Molecular Genotyping of Historical Barley Landraces Reveals Novel Candidate Regions for Local Adaption

Aslan Selçuk, Nils Forsberg, Jenny Hagenblad and Matti W. Leino

Linköping University Post Print

N.B.: When citing this work, cite the original article.

Original Publication:

Aslan Selçuk, Nils Forsberg, Jenny Hagenblad and Matti W. Leino, Molecular Genotyping of Historical Barley Landraces Reveals Novel Candidate Regions for Local Adaption, 2015, Crop science, (55), 6, 2766-2776.

<http://dx.doi.org/10.2135/cropsci2015.02.0119>

Copyright: Crop Science Society of America

<https://www.crops.org/>

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-122553>

Molecular genotyping of historical barley (*Hordeum vulgare* L.) landraces reveals novel candidate regions for local adaption

Selçuk Aslan, Nils E G Forsberg, Jenny Hagenblad and Matti W Leino*

S. Aslan, Swedish University of Agricultural Sciences, Department of Plant Biology, Box 7080, SE-750 07 Uppsala, Sweden; N.E.G. Forsberg, Norwegian University of Science and Technology, Department of Biology, N-7491 Trondheim, Norway; N. E. G. Forsberg, J. Hagenblad and M. W. Leino, IFM-Biology, Linköping University, SE-581 83 Linköping, Sweden; M. W. Leino, Swedish Museum of Cultural History, SE-643 98 Julita, Sweden. S. Aslan and N. E. G. Forsberg made equal contributions to the article and are joint first authors.

* Corresponding author: (Matti.Leino@nordiskamuseet.se).

Abstract

Previous research has shown that barley landraces from Northern Europe form genetically distinct latitudinal groups, suggesting that adaption plays an important role in shaping the geographical distribution of genetic diversity. Here, we investigate how Northern European barley landraces relate to landraces from other parts of Europe and whether candidate genes and genetic regions for climate adaption can be identified. For this purpose 27 barley landraces, available as 120-140 year old seed collection specimens, were genotyped with a 384 SNP semi-high-throughput assay. Bayesian genetic clustering and principal component analyses showed that barley landraces from the Nordic countries form a homogenous and genetically distinct group relative to barley from Central Europe and the Mediterranean area. To search for patterns of adaption, polymorphic positions in the flowering time determining genes *HvCO1*, *HvFT1*, *Ppd-H1* and *VRN1-H1* were genotyped through Sanger sequencing. Genotyping of *Ppd-H1* confirmed the previously known allele distribution with the responsive allele present in the South and the non-responsive allele in the North. The other three genes were in general much more variable in Central and Southern Europe compared to the North. Neither of the flowering time genes showed any geographically correlated variation within the Nordic countries. To identify additional putative genetic regions associated with climatic adaption in Northern European barley landraces we correlated allelic frequencies in the 384 SNP set with climatic variables using the software *Bayenv*. This allowed us to identify five SNPs putatively associated with length of growth season and two SNPs putatively associated with precipitation. The results demonstrate how historical and ancient crop plant DNA can be used to understand how genetic variation has been geographically distributed during historical spread of crops and the molecular identity of adaption.

Keywords: aged DNA, flowering time, historical samples, *Hordeum vulgare*, landraces, SNP genotyping, *VRN-H1*, *HvFT1*, *Ppd-H1*, *HvCO1*, *Vrs1*.

Introduction

Barley (*Hordeum vulgare* L.) is cultivated in different regions throughout the world, many times in climates that are very different from that found in the species' native area. Within Europe only, length of growth season, photoperiod, temperature and precipitation all vary distinctly. To enable the crop's widespread cultivation, substantial adaption has been required. Such wide adaptability suggests that barley should have the genetic resource basis for meeting future climate changes, but a deeper understanding of the genes involved in climate adaption is needed for efficient breeding. Recently, large-scale genome-wide association studies (GWAS) using genebank genetic resources have been initiated (see Muñoz-Amatriaín et al., 2014; Pasam et al., 2014 and references therein) for this purpose. Barley landraces from some regions, such as Northern Europe, are, however, poorly represented in genebanks (Jones et al., 2008a; Pasam et al., 2014) and the adaption to these specific climates might go undetected. We show here how historical seed samples can be exploited to fill this gap.

The geographical distribution of neutral genetic diversity in barley has been the focus of a number of studies. Clear geographic clustering of European landrace barley has been shown, both on a continental scale (Jones et al., 2011, Pasam et al., 2014) and in certain regions, such as Spain (Yahiaoui et al., 2008), Sardinia (Rodriguez et al., 2012) and Fennoscandia (Leino & Hagenblad, 2010; Forsberg et al., 2015). This suggests limited gene-flow which would facilitate the evolution of local adaption in response to climatic differences. In Fennoscandia, clear latitudinal structuring shared across several countries can be observed (Forsberg et al., 2015). This clustering is in stark contrast with the historically documented longitudinal movement of seeds, and suggests that adaption to agro-climatic conditions may indeed play an important role in shaping the distribution of genetic diversity.

