described fully elsewhere (Close et al., 2009). DNA from young landrace and wild plants was extracted with the DNeasy plant mini-preparation kit (QiaGen, Hilden, Germany). Genotyping was then undertaken using the Illumina GoldenGate BeadArray technology (Fan et al., 2003), and the resulting data were processed and manually inspected using the BeadStudio 3.1.3.0 software package (Illumina Inc., San Diego, CA, USA). Our most important concerns during data analysis were two-fold: first, to compare barley categories with each other at an overall genome level and, second, to study chromosome-level effects. To this end, we employed a number of statistical approaches, as described below.

Data analysis: overall genome-level comparisons

As noted above, the different sampling approaches (a degree of population nesting for landrace accessions) mean that comparisons of our cultivated and wild accessions should be interpreted with some caution. Bearing this proviso in mind, a Bayesian method, HICKORY 1.1 (Holsinger et al., 2002), set at default parameters, was used to generate Nei’s (1978) genetic diversity estimate ($H$) with ‘credibility intervals’ for landrace and wild barley by geographic region (treating regions as populations). FSTAT 2.9.4 (Goudet, 1995), modified through personal communication with Jerome Goudet to deal with > 1000 markers, was also used to generate allelic richness values ($A$) for the same samples as an alternative measure of diversity, correcting for different population sizes by rarefaction.

To assess patterns of differentiation between barley categories, a Bayesian structure analysis of landrace and wild individuals was undertaken with the STRUCTURE 2.2 software package (Falush et al., 2007). Analysis was based on 25 000 ‘burn-in’ replications and a further 25 000 Markov chain Monte Carlo (MCMC) steps (sufficient to ensure the convergence of key parameters) and different levels of $K$ (the number of assumed groups); other options were kept at default settings. STRUCTURE analysis was not originally designed for use on predominantly inbreeding species, such as barley, and results should therefore be interpreted with caution, and compared with other methods.
of analysis [analysis with the InStruct software package, designed to help take account of selfing (Gao et al., 2007) produced similar results not reported here]. Despite its limitations, the STRUCTURE approach has been used widely on barley to reveal interesting biological features (e.g. Rostoks et al., 2006; Morrell & Clegg, 2007; Saisho & Purugganan, 2007; Roy et al., 2010). Principal co-ordinate analysis (PCoA) based on Hamming’s distance (= 1-simple matching) between pairs of genotypes was also undertaken to structure diversity among individuals using the PAST 1.91 software suite (Hammer et al., 2001). Calculation of the standardized Mantel statistic (R MT, using a randomization test with 5000 permutations to assign significance to estimates) for genetic and geographic (based on latitude and longitude co-ordinates) distances between landrace populations was also undertaken in PAST 1.91. A similar analysis was undertaken for wild accessions based on individual, rather than population, collection site coordinates. Finally, the generation of pairwise F ST values between barley categories and regions (treating regions as populations for each category) was carried out with ARLEQUIN 3.11 (Excoffier et al., 2005), and significance values for F ST estimates were assigned on the basis of a randomization test with 10 000 permutations.

Data analysis: chromosome-level effects

The investigation of human and natural selection processes can take a number of approaches for analysis (see Oleksyk et al., 2010 for the range of techniques applied in human genetics). Our assessment of chromosome-level effects for mapped markers in barley used several methods. ARLEQUIN 3.11 was used to generate F ST values between landrace and wild categories for polymorphic SNPs, treating categories as populations and based on Weir & Cockerham’s (1984) θ statistic. These F ST estimates were plotted against map positions along each of barley’s seven chromosomes with EXCEL 2007. The F ST value for each of a group of category-diagnostic SNPs (defined as having a frequency difference between landrace and wild categories Δfw ≥ 0.9, see more below) was then compared with the mean F ST value of the 10 markers that flanked it on the chromosome (five adjacent markers each side of the diagnostic marker). This was performed in order to assess whether subspecies’ differentiation extended to the context of the neighboring chromosomal regions, an important issue in understanding evolutionary processes.

To assess genetic variation differences along chromosomes, EXCEL 2007 was employed for mapped SNPs to generate and plot Nei’s (1978) genetic diversity (H) for landrace and wild categories (treating each category as a single sample). To provide an indication of diversity trends, estimates were plotted as rolling averages of 15 consecutive values stepped by each SNP, based on the inclusion of seven estimates on each side of a central marker to which the mean value was assigned (the ‘sliding window’ method; Rostoks et al., 2006). To assess the significance of any diversity difference observed between landrace and wild stands, a two-tailed t-test using the same sliding window approach for 15 consecutive individual locus H values was also undertaken in EXCEL 2007. To estimate average linkage disequilibrium (LD) (ZnS, the mean of r2 values of all paired SNP comparisons; Hill & Robertson, 1968; Kelly, 1997) along landrace and wild barley chromosomes, DNASP 5.00.07 (Librado & Rozas, 2009) was used. Estimates were based on a minimum minor allele frequency of 0.1 for mapped markers. The significance of ZnS was determined by coalescent simulation based on the calculated recombination parameter R and 1000 replications.