One of the most studied genes responsible for environmental adaption in barley is the photoperiod response gene *Ppd-H1*, which triggers flowering when barley is exposed to

lengthening days. A single base-pair mutation of the *Ppd-H1* gene results in a non-responsive allele (*ppd-H1*) delaying flowering time under long-day conditions (Turner et al., 2005; Jones et al., 2008b). The distribution of the two alleles has been studied in both extant and historical barley landraces from Europe (Jones et al., 2008b; Lister et al., 2009; Jones et al., 2011; Jones et al., 2012; Lister et al., 2013) with a clear separation between Southern Europe, where responsive *Ppd-H1* dominates and Northern Europe where the non-responsive *ppd-H1* is abundant. It has been suggested that the responsive allele would be advantageous in the South where drought requires rapid transfer into reproductive stage of the plant whereas in the North cooler and wetter conditions allow for a longer vegetative stage and thus higher yield potential (Jones et al., 2012).

Although *PPD-H1* is a major regulator of flowering time in barley, other genes in the pathway leading to a shift from vegetative to reproductive stage have also been described (see reviews by Cockram et al., 2007a; Trevaskis et al., 2007). Key genes, besides *Ppd-H1*, are the CONSTANS orthologue *HvCO1* (Campoli et al., 2012), the Floral Locus T gene *HvFT1* (Faure et al., 2007) and the vernalization gene *VRN-H1* (von Zitzewitz et al., 2005). Stracke et al., (2009) partially sequenced coding and non-coding intervals of three of these genes, i.e. *Ppd-H1*, *HvCO1* and *HvFT1*, in a collection of 220 spring barley accessions. The *Ppd-H1* sequence was highly variable and haplotypes were strongly associated with flowering time. *HvCO1* and *HvFT1* were both less variable and had weaker associations with flowering time, but a significant influence on flowering time was found when studying the interaction between *HvCO1* and *Ppd-H1* haplotypes. Stracke et al. (2009) further found that the haplotype frequencies for the three genes differed between continents but did not investigate the haplotype distribution within Europe. Studying sequence variation in *VRN-H1* in European barley, Cockram et al., (2007b) found that *VRN-H1* variation explained (together with *VRN-H2*) the spring/winter growth habit, but also that sequence variation might influence flowering

time among spring varieties. These *VRN-H1* alleles, associated with early flowering, are however mostly found in wild germplasm (Cockram et al., 2011).

Genebank collections of barley landraces from Northern Europe are very limited (Jones et al., 2008a; Pasam et al., 2014). In addition, several studies have shown that accessions from genebanks might have lost some genetic integrity and that historical material provides clearer phylogeographic pictures (Lister et al., 2009; Hagenblad et al., 2012; Forsberg et al., 2015). We thus chose to use seeds sampled in the 19th century, unaffected by 20th century plant breeding, seed trade or *ex situ* preservation as a source of DNA for this study. By using seed samples pre-dating the origin of modern breeding practices we can be confident there has been no introgression of modern varieties into our landrace material, as has been detected in other studies of landrace material (Bellucci et al., 2014). The seeds were sampled directly on farms in the late 19th century, usually in large quantities (thousands of seeds), and have never been regenerated. The seed samples can thus be considered to be representative 19th century landrace *in situ* accessions. Although the seeds are not viable (Leino et al., 2009), making traditional phenotype studies impossible, they are, for most parts of the region, the only available landrace material for studying adaption in the Nordic countries.

The latitudinal genetic structuring previously found in historical Fennoscandian barley (Forsberg et al., 2015) suggests that climate adaption has contributed to genetic structuring. In this study, we first contrasted the genetic diversity of the Nordic material to that present in historical material of landrace barley from more southern latitudes to assess the genetic relationship between Fennoscandian barley and barley from other parts of Europe. Second, we genotyped polymorphic positions in some known flowering regulatory genes to search for signs of adaption. Third, we used SNP genotyping data to find additional genetic targets for climate adaption.

Materials and methods

Plant material and DNA-extraction

Seeds from 27 historical accessions of landrace barley were selected for this study (Table 1). The accessions were chosen to represent a geographic span from the northernmost edge of Sweden to the Upper Egypt region in the south. Of the chosen barleys, 24 were taken from the 19th century seed collections of the Swedish Museum of Cultural History (NM; Leino et al., 2009) and the remaining three were selected from herbarium specimens collected by the Swedish physician and botanist Albert Atterberg (ATT) during the late 19th century (Atterberg, 1899; Leino, 2011). The ATT material was available as single ears from three individuals of each accession and the NM material was available as several thousand seeds from a mixture of individuals.

DNA extraction was performed in an area isolated from other DNA work using FastDNA Spin Kits and the FastPrep Instrument (MP Biochemicals, Solon, OH, USA), with necessary precautions taken for handling of historical material (Leino et al., 2009). Negative controls were included in each of the extraction series.

Genotyping with SNP markers

Six individual seeds from each accession were genotyped with the markers in the C-384 SNP set developed by Moragues et al. (2010). SNP data were generated using an Illumina Golden Gate assay (Illumina Inc., San Diego, CA, USA). The resulting data were processed and studied with the Bead Studio 3.1.3.0 software package (Illumina Inc., San Diego, CA, USA).

Genotyping polymorphic positions in flowering related genes

Single seed DNA extracts were genotyped for flowering related genes. A set of PCR primers was designed to amplify short fragments in the genes *HvCO1*, *HvFT1*, *VRN-H1* and *Ppd-H1* (Supplementary Table S1). The degraded nature of historical DNA only allows for short amplicons, and hence only parts of the genes could be readily sequenced. When primer pairs repeatedly failed to amplify DNA, redesign of primers was carried out at least twice. The best working primer pairs are presented in Supplementary Table S1. The regions sequenced were chosen to cover previously identified polymorphic positions putatively associated with functionality (Table 2). PCRs were performed in a total reaction volume of 20 µl containing 2 mM MgCl₂, 0,2 µM forward primer, 0,2 µM reverse primer, 0,4 µM each of dNTPs and 1 U Dream Taq DNA polymerase and supplied buffer (Fermentas). The PCR cycling parameters were: denaturation for 2 min 30 s at 94 °C followed by 30 cycles of 30 s of denaturation at 94 °C, 1 min annealing at 50-59 °C, 30 s of extension at 72 °C and final elongation step of 72 °C for 5 min. PCRs were carried out in two stages with 2 µl from the first reaction being used as template in the second reaction. Fragments covering *Vrs1* mutations, segregating six-row barley from two-row barley, were amplified according to Leino & Hagenblad (2010).

Amplicons were either directly Sanger sequenced using the PCR primers (Macrogen, South Korea) or in the case of *Ppd-H1* genotyped with a CAPS (Cleaved Amplified Polymorphic Sequence) assay. The *Ppd-H1* amplicons were restricted with FastDigest *Hha1* enzyme (Fermentas) according to the manufacturer's instructions and electrophoresed. If the SNP_2432 is a T (non-responsive allele) the result is a 120 bp fragment whereas a C (responsive allele) results in restriction into two fragments of 47 bp and 73 bp.

Statistical analysis

The spatial genetic structure of the accessions was studied with Principal Component Analysis (PCA) performed using the R statistical software (R Development Core Team, 2013, version 3.02). Spatial genetic structure was also investigated using the software Structure 2.3.3 (Pritchard et al., 2000) with the admixture model and 25 000 burn-in iterations followed by 50 000 MCMC iterations. The number of clusters, K, best describing the data was determined from the H' value calculated by CLUMPP (Jakobsson et al, 2007) and the ΔK value (Evanno, 2005). The results from 10 independent runs of STRUCTURE for each K were merged using CLUMPP and graphical output was obtained using the software ArcGIS Desktop software (ESRI, 2011).

Environmental correlations

Environmental correlations with SNP distribution were evaluated using the *Bayenv* 2.0 software (Günther and Coop, 2013). As a proxy for the historical climate crops have adapted to we used environmental data on daily temperatures and precipitation for the 30 year period of 1961-1990, extracted from the Global Land Surface Databank (http://gosc.org/GLOBAL_SURFACE_DATABANK/GBD.html). Average yearly precipitation was calculated for the same 30 year period, when available, and was used as a measure of regional precipitation. Proportion of days with lowest daily temperatures over 5°C was used as a proxy of growth season length. The NM662 accession, which was represented by a single individual, was excluded for the environmental correlation. *Bayenv* 2.0 was run for 50 000 iterations and the runs were repeated 20 times for each SNP averaging the results for a more consistent result, as suggested by Blair et al. (2014).

Results

Genotyping of historical seed

In spite of being significantly fragmented, the DNA of the historical samples was successfully genotyped for SNP markers using the Illumina GoldenGate 384 SNP assay. The reproducibility of the method was assessed through genotyping four individual seeds from the same ear of the ATT9 accession. After excluding missing data (3 SNPs) this resulted in completely identical genotypes for all individual seeds. In the total dataset, SNPs with more than 20 % missing data were removed from further analysis leaving 309 SNPs to be used in the downstream analyses. Individuals with more than 50 % missing data were excluded; leaving 143 individual seeds from 27 accessions with an average level of missing data of 3.54 %.

The PCR-amplification of fragments used for genotyping the flowering time polymorphisms through Sanger sequencing proved consistently less sensitive than the high-throughput SNP genotyping method. Successfully amplified fragments were between 90 and 200 bp, and no obvious correlation between fragment length and amplification success was observed. Overall, 89 % of the fragments covering flowering time polymorphisms were successfully amplified and sequenced. *Ppd-H1*, *HvFT1* SNPs 1572 and 1599 as well as *VRN-H1* SNPs 191 and 258 were successfully genotyped in all accessions (Supplementary Table S2).