Results and Discussion

SNPs reveal higher overall genetic diversity in wild than in landrace barley

The 1135 of 1536 BOPA1 SNPs that could be reliably called in at least 95% of both landrace and wild barley individuals were retained for data analysis, providing, in total, > 500 000 genotype assignments (see Table S1). Of the 1135 retained loci, 1074 were polymorphic. The mean levels of missing data (ambiguous calls) and the occurrence of heterozygous states across these loci were very low (< 0.2% and < 0.006%, respectively). Genome-level genetic diversity in wild barley (no account for chromosome-level partitioning) was greater than in landrace barley (mean H across regions of 0.200 and 0.161, respectively; Table 1), although the difference in allelic richness between categories (wild, 1.886; landrace, 1.768) was not statistically significant (P = 0.386, based on a two-tailed randomization test with 5000 permutations in FSTAT). This test was, however, limited in power by having only three regions within categories against which to make comparisons. A further test of the difference in diversity (based on H values) between landrace and wild barley categories was possible through nesting SNPs on a chromosome-by-chromosome basis (mean H estimates by chromosome given in Table 2). This test offered greater statistical power as values could be compared across seven chromosomes. In this case, a two-tailed t-test revealed a significant difference between categories (P < 0.01).

Our BOPA1 data were consistent with those of Russell et al. (2004), who sequenced 23 genes associated with grain germination. They found higher nucleotide variation in a small subset of wild accessions compared with both landrace barley from Syria and Jordan and advanced spring cultivars (mean haplotype diversity: wild, 0.517; landrace, 0.386; advanced cultivars, 0.284). In our analysis, landrace material had c. 80% of the diversity of wild barley (based on
genome-level $H$ values), whereas, in Russell et al.’s (2004) study, the equivalent comparison showed a rather similar 75% of diversity in the landrace accessions. Our estimates were, however, in contrast with those of Jana & Pietrzak (1988), who undertook the only other significant site-matched comparison of wild and landrace barley within the Fertile Crescent. They found that isozyme diversity in wild accessions was somewhat lower than in landrace material.

Table 1: Summary data for genome-level analysis of 1135 single nucleotide polymorphism (SNP) loci scored in landrace and wild categories of barley collected from the same geographic regions in Jordan and Syria.

<table>
<thead>
<tr>
<th>Material</th>
<th>$N$</th>
<th>PL</th>
<th>$H$ (95% credible interval)$^a$</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total accessions</td>
<td>448</td>
<td>1074</td>
<td>0.198 (0.197–0.199)</td>
<td>1.768$^e$</td>
</tr>
<tr>
<td>Landrace accessions$^b$</td>
<td>317</td>
<td>914</td>
<td>0.185 (0.183–0.187)</td>
<td>1.548$^f$</td>
</tr>
<tr>
<td>Region 1 (8 populations)$^c$</td>
<td>112</td>
<td>812</td>
<td>0.170 (0.167–0.172)</td>
<td>1.520$^f$</td>
</tr>
<tr>
<td>Region 2 (7 populations)$^c$</td>
<td>84</td>
<td>787</td>
<td>0.129 (0.128–0.131)</td>
<td>1.410$^f$</td>
</tr>
<tr>
<td>Region 3 (9 populations)$^c$</td>
<td>121</td>
<td>701</td>
<td>0.161</td>
<td>1.493</td>
</tr>
<tr>
<td>Mean across regions$^d$</td>
<td>105.6</td>
<td>767</td>
<td>0.161</td>
<td>1.493</td>
</tr>
<tr>
<td>Wild accessions$^b$</td>
<td>131</td>
<td>1006</td>
<td>0.229 (0.227–0.231)</td>
<td>1.886$^e$</td>
</tr>
<tr>
<td>Region 1$^c$</td>
<td>14</td>
<td>540</td>
<td>0.164 (0.159–0.169)</td>
<td>1.461$^f$</td>
</tr>
<tr>
<td>Region 2$^c$</td>
<td>29</td>
<td>769</td>
<td>0.206 (0.203–0.210)</td>
<td>1.606$^f$</td>
</tr>
<tr>
<td>Region 3$^c$</td>
<td>88</td>
<td>958</td>
<td>0.230 (0.228–0.232)</td>
<td>1.684$^f$</td>
</tr>
<tr>
<td>Mean across regions$^d$</td>
<td>43.7</td>
<td>756</td>
<td>0.200</td>
<td>1.584</td>
</tr>
</tbody>
</table>

$N$, number of individuals tested; PL, number of polymorphic loci; $H$, mean genetic diversity calculated according to Nei’s (1978) estimate and using a Bayesian method; A, allelic richness after rarefaction.

$^a$Based on polymorphic markers only.

$^b$All accessions within a barley category treated as a single sample.

$^c$All accessions within a region treated as a single sample.

$^d$Arithmetic average of three regional values.

$^e$Rarefaction to a minimum sample size of 118 (minimum number of complete genotypes at any one SNP in either barley category).

$^f$Rarefaction to a minimum sample size of 10 (minimum number of complete genotypes at any one SNP in any region in either barley category).