Phylogeographic structure across Europe

The Bayesian clustering algorithm implemented in the software Structure was used to assess the data for phylogeographic structure and genetically group six individuals from each accession (three in the case of the ATT accessions, and one in the case of NM662). Clustering into two groups had the best support from both ΔK and H' values (Supplementary Table S3). The two groups had geographically well-defined distributions (Figure 1A). All Fennoscandian

accessions (except NM270 Piikis) and the German accession NM2152 Oderbrücker clustered primarily in the same cluster (yellow). A second cluster (grey) included all remaining accessions.

A clustering similar to that found by Structure could be observed along the first PC in a PCA of the accessions' genotypes (Supplementary figure S1A). Along the second PC the three accessions NM270 Piikis, ATT13 Siebenbürgen and ATT16 Tyrolen clustered separately from all other accessions. Row-type is one possible explanation for the separation along the second PC. We thus examined the row-type of the samples from morphology and through genotyping. Complete ears were only available for the four Atterberg (ATT) accessions, all of which were six-row. The remaining samples were available as threshed seed, but seed size and presence of a curved morphology can indicate row-type. This examination suggested NM270 Piikis to be two-row but all remaining accessions as six-row. Three mutations in the row-type determining gene *Vrs1* are known that result in a six-row phenotype (Komatsuda et al., 2007). Genotyping these mutations showed that most accessions were indeed six-row, carrying the causative deletion in position 1353 (Supplementary Table S2). However, the accession ATT16 (but not ATT13) was wild-type in all three known mutation sites and should thus be two-row in spite of its appearance. In summary, the clustering along PC2 (Supplementary figure S1A) seem associated with the two-row barleys, but in a cryptic manner. The Finnish NM270 seems phenotypically two-row but not genetically so, the Austrian ATT16 is phenotypically six-row, but has a genotype that should result in a two-row phenotype, whereas the Romanian ATT13 has both phenotypic and genetic evidence for a six-row phenotype.

When we analyzed only accessions from Sweden and Finland, we obtained similarly high support for two and three clusters respectively from H' values while ΔK supported two or four clusters (Supplementary Table S3). At $K = 3$ populations showed the latitudinal clustering suggested by previous studies (Figure 1B, Forsberg et al., 2015). At $K = 4$ a northern cluster

and a southern cluster were found with NM270 Piikis and NM625 Toppmyra assigned to individual clusters (data not shown). A similar picture was shown by a PCA of the accessions' genotypes where the distribution of the Fennoscandian accessions along the first two PCs showed a distribution similar to, and significantly correlated with, the latitudinal origin of the accessions (Supplementary figure S1B). PC1 was correlated with latitudinal origin ($p < 0.05$, $r^2 = 0.255$) and PC2 was correlated with both the latitudinal ($p < 0.001$, $r^2 = 0.552$) and longitudinal origin ($p < 0.05$, $r^2 = 0.451$) of the accessions.

Geographical distribution of flowering time alleles

The latitudinal structuring of the Fennoscandian accessions suggested a causative role of climate adaption. We thus genotyped single individuals from each accession for variable sequence positions in genes previously suggested to regulate flowering induction in barley, and thus putatively involved in the adaption of landraces.

No causative SNP has yet been identified in *HvCOI*, but several polymorphic positions putatively associated with flowering time have been suggested. We genotyped six of these polymorphisms, four in introns and two in exons (Table 2). Five of the six were monomorphic in the successfully genotyped accessions. It should be noted, however, that the rates of PCR failure were relatively high for genotyping these polymorphisms and some variation might have escaped detection. The sixth marker (HvCO1_885) was polymorphic, but without obvious geographical distribution (Figure 2A). In Fennoscandia all accessions except NM613 Sundby, had identical genotypes.

Five variable positions in the *HvFT1* promoter and intron sequence were genotyped (Table 2; Supplementary Table S2). One of the five (HvFT1_1572) was monomorphic, while the other

four displayed polymorphisms. In Fennoscandia all genotypes were identical with the exception of NM727 Sandön. In contrast, *HvFT1* proved much more variable in Central and Southern Europe (Figure 2B).

The vernalization gene *VRN-H1*, a gene that also acts as a flowering regulator, was analyzed for four positions (Table 2). The four sites were found to be rather variable, particularly in Central and Southern Europe (Figure 2C). Only VRN-H1_191 had a clear geographic differentiation where all Fennoscandian accessions, except NM270 Piikis, shared the same allele, an allele that was not found among the more southern accessions.

The causative SNP in position 2432 in the photoperiod response gene *Ppd-H1* was genotyped using a CAPS assay. All Fennoscandian accessions, without exception, carried the non-responsive allele (Figure 2D, Supplementary Table S2). Among the remaining accessions from Central and Southern Europe the responsive allele dominated, with exception of the accessions from Spain (NM2180) and Romania (ATT13). Notably, the relatively northern accession from Scotland (NM2191) also carried the responsive allele (Figure 2D).

Identification of additional variation associated with climate adaption

As the flowering time genes showed little or no variation across Fennoscandia, we screened the SNP dataset used for investigating population structure in the Fennoscandian countries for genetic variation linked to climate. The Illumina GoldenGate assay SNPs genotyped are derived from expressed sequence tags (ESTs) and thus correspond to expressed genes. The software *Bayenv* uses Bayesian statistics to rank genetic polymorphisms by their allele frequency correlation with environmental variables. As environmental variables, we used growth season length and mean yearly precipitation calculated from data from the meteorological stations closest to the sample origins.

We identified five SNPs correlated with length of growth season and two precipitation correlated SNPs (Figure 3; Supplementary Table S2), neither of which were fixed within the accessions studied. None of the identified SNPs correlated with both length of growth season and precipitation. The strongest correlation with growth season length was found for SNP 11_10030, positioned in a gene with unknown function, and neither of the SNPs correlated with growth season were located in well-characterized proteins or genes with known effect on flowering time or vernalization response (Figure 3A). The strongest precipitation correlated SNP was 11_10093 located in a gene corresponding to a drought-inducible protein (Close et al., 2009) (Figure 3B).

Discussion

As agriculture spread across Europe, cultivated crops were repeatedly facing new climatic conditions. In particular, this was the case when agriculture entered Scandinavia at the northernmost fringe of Europe, about 4000–3700 BC. Here, extreme conditions with temperature-limited growth seasons and particular light regimes required specific adaption of crop species. Recent archaeological evidence shows that the agricultural expansion into the North was rapid (Sørensen & Karg, 2012). The movement of crops thus likely caused repeated genetic bottlenecks with strong selective pressure on a few adapted genotypes. In this study, accessions from Northern Europe formed a more homogenous group than Central and Southern Europe, both for SNP markers (Supplementary figure 2A) and the flowering time genes. This result is in congruence with Jones et al. (2011), studying extant landraces with SSR markers, who also showed a relative homogeneity among Fennoscandian barley, compared to barley from other parts of Europe.

Although geography could explain genetic clustering to a large extent, the three deviating accessions NM270 Piikis, ATT13 Siebenbürgen and ATT16 Tyrolen suggested that also row-type has a major influence. Previous genotyping of European barley have indeed shown that two-rowed and six-rowed barley are genetically diverse groups (Yahiaoui et al., 2008; Malysheva-Otto et al., 2006; Kolodinska Brantestam et al 2007; Hagenblad & Leino, 2010). The analyses of phenotype and *Vrs1* mutations known to result in a six-row phenotype (Komatsuda et al., 2007) were, however, not fully concordant and indicate that row-type is under complex control of more genes than just *Vrs1*. Alternative mechanisms for control of row-type, besides *Vrs1*, have been suggested (Palmer et al., 2009; Koppolu et al., 2013; Cuesta-Marcos et al., 2010). For example, *Vrs4* was recently identified as a regulator of *Vrs1* expression, but has low natural variation and no causative polymorphisms have been identified yet (Koppolu et al., 2013). It is also possible that the aberrant accessions carry other allelic variants of *Vrs1*. Cuesta-Marcos et al. (2010) identified a *Vrs1.a4* allele associated with a six-row phenotype but without detecting a causative mutation.

In this study, we chose to work with historical materials assembled from 19th century seed collections. For the Nordic area, historical crop materials are superior to extant genebank accessions with regards to both geographical coverage and detailed passport data. The genotyping difficulties caused by DNA degradation are offset by an indisputable landrace origin with no risk of identity loss during the past century of plant improvement and more recent genebank conservation. As in Forsberg et al. (2015) we show that efficient and reliable SNP genotyping of historic plant DNA can be performed with a high-throughput method, such as the Illumina Golden Gate assay. Recently, Lister et al. (2013) validated another high-throughput SNP genotyping assay, Kompetitive Allele Specific PCR (KASP), for use on historic plant DNA. Both these methods seem to outperform traditional Sanger sequencing for SNP genotyping, in terms of success rate and efficiency. Alternatively, pyrosequencing could

offer another choice for cost-efficient SNP genotyping in barley (Silvar et al., 2011). These methods do, however, require development of assays and when the targets are specific positions in the genome, Sanger sequencing might still be an economically preferable choice.

Lister et al. (2009; 2013) have previously shown that the geographical distribution of *Ppd-H1* alleles is painted much clearer with 19th century herbarium material than with extant genebank landraces. The same clear pattern for *Ppd-H1* in Fennoscandia is obtained in this study, where all Fennoscandian barleys were found to have the non-responsive *ppd-H1* allele, contrasting extant Fennoscandian barley landraces with scattered presence of the responsive *Ppd-H1* allele (Jones et al., 2008b).