Table 2: Summary data for chromosome-level analysis of genetic diversity ($H$), genetic differentiation ($F_{ST}$), and linkage disequilibrium ($Z_{ns}$) for landrace and wild categories of barley collected from the same geographic regions in Jordan and Syria and scored for 1135 single nucleotide polymorphism (SNP) loci.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>$H$ landrace$^a$</th>
<th>$H$ wild$^a$</th>
<th>Mean $F_{ST}$ all SNPs$^b$</th>
<th>Mean $F_{ST}$ diagnostic SNPs$^c$</th>
<th>Mean $F_{ST}$ 10 flanking SNPs$^d$</th>
<th>$Z_{ns}$ landrace$^b$</th>
<th>$Z_{ns}$ wild$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1H</td>
<td>0.172 (121)</td>
<td>0.218 (121)</td>
<td>0.270 (119)</td>
<td>0.936 (5)</td>
<td>0.315</td>
<td>0.120 (19)***</td>
<td>0.023 (14)</td>
</tr>
<tr>
<td>2H</td>
<td>0.209 (179)</td>
<td>0.208 (179)</td>
<td>0.238 (164)</td>
<td>No diagnostic</td>
<td>Not available</td>
<td>0.053 (42)***</td>
<td>0.038 (34)***</td>
</tr>
<tr>
<td>3H</td>
<td>0.183 (170)</td>
<td>0.222 (170)</td>
<td>0.250 (164)</td>
<td>0.958 (3)</td>
<td>0.212</td>
<td>0.068 (26)***</td>
<td>0.066 (39)***</td>
</tr>
<tr>
<td>4H</td>
<td>0.160 (132)</td>
<td>0.214 (132)</td>
<td>0.308 (125)</td>
<td>0.958 (6)</td>
<td>0.299</td>
<td>0.080 (18)***</td>
<td>0.031 (20)***</td>
</tr>
<tr>
<td>5H</td>
<td>0.195 (207)</td>
<td>0.219 (207)</td>
<td>0.229 (201)</td>
<td>0.908 (2)</td>
<td>0.268</td>
<td>0.153 (41)***</td>
<td>0.057 (36)***</td>
</tr>
<tr>
<td>6H</td>
<td>0.176 (138)</td>
<td>0.197 (138)</td>
<td>0.266 (129)</td>
<td>0.973 (1)</td>
<td>0.321</td>
<td>0.076 (24)***</td>
<td>0.034 (22)***</td>
</tr>
<tr>
<td>7H</td>
<td>0.213 (112)</td>
<td>0.242 (112)</td>
<td>0.289 (105)</td>
<td>0.913 (4)</td>
<td>0.258</td>
<td>0.066 (29)***</td>
<td>0.022 (26)***</td>
</tr>
<tr>
<td>All</td>
<td>0.187 (1059)</td>
<td>0.217 (1059)</td>
<td>0.260 (1007)</td>
<td>0.940 (21)</td>
<td>0.281</td>
<td>0.073 (199)***</td>
<td>0.028 (191)***</td>
</tr>
</tbody>
</table>

$H$, the mean genetic diversity calculated according to Nei’s (1978) estimate, was calculated for each chromosome employing individual locus values for mapped markers as used to generate chromosome-level diversity profiles (Fig. 5). In $F_{ST}$ calculations, diagnostic SNPs were defined as those with a frequency difference ($D_{lw}$) between landrace and wild categories of $\geq 0.9$ for alternate character states (see Results and Discussion for further information). $F_{ST}$ profiles along chromosomes, from which values in this table were derived, are shown in Fig. 4. $Z_{ns}$ values were calculated based only on those SNPs with a minimum minor allele frequency of 0.1 in both landrace and wild barley and no missing data points across all accessions in a category. The significance of the $Z_{ns}$ values is shown (***, $P < 0.001$; **, $P < 0.01$, NS, not significant).

$^a$Based on all mapped SNPs, number given in parentheses.

$^b$Based on polymorphic SNPs, number given in parentheses.

$^c$Based on diagnostic SNPs, number given in parentheses.

$^d$Based on the five adjacent markers on each side of each diagnostic SNP.

$^e$Value for all mapped SNPs without chromosomal subdivision outside parentheses, arithmetic average from seven chromosomes in square parentheses.
across 12 collection sites in Jordan (mean $H$ of 0.293 and 0.367, respectively; our calculations based on a summary of their given data). Their comparison was, however, constrained by low marker resolution and very limited genome coverage, issues overcome using the BOPA1 platform.

In a comparative study of landraces and modern varieties by Moragues et al. (2010), BOPA1 SNPs showed ascertainment bias towards advanced cultivars, presumably because the barley lines used in SNP identification and screening came mostly from this latter germplasm group (Rostoks et al., 2005, 2006; Close et al., 2009). The landrace material in our current study is clearly more related to advanced cultivars than are the wild accessions used. This raises the prospect of ascertainment bias skewing our landrace–wild comparison, for example through greater ‘pruning’ of rarely polymorphic markers in wild germplasm (Hamblin et al., 2007), which would result in an underestimation of the genetic diversity of the progenitor. We were unable to exclude the possible effects of ascertainment bias from our analysis, but our data that showed higher levels of genetic variation in wild material would suggest that the relative pruning of SNPs in wild compared with landrace barley must be limited. Furthermore, the difference in diversity levels between landrace and wild barley in our study was similar to that found in previous work (Russell et al., 2004). Further consideration of SNP profiles supports the position of limited pruning in wild material as, compared with landraces, wild barley had a smaller fraction of markers with highly asymmetrically distributed character states, and a greater proportion with intermediate frequency differences (see Fig. S1). Future work to better understand the possible effects of ascertainment bias is required and will compare the results from BOPA1 with data from a second array of 1536 SNPs, BOPA2, a platform compiled with greater emphasis on the use of wild barley in SNP selection (Close et al., 2009), as well as with nuclear and chloroplast simple sequence repeat (SSR) and DArT profiles being assembled by the authors. Regardless, it seems unlikely that ascertainment bias will differentially influence the landrace–wild comparison between different geographic regions within the Fertile Crescent, a factor that could be important in the accurate determination of cultivated barley origins.