In addition to *Ppd-H1*, we also genotyped a set of polymorphic positions in the genes *HvCO1*, *HvFT1* and *VRN-H1* previously suggested to be associated with flowering time (Stracke et al., 2009; Cockram et al., 2007b). These genes also contrast clearly between Northern Europe on one hand and Central and Southern Europe on the other. In general, more variation was found among the accessions from Central and Southern Europe. Two positions, VRN-H1_191 and HvFT1_1801 (and HvFT1_1915 in strong LD with HvFT1_1801; Stracke et al. 2009), have geographic distributions similar to *Ppd-H1*, with one variant in the North and another in Central and Southern Europe. The lack of clear geographic patterns in the other flowering time SNPs need, however, not rule out their association with climatic adaption. Ma et al. (2010) studying populations of aspen (*Populus tremula*) showed a gradual change in allele frequencies of four flowering time associated SNPs over different latitudes rather than abrupt transitions between presence and absence of an allele. It would be interesting to genotype additional individuals from each location for the same SNPs that we have studied here to see if gradual allele frequency changes show a clearer geographic pattern than the distribution of the single individuals studied here. Whether the SNPs we have studied have any actual effect on gene

function and adaption, are linked to adaptive polymorphisms or are just artefacts of genetic drift is not clear and clarifying this will require extensive sequencing and phenotyping of extant barley cultivars.

The polymorphisms at HvFT1_1801 and HvFT1_1915 have been suggested to be diagnostic for winter/spring growth habit in barley (Yan et al., 2006). The same polymorphism was recorded by Faure et al. (2007). In Spanish landraces of winter barley the HvFT1_1801-1915 polymorphisms have been found to be geographically structured with the A-G haplotype (here shown to prevail in Fennoscandia) more common in the North and conferring delayed flowering. Recent reports show that expression of *HvFT1*, with influence on flowering time, is also regulated by copy number variation (Nitcher et al., 2013). Of particular interest is the finding that some Fennoscandian cultivars, derived from native landraces, have a unique genetic constitution with a single promoter and multiple copies of exon 1 and exon 3, conferring high *HvFT1* expression and rapid flowering (Loscos et al., 2014).

Based on the classification by Yan et al. (2006) all Fennoscandian landraces are fixed for the *HvFT1* spring allele, whereas the Central and Southern European landraces have the winter allele. In Northern Europe spring types are almost exclusively cultivated, yet in other parts of Europe both spring and winter (and facultative) types occur (Cockram et al., 2011; Jones et al., 2011). With the winter allele of *HvFT1* present, growth habit is instead regulated by *VRN-H1* and *VRN-H2* allele combinations (Cockram et al., 2007b; Cockram et al., 2011). *VRN-H1* diversity also consists of variation in the length of the first intron (Cockram et al. 2007b), while functionality of *VRN-H2* is determined by gene deletions (von Zitzewitz et al., 2005). Unfortunately, the degradation of DNA in the historical material complicates the interpretation of genotyping for deletions.

By comparing with genotyping performed in modern barley cultivars from Fennoscandia it is possible to investigate whether flowering time genotypes have changed during 20th century

breeding. Loscos et al. (2014) genotyped eleven Fennoscandian cultivars, all of which carried the non-responsive *ppd-H1* allele, similar to the historical material. In *HvFT1*, nine of the eleven cultivars (all six-row) had the HvFT1_1801-1915 A-G haplotype. The remaining two, with the T-C haplotype, were two-row barleys. In our historical material, all Fennoscandian accessions, including the two-row NM270 Piikis had the A-G haplotype. For *HvCO1* no data from Fennoscandian cultivars are available, but we note that our historical material, with exception of NM613 Sundby, correspond to haplotype 3, as defined by Stracke et al. (2009). This haplotype was present in 15% of Stracke's 100 European cultivars, where instead haplotype 1 dominated. Cockram et al. (2007b) included 82 Fennoscandian cultivars in their screening of *VRN-H1* diversity. Of the four tested polymorphic sites in our study, the same genotypes dominated in Cockram's screening. However, we note a greater diversity among the modern cultivars in comparison to the historical material, in particular for SNPs VRN-H1_191 and VRN-H1_16723. This is partly explained by a higher number of two-row cultivars in the modern cultivar screen, but more variation is also found among six-row cultivars. In conclusion, modern cultivars from Fennoscandia largely carry the same genotypes in flowering time genes as did their ancestral landraces. Some introgression of other alleles through breeding programs using exotic germplasm is suggested, partly as a consequence of the increased use of two-row barley in Fennoscandia during the 20th century.

Structure analysis of the Fennoscandian accessions used in this study (Figure 1B) confirmed the latitudinal clustering previously found in a similar material (Forsberg et al, 2015). Our PCA for the Fennoscandian data (Supplementary figure S1B) also verified the latitudinal ordering found by Leino and Hagenblad (2010) using a different type of genetic marker. In spite of the clear latitudinal clustering of different accessions within Fennoscandia, we found no obvious variation explaining adaption to the various climatic conditions in the region among the variable positions in flowering time genes investigated. This finding does not exclude that

additional variation exists within these known flowering time genes also within Fennoscandia. It is, however, likely that these genes are either without variation causing flowering time adaption or that variation within and among populations is very limited. LD has been shown to be strong both within these genes (Stracke et al. 2009) and in the barley genome in general in this type of material (Forsberg et al. 2015). Additionally the haplotypes found showed a high level of similarity. Instead, we tested if genetic regions associated with climate adaption could be found among the set of non-flowering time SNP markers. Although less than 400 SNP markers scattered over the huge barley give a poor resolution, LD in this plant material is extensive and could be found over genetic distances more than 100 cM (Forsberg et al., 2015). The selection that can be expected to have acted on climate adaption genes should further have increased LD. Thus, SNPs associated with climate variables should indicate candidate regions for climatic adaptations.

For this purpose, we used the software *Bayenv*. This program has previously been successfully applied on various organisms to identify loci correlated with climatic adaptations (examples, see Günther & Coop, 2013). The SNP most strongly associated with differences in climate is the SNP 11_10093. This SNP was correlated with variation in precipitation and is located in a gene corresponding to a drought-inducible protein (Close et al., 2009). Recently, Hu et al. (2013), using a transgenic approach, showed that the wheat orthologue of this gene confers tolerance to drought stress. Of the remaining six SNPs identified as associated with variation in length of growth season and precipitation neither is located in genes previously indicated as involved in climate adaption. This should not be surprising as we are only studying few SNPs spread over the barley genome. The large amount of LD previously shown in barley from this region means associations can be found with markers located tens of cM away from the actual SNP. The SNPs detected here should be considered as markers for candidate regions in the genome rather than de facto causative SNPs. Genotyping of a larger set of accessions with more densely

located markers in the candidate regions detected, preferably with the addition of phenotypic data, will be needed to shed light of the genes causing adaption to variations in the length of growth season and precipitation.

In the present study, we have shown how historical samples of locally adapted crop landraces, although no longer viable, can be used to identify genetic variation associated with local adaption. Genetic analyses of historical specimens thus constitute an alternative and complementary approach to the analysis of phenotypic data (Keilwagen et al., 2014). On a continental scale Northern European barley landraces are genetically homogenous. In spite of this, genetic markers for potential adaptive variation can be detected, also with a low genomic coverage of SNPs.

Acknowledgements

This research was funded by the Lagersberg foundation and the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS). Maria Lundström and Tytti Vanhala are acknowledged for technical assistance.

References

- Atterberg, A. 1899. Varieties and Forms of Barley. (In German.) J. f. Landwirtschaft 47:1-44.
- Bellucci, E., E. Bitocchi, D. Rau, L. Nanni, N. Ferradini, A. Giardini, M. Rodriguez, G. Attene and R. Papa. 2013. Population structure of barley landrace populations and gene-flow with modern varieties. PloS One, 8:e83891.

Blair, L. M., J.M. Granka, and M.W. Feldman. 2014. On the stability of the Bayenv method in assessing human SNP-environment associations. *Human genomics*, 8:1.

Campoli, C., B. Drosse, I. Searle, G. Coupland, and M. von Korff. 2012. Functional characterisation of *HvCO1*, the barley (*Hordeum vulgare*) flowering time ortholog of CONSTANS. *Plant J.* 69: 868-880.

Casas, A.M., A. Djemel, S. Yahiaoui, L.J. Ponce, B. Contreras-Moreira, M.P. Gracia, J.M. Lasa and E. Igartua (2011). HvFT1 (VrnH3) drives latitudinal adaption in Spanish barleys. *Theor Appl Genet* 122:1293-1304.

Cockram J., H. Jones, F.J. Leigh, D. O'Sullivan, W. Powell, D.A. Laurie, and A.J. Greenland. 2007a. Control of flowering time in temperate cereals: genes, domestication, and sustainable productivity. *J. Exp. Bot.* 58:1231-1244.

Cockram, J., E. Chiapparino, S.A. Taylor, K. Stamati, P. Donini, D.A. Laurie, and D.M. O'Sullivan. 2007b. Haplotype analysis of vernalization loci in European barley germplasm reveals novel *VRN-H1* alleles and a predominant winter *VRN-H1/VRN-H2* multi-locus haplotype. *Theor. Appl. Genet.* 115: 993–1001.

Cockram, J., H. Jones and D.M. O'Sullivan. 2011 Genetic variation at flowering time loci in wild and cultivated barley. *Plant Genet Resour* 9:264-267.