**SNPs reveal clear differentiation between landrace and wild barley and indicate secondary contact between the two**

STRUCTURE analysis (with $K$ set at 2) of BOPA1 data revealed $Q$ profiles that generally demonstrated a clear difference in the genetic composition of landrace and wild barley in Jordan and Syria (Fig. 2a), with the primary division of diversity between the two categories and a sharp transition in $Q$ values. The two gene pools are therefore, in general, discrete and the pattern revealed by SNPs is illustrative of sympatric speciation. There are, however, exceptions. Of particular note, eight accessions with unusual, mixed $Q$ profiles (setting $\Delta Q_{lw}$ at $\leq 0.5$) were observed (Fig. 2b); four of these intermediate individuals originated from material identified during collection as wild germplasm and four as landrace individuals. PCoA confirmed the STRUCTURE results, with landrace and wild categories of barley again, in general, clearly differentiated into discrete entities (Fig. 3). Again, however, the eight individuals identified as unusual in STRUCTURE were placed in intermediate positions in ordination. In particular, this was the case for the four atypical ‘wild’ accessions identified by STRUCTURE, one of which was collected from northeastern Syria (Region 1) and three of which came from Jordan–southern Syria (Region 3) (refer to Fig. 1 for regional locations). Interestingly, these individuals were separated on the second principal coordinate of PCoA in the same direction.

![Fig. 2](https://www.newphytologist.com/191-564-578)

**Fig. 2** STRUCTURE results ($K = 2$) for genome-level analysis of 1135 single nucleotide polymorphism (SNP) loci in landrace ($N = 317$) and wild ($N = 131$) categories of barley collected from the same geographic regions in Jordan and Syria. Results are ordered by category (a) and by membership coefficient ($Q$) values (b). A clear distinction between landrace and wild categories was evident from the $Q$ profiles. Eight accessions with unusual, mixed profiles ($\Delta Q_{lw}$ set at $\leq 0.5$) are highlighted in (b). Four of these individuals originated from material identified during collection as wild, four as landrace. STRUCTURE analysis for $K = 3$ and 4 (results not shown) revealed that the next most evident levels of differentiation were within landrace material (similar to observations from principal co-ordinate analysis, see Fig. 3).
as the corresponding landrace stands collected from these regions. Separation of the four unusual ‘landrace’ accessions identified in STRUCTURE, although evident in ordination, was less prominent.

PCoA indicated a higher degree of substructuring of genetic variation between sampled regions for landrace than for wild barley. Greater structuring in the former case was confirmed by a comparison of Mantel tests of genetic and geographic distances ($r_M$ of 0.707 and 0.068 for landrace and wild barley, respectively; $P < 0.001$ and $P < 0.05$, respectively). Data on landrace barley were therefore consistent with nuclear and chloroplast SSRs that revealed significant geographic structuring of genetic variation in Jordan and Syria (Russell et al., 2003; based on 21 polymorphic nuclear SSRs and one polymorphic chloroplast SSR). It is important to note, however, that the greater overall spread of values for landrace compared with wild barley in our PCoA (along the second principal coordinate, Fig. 3) should not be interpreted in terms of relative overall genetic diversity of the two categories. As noted above, wild barley had higher total genome-level genetic diversity (Table 1). The apparent anomaly between diversity values and ordination plots appears to be a result of polymorphism being more structured into discrete haplotypes (less randomly assorted) in the cultivated gene pool than in wild barley (Russell et al., 2004), as indicated by the data on LD levels presented below.

Both cultivated and wild barley types display limited outcrossing (Abdel-Ghani et al., 2004). In addition, the brittle rachis of the wild spike (Pourkheirandish & Komatsuda, 2007) shatters before cultivated barley is harvested in sympatric stands. Furthermore, there is marked variation in flowering time between wild barley and landraces (the former generally flower earlier; E. Weltzien, personal observation during collection of tested landraces). All of these factors suggest that hybridization between landrace and wild barley should be rare. Various authors have, however, suggested occasional intermingling in the Fertile Crescent (e.g. Jana & Pietrzak, 1988; Badr et al., 2000; Pourkheirandish & Komatsuda, 2007), although few genetic data are available to confirm or reject this hypothesis. Our SNP data are able to provide data on this question. Despite predominant landrace and wild category separation, as noted above, SNPs revealed intermediate STRUCTURE and PCoA profiles for eight accessions. The most obvious explanation of the origin of these ‘mixed’ genotypes is that they result from a degree of secondary contact between landrace and wild stands in Jordan and Syria. Being able to respond to environmental alterations through hybridization may be an important means of adaptation to anthropogenic climate change.

Fig. 3 Principal co-ordinates for genome-level analysis of 1135 single nucleotide polymorphism (SNP) loci in landrace ($N = 317$) and wild ($N = 131$) categories of barley sampled from the same three geographic regions in Jordan and Syria. Landrace and wild categories were generally clearly differentiated and a degree of genetic structuring between regions was also evident, especially in landrace material. Eight unusual individuals identified in STRUCTURE analysis (four ‘wild’ and four ‘landrace’ accessions, Fig. 2) are circumscribed by category. R1L ($N = 112$), Region 1 landrace; R2L ($N = 84$), Region 2 landrace; R3L ($N = 121$), Region 3 landrace; R1W ($N = 14$), Region 1 wild; R2W ($N = 29$), Region 2 wild; R3W ($N = 88$), Region 3 wild. It should be noted that the greater overall spread of values (along the second principal coordinate) of landrace accessions compared with wild barley does not indicate higher overall genetic diversity in the first category, but higher linkage disequilibrium in this material (see Results and Discussion for further information).
change, alteration of farmer requirements and other current threats to barley in traditional agriculture (Ceccarelli et al., 1987; Jarvis & Hodgkin, 1999; Mercer & Perales, 2010). Our data, especially on the presence of unusual ‘landrace’ accessions, suggest at least a limited role for wild–landrace hybridization in providing future adaptive capacity in locally cultivated barley in Jordan and Syria, as suggested by Ceccarelli et al. (1987). Our estimate of eight accessions as products of hybridization between the two categories may well be an underestimate, as we only considered those accessions with the most mixed Q profiles to be the results of interaction. In future, hypotheses of secondary contact may be tested further by an assessment of 21 ‘category-diagnostic’ SNPs identified in this study, for which a frequency difference (Δf) between landrace and wild categories of ≥ 0.9 for alternate character states was observed (see Table S1). Clearly, the determination of these 21 markers as diagnostic of categories assumes that the bulk of the collections analyzed in our study were identified correctly when they were first sampled, and are indeed ‘true’ landrace or ‘true’ wild accessions, as otherwise there is a danger of circularity in specifically assigning marker states (for a discussion on this, see, for example, Dawson et al., 1996; Lowe et al., 2000). We are not able to eliminate circularity entirely, but multiple field trials of (apparent) landrace populations have confirmed the identity of the majority as such (Ceccarelli et al., 1987; Weltzien, 1988; van Leur et al., 1989), and WBDC accessions have been morphologically typed in nursery experiments as true to type (B. Steffenson, personal observations).

To explore further the evolutionary relationship between landrace and wild barley subspecies, we assessed the difference in chromosomal context of the 21 category-diagnostic SNPs through the calculation of individual locus FST values of all mapped markers (Table 2 and Fig. 4). As expected on the basis of their asymmetric distributions, the FST values for diagnostic markers were outliers, with values ranging from 0.885 to 0.977 and a high mean of 0.940 (Table 2). More interestingly, the mean FST value for the 10 SNPs that flanked each of our diagnostic markers (0.281) was, however, only marginally different from the mean value for all 1007 mapped SNPs (0.260). Our data therefore indicated that diagnostic SNPs are not set in chromosomal regions that are otherwise particularly differentiated between wild and cultivated barley for the BOPA1 platform. There is no evidence that diagnostic SNPs are within ‘blocks’ of genes that define barley types, as might have been anticipated in a predominantly selfing species where evolutionary histories between proximate loci are more correlated through limited recombination (Moore & Stevens, 2008; see more on recombination below). Overall, our data are consistent with an ancient division between subspecies of landrace and wild barley, followed by some limited secondary contact within current locations.

SNPs identify significant chromosome-level differences in genetic diversity and shed light on domestication processes

To investigate human and natural selection processes in landrace and wild barley with the BOPA1 platform, we calculated rolling genetic diversity estimates of mapped SNPs along chromosomes (Fig. 5). Differences in rolling diversity values by chromosome position, chromosome and barley category were evident. Overall, diversity estimates varied less along chromosomes for wild material than for landrace barley. At least in part, this may reflect different levels of LD between barley categories (lower levels in wild barley would lead to a more even distribution of diversity), a hypothesis that was tested through the calculation of chromosome-level ZnS values (Table 2). These indicated that LD is highly significant in both barley categories, although absolute values were sometimes relatively low (this apparent paradox reflects the relatively large numbers of loci sampled and included in pairwise comparisons). Only in the case of wild barley chromosome 1H was the estimate not significant. Corresponding to a more even distribution of diversity along chromosomes in wild barley, and in accordance with other comparative studies (Morrell et al., 2005; Caldwell et al., 2006; Waugh et al., 2009), our data revealed lower LD in wild than in landrace accessions.

The level of LD is known to be affected by population structure, the varying evolutionary histories of different gene pools, mating systems and the strength of selection, among other factors (Morrell et al., 2005; Caldwell et al., 2006; Kim et al., 2007; Mitchell-Olds et al., 2007; Stinchcombe & Hoekstra, 2008; Song et al., 2009). In our analysis, the difference observed between barley categories may in part represent the greater regional genetic structure observed in landraces (Fig. 3); however, for both categories of material, ZnS estimates without the specific chromosomal assignment of SNPs were lower than the mean values with chromosomal assignment (Table 2). This indicates that a proportion of the LD observed in our study is genuinely a result of physical linkage on chromosomes and not just down to geographic structuring. Another contribution to the observed difference between categories may be the different sampling strategies employed (more ‘diffuse’ individual sample points for wild accessions, compared with 24 population collection sites for landrace material), and further research on this point is required.

Regardless of origin, our observations on LD differences are important because both sets of barley used in the current study are key resources for crop improvement through association-based gene mapping (Waugh et al., 2009). The level of LD is a critical factor in determining appropriate strategies for the use of advanced cultivars, landraces and wild materials in modern breeding approaches, especially in the elimination of ‘false positive’ associations between markers and traits.
(Aranzana et al., 2005; Zhao et al., 2007; Stinchcombe & Hoekstra, 2008; Waugh et al., 2009). Our data show compensation for genetic structure to be of particular relevance in landrace barley. The difference in LD observed here is of particular interest because the same spatial range for landrace and wild germplasm categories was deliberately employed in order to remove a confounding macrogeographic component that is generally present in such comparisons.