Close, T.J., P.R. Bhat, S. Lonardi, Y. Wu, N. Rostoks, L. Ramsay, A. Druka, N. Stein, J.T. Svensson, and S. Wanamaker. 2009. Development and implementation of high-throughput SNP genotyping in barley. *BMC Genomics* 10:582.

Cuesta-Marcos, A., P. Szűcs, T.J. Close, T. Filichkin, G.J. Muehlbauer, K.P. Smith, and P.M. Hayes. 2010. Genome-wide SNPs and re-sequencing of growth habit and inflorescence genes in barley: implications for association mapping in germplasm arrays varying in size and structure. *BMC Genomics* 11:707.

ESRI (2011). ArcGIS Desktop: Release 10. Environmental Systems Research Institute: Redlands, CA.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14:2611-2620.

Faure, S., J. Higgins, A. Turner, and D.A. Laurie. 2007. The *FLOWERING LOCUS T*-like gene family in barley (*Hordeum vulgare*). *Genetics* 176:599–609.

Forsberg, N.E.G., J. Russell, M. Macaulay, M.W. Leino, and J. Hagenblad. 2015. Farmers without borders - genetic structuring in century old barley (*Hordeum vulgare*). *Heredity* 114:195–206.

Günther, T., G. Coop. 2013. Robust identification of local adaption from allele frequencies. *Genetics* 195: 205-220.

Hagenblad J., J. Zie, and M.W. Leino. 2012. Exploring the population genetics of genebank and historical landrace varieties. *Genet. Res. Crop. Evol.* 59:1185-1199.

Hu, W., C. Huang, X. Deng, S. Zhou, L. Chen, Y. Li, C. Wang, Z. Ma, Q. Yuan, Y. Wang, R. Cai, X. Liang, G. Yang, and G. He. 2013. TaASR1, a transcription factor gene in wheat, confers drought stress tolerance in transgenic tobacco. *Plant Cell Environ.* 36:1449-1464.

Jakobsson, M., and N.A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801-1806.

Jones H., D.L. Lister, M.A. Bower, F.J. Leigh, L.M. Smith, and M.K. Jones. 2008a. Approaches and constraints of using existing landrace and extant plant material to understand agricultural spread in prehistory. *Plant Gen. Res.* 6:98–112.

Jones H., F.J. Leigh, I. Mackay, M.A. Bower, L.M. Smith, M.P. Charles, G. Jones, M.K. Jones, T.A. Brown, and W. Powell. 2008b. Population-based resequencing reveals that the flowering time adaption of cultivated barley originated east of the Fertile Crescent. *Mol. Biol. Evol.* 25:2211-2219.

Jones, H., P. Civián, J. Cockram, F.J. Leigh, L.M. Smith, M.K. Jones, M.P. Charles, J.-L. Molina-Cano, W. Powell, G. Jones, and T.A. Brown. 2011. Evolutionary history of barley cultivation in Europe revealed by genetic analysis of extant landraces. *BMC Evol. Biol.* 11:320.

Jones, G., H. Jones, M.P. Charles, M.K. Jones, S. Colledge, F.J. Leigh, D.A. Lister, L.M.J. Smith, W. Powell, and T.A. Brown. 2012. Phylogeographic analysis of barley DNA as evidence for the spread of Neolithic agriculture through Europe. *J. Archae. Sci.* 39:3230-3238.

Keilwagen, J., B. Kilian, H. Özkan, S. Babben, D. Perovic, K.F. Mayer, A. Walther, C. H. Poskar, F. Ordon, K. Eversole, A. Börner, M. Ganai, H. Knüpffer, A. Graner and S. Friedel. 2014. Separating the wheat from the chaff—a strategy to utilize plant genetic resources from ex situ genebanks. *Scientific reports*, 4:5231.

Kolodinska Brantestam, A., R. von Bothmer, C. Dayteg, I. Rashal, S. Tuveesson, and J. Weibull. 2007. Genetic diversity changes and relationships in spring barley (*Hordeum vulgare* L.) germplasm of Nordic and Baltic areas as shown by SSR markers. *Genet. Res. Crop Evol.* 54:749-758.

Komatsuda, T., M. Pourkheirandish, C. He, P. Azhaguvel, H. Kanamori, D. Perovic, N. Stein, A. Graner, T. Wicker, A. Tagiri, U. Lundqvist, T. Fujimura, M. Matsuoka, T. Matsumoto, and M. Yano. 2007. Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc. Natl. Acad. Sci. USA.* 104:1424-1429.

Koppolu, R., N. Anwar, S. Sakuma, A. Tagiri, U. Lundqvist, M. Pourkheirandish, T. Ruttend, C. Seilere, A. Himmelbach, R. Ariyadasaa, H. M. Youssef, N. Stein, N. Sreenivasulu, T. Komatsuda and T. Schnurbusch. 2013. Six-rowed spike4 (*Vrs4*) controls spikelet determinacy and row-type in barley. *