Diversity profiles (Fig. 5) indicated that, in particular chromosomal regions, both landrace and wild barley exhibited reductions in genetic variation that may represent population bottlenecks and selection processes. A comparison of landrace and wild barley demonstrated that, on occasions, contrasting levels of diversity were revealed (Fig. 5). In total, 141 of 1059 mapped SNPs (13%) showed a significant difference ($P \leq 0.05$) in diversity between

Fig. 4 Chromosome-level profiles of genetic differentiation ($F_{ST}$) between landrace ($N = 317$) and wild ($N = 131$) categories of barley collected from the same geographic regions in Jordan and Syria, based on 1007 polymorphic, mapped, single nucleotide polymorphisms (SNPs). Arrows represent the locations of 21 category-diagnostic SNPs with a frequency difference ($D_{lw}$) between landrace and wild accessions of $\geq 0.9$ for alternate character states. $F_{ST}$ values for SNPs adjacent to diagnostic markers were generally at the ‘background’ level, and there is therefore no evidence that landrace–wild differentiation for SNPs extends to adjacent chromosomal regions (see further data in Table 2). SNPs are given in map order (starting with the lowest cM value at the bottom for each chromosome; positions of first and last SNPs given), but not on a cM scale (i.e. pairs of adjacent points on the vertical axis of profiles can be different distances apart in cM). The length of each chromosome in the figure therefore simply represents the number of polymorphic SNPs scored (see Materials and Methods for further information).
Fig. 5 Chromosome-level profiles of Nei’s (1978) genetic diversity (H) for landrace (blue line, N = 317) and wild (red line, N = 131) categories of barley collected from the same geographic regions in Jordan and Syria, based on 1059 mapped single nucleotide polymorphism (SNPs); 1007 of these SNPs were polymorphic and 52 monomorphic. Diversity estimates, which are rolling averages based on 15 adjacent loci (see Materials and Methods for further information), indicated differences within barley categories in the levels of variation revealed along chromosomes. Differences between landrace and wild materials were also evident; red and blue arrows represent positions at which rolling t-tests of H estimates indicated these differences to be statistically significant (red arrow, wild significantly more diverse than landrace accessions; blue arrow, vice versa). Black arrows represent the locations of 20 diagnostic SNPs also shown in Fig. 4 (one diagnostic SNP not shown as trimmed from map when using rolling averages). Overall, chromosomes in wild accessions were more diverse than in landrace material, except in the case of 2H (mean diversity for all markers across this chromosome approximately equal for categories, H = 0.209 and H = 0.208, respectively; see values given in Table 2). SNPs are given in map order (starting with the lowest cM value at the bottom for each chromosome; positions of first and last SNPs given), but not on a cM scale (i.e. pairs of adjacent points on the vertical axis of profiles can be different distances apart in cM). The length of each chromosome in the figure therefore represents the number of SNPs for which rolling diversity values could be calculated (as rolling averages were based on 15 loci, the first and last seven SNPs on each chromosome were excluded from the profiles). The approximate chromosomal locations of a subset of genes known to be important in barley evolution and adaptation are indicated, based on the following references (notations taken from quoted papers, see references therein for further information): a, Wenzl et al. (2006): ari-eGP (dwarfing), Eam1 (early maturity, long days), Ppd-H1 (photoperiod response), Sd (or denso, dwarfing); b, Pourkheirandish & Komatsuda (2007): btr1 and btr2 (nonbrittle rachis), nud (naked caryopsis), sgh1 (or Vrn-H2), Sgh2 (or Vrn-H1) and Sgh3 (or Vrn-H3) (all reduced vernalization), Ppd-H2 (photoperiod response), vr1 (six-rowed spike). Differences in the levels of variation revealed around these locations are discussed in the text.
categories. In 132 of these cases, wild barley was more diverse than landrace material, whereas in only nine cases was this pattern reversed, corresponding to expectations based on whole genome and whole chromosome analysis of diversity (estimates in Tables 1 and 2). Data showing that, for the majority of significant differences, it is the wild material that is more diverse are doubly relevant in the context of possible ascertainment bias which would be expected to lower diversity in the wild category (see discussion above). On occasions, lower landrace compared with wild diversity corresponded with known centromeric regions, for example, on chromosomes 1H and 4H (centromeres at 60 and 70 cM, respectively, according to Wenzl et al., 2006).

Of most interest is a comparison of SNP diversity patterns along chromosomes with the locations of mapped genes known to be important in barley evolution and adaptation. Pourkheirandish & Komatsuda (2007), for example, described some of the key traits associated with the transition from wild to cultivated barley that have been mapped with more or less accuracy. Included are a nonbrittle rachis, six-rowed spike, naked caryopsis, reduced vernalization and photoperiod insensitivity (Fig. 5). On occasions, the genes involved have been sequenced and the mechanism of the wild to domestic transition has been explored in some detail (e.g. Komatsuda et al., 2007; Jones et al., 2008).

Probably the most important feature connected to domestication is a nonbrittle rachis, a recessive character which results in grains remaining longer on plants after maturation, allowing the efficient harvest of cultivated compared with wild barley. This character is associated with two tightly linked genes, btr1 and btr2, on the short arm of chromosome 3H. Our analysis indicated a significant reduction in diversity in landrace barley compared with wild germplasm in this region, perhaps indicating a domestication bottleneck. Our data are consistent with a decrease in variation at this location revealed by BOPA1 SNPs in a previous study of cultivated barley (Rostoks et al., 2006), in this instance conducted on elite northwestern European types.

The appearance of a six- rather than two-rowed spike during domestication is associated with a number of genes, including recessive vrs1 on chromosome 2H. The wild-type Vrs1 allele is known to encode a transcription factor that includes a homeodomain with a closely linked leucine zipper motif, and a number of independent mutations in vrs1 have occurred in the origin of six-rowed cultivars (Komatsuda et al., 2007). Our data did not indicate significant differences in diversity between landrace and wild categories in the region around this gene, which is not surprising given that the landrace accessions assessed here were two-rowed. Nor did our data reveal significant differences around the single recessive gene, nud, located on chromosome 7H, which controls the naked caryopsis character. This trait, which allows the easy separation of husks on threshing, is found in certain barley cultivars and apparently represents a mutation that occurred early in the domestication process (at least 8000 BP; Pourkheirandish & Komatsuda, 2007).

Conversely, our data provided evidence for higher diversity in wild barley in the chromosomal regions around the genes vgh1 (or Vrn-H2), Sgh2 (Vrn-H1) and Sgh3 (Vrn-H3) (on chromosomes 4H, 5H and 7H, respectively), especially in the first and third cases, where the difference from landrace barley was statistically significant. These genes are associated with a reduced vernalization requirement in spring barley that developed at some stage during domestication. As another example in our summary of key domestication genes, photoperiod insensitivity in cultivated types has involved the accumulation of mutations in the Ppd-H1 and Ppd-H2 loci on chromosomes 2H and 1H, respectively (Pourkheirandish & Komatsuda, 2007). BOPA1 data indicated significantly higher diversity in wild barley around Ppd-H2, but not around Ppd-H1 (unlike Jones et al., 2008, who found greater diversity at the Ppd-H1 gene in wild material; possibly the lack of difference in our study is a case of ascertainment bias, see third paragraph of the section ‘SNPs reveal higher overall genetic diversity in wild than in landrace barley’).

Our data revealed additional significant differences in diversity between landrace and wild accessions at chromosomal regions known to be important in the development of the barley crop, such as around the Sd (dwarfing) gene on chromosome 3H, and in other parts of the genome. A comparison of rolling H values with the positions of our 21 category-diagnostic markers also provided interesting insights (Fig. 5). Diagnostic markers are, by definition, associated with relatively low diversity levels within barley categories because they approach fixation. However, consistent with the rapid ‘fall off’ of FST values observed around diagnostic markers (Fig. 4), rolling H values around the same SNPs did not reveal lower diversity than elsewhere. Nor does there appear to be any particular association between diagnostic markers and regions of the genome considered to be important in domestication. This supports a hypothesis that polymorphism at diagnostic markers is the result of stochastic sorting processes in the evolution of the barley crop, rather than an association with domestication traits directly. Further analysis is required to assess differences in genetic diversity between advanced cultivars, landraces and wild barley at mapped chromosome positions, in different geographic locations and under varying environmental conditions. This will identify new candidate chromosomal regions for responses to human and environmental selection processes, and will temporally dissect domestication traits, facilitate gene discovery, enhance future breeding programs and effectively guide conservation efforts (Rostoks et al., 2006; Burger et al., 2008; Verhoeven et al., 2008).
SNPs suggest limited geographic origins for cultivated barley

The number of origins involved in crop domestication is an important concern in agriculture because patterns of crop evolution guide future improvement strategies (Burger et al., 2008; Clement et al., 2010; Li et al., 2010). Some authors have stressed a single origin for cereals in the Fertile Crescent (Salamini et al., 2004), whereas others have placed emphasis on a more complex series of domestication events occurring within the distributions of wild progenitors (Brown et al., 2008). Where the ranges of present-day landraces and wild progenitors overlap, the most effective means to study the origins of cultivation is through geographically matched sampling of both gene pools. In the case of a predominantly local origin, landraces should be more genetically similar to geographically proximate wild stands than to other germplasm, although this will not be the case if landraces are derived from elsewhere. Despite the theoretical utility of this approach for the study of origins, it has predominantly been applied to incipient plant domesticates only (e.g. Kelly et al., 2004; Hollingsworth et al., 2005; Dawson et al., 2008; although see, for example, the exception of soybean: Li et al., 2010), partly because landraces and wild progenitors of long cultivated species frequently no longer co-occur. Even when distributions are sympatric, however, there has been a failure to embrace a matched sampling approach. The case of barley is a good example: a number of genetic studies have compared wild and cultivated stands in order to delineate origin (e.g. Badr et al., 2000; Wei et al., 2005; Kilian et al., 2006; Saisho & Purugganan, 2007), but nonmatched distributions of samples have confounded interpretation, as the geographic component of variation between categories has not been eliminated. Only one significant region-wide comparison of site-matched wild and landrace barley has been reported, but this study was limited in resolution and in genome coverage (see second paragraph of the section ‘SNPs reveal clear differentiation between landrace and wild barley and indicate secondary contact between the two’, $R_s = 0.707$, $P < 0.001$). Cultivated populations therefore appear to have been ‘in situ’ for a considerable period of time, although stands are clearly composed of mixed genotypes (as is evident from the spread of individual points within regions in ordination, Fig. 3) and present-day farmers are known to replace their seed by transfer over quite large distances (over 100 km; E. Weltzien, discussions during field sampling).

Our observation in the current study of geographically correlated SNP variation in both landrace and wild samples of barley provides the opportunity to explore origins of cultivation through a matched sampling method. A between-category comparison of genetic distances (see Table S2) indicated that landrace barley from all three geographic regions tested was most similar to wild barley from a single area, namely Jordan–southern Syria (Region 3), indicating that, of the material assessed here, this is the most likely origin of extant primitive cultivars in both countries. In other words, within two regions – northwestern and northeastern Syria (Regions 1 and 2) – geographically matched landrace and wild material did not demonstrate the lowest pairwise distance of all possible landrace and wild comparisons, as would have been anticipated if cultivated material was of local origin to these regions.

Clearly, our data provide information on likelihoods of origin only for the tested countries and do not exclude domestication events outside them (e.g. Morrell & Clegg, 2007); further research on origins will need to extend the use of the BOPA1 assay to geographically matched landrace and wild accessions collected from throughout the Fertile Crescent (also including, for example, Iran, Iraq, Israel, Lebanon, Turkey) in order to provide more definitive information. In addition, sample sizes across regions varied considerably for wild accessions in our comparison, and the inclusion of further wild individuals, especially from northeastern Syria (Region 1, $N = 14$ only in our study), is advised in order to exclude artifacts in analysis in which pairwise $F_{ST}$ values may be artificially inflated when sample sizes are small.

Although our data therefore provide support for the importance of cross-regional germplasm transfer events within the Fertile Crescent during barley domestication, the genetic structure observed among landrace populations shows a remarkable correspondence with the geographic locations of sample sites throughout the whole of Jordan and Syria (see second paragraph of the section ‘SNPs reveal clear differentiation between landrace and wild barley and indicate secondary contact between the two’, $R_s = 0.707$, $P < 0.001$). Cultivated populations therefore appear to have been ‘in situ’ for a considerable period of time, although stands are clearly composed of mixed genotypes (as is evident from the spread of individual points within regions in ordination, Fig. 3) and present-day farmers are known to replace their seed by transfer over quite large distances (over 100 km; E. Weltzien, discussions during field sampling).

The maintenance of population genetic structure in the face of human seed exchange, which could potentially have swamped pre-existing landraces, suggests that local adaptation and human selection for plant performance, under varying rainfall and soil type, have occurred across Jordan and Syria (Ceccarelli et al., 1987; van Leur et al., 1989; Weltzien, 1989). Such observations support the utility of landraces as sources of adaptive variation to combat the drought and heat stress associated with anthropogenic climate change, important concerns that are a focus of attention in current breeding programs (Weltzien & Fischbeck, 1990; Brown et al., 2008; von Korff et al., 2008).

Final remarks

Given the marker technologies now developed for barley, the availability of extensively sampled wild and landrace gene pools, and considerable comparative phenotypic data collected from controlled field trials conducted over many sites and seasons (Waugh et al., 2009), ongoing studies on the species continue to present a useful model for the investigation of the evolution, adaptation and spread of the...
world’s most important crops (Wenzl et al., 2006; Close et al., 2009). Combining population genomic data from landrace and natural populations with information from field trials, for example, provides a powerful approach to identify the genes responsible for adaptive phenotypes (Stinchcombe & Hookstra, 2008; Waugh et al., 2009). The beneficial effects of exotic genes from wild and landrace barley in advanced cultivars have been demonstrated, although significant challenges remain in the identification and introgression of favorable exotic genes – for biotic and abiotic stresses, quality traits and agronomic characteristics – into breeding programs (von Korff et al., 2006). These challenges can be aided by the use of SNPs in marker-assisted selection (Jones et al., 2009) and through new approaches to analysis, such as those described in this article.

Acknowledgements

We thank Joy K. Roy, Shiaoman Chao and Timothy Close for providing genotype data of the World Barley Diversity Collection, Stephanie Dahl for technical assistance, and Jordi Comadran and Bill Thomas for informative discussions. We are also grateful for the comments of three anonymous reviewers that improved the manuscript. This work was supported by grants BSBR (ERA-PG) ExBarDiv BB/E024726/1 and Generation Challenge Programme (GCP) G3007.02. SCRI received a grant in aid from the Scottish Government.

References


**Supporting Information**

Additional supporting information may be found in the online version of this article.

**Fig. S1** Genome-level frequency differences ($\Delta f$) between alternate character states for 1135 single nucleotide polymorphism (SNP) loci in landrace ($N = 317$) and wild ($N = 131$) categories of barley sampled from the same geographic regions in Jordan and Syria.

**Table S1** Passport and barley oligonucleotide pool assay 1 (BOPA1) single nucleotide polymorphism (SNP) data at 1135 loci for 317 landrace and 131 wild accessions of barley collected from the same three geographic regions in Jordan and Syria.

**Table S2** Pairwise genetic distance ($F_{ST}$) values for genome-level analysis of 1135 single nucleotide polymorphism (SNP) loci scored in landrace and wild barley accessions collected from the same three geographic regions in Jordan and Syria.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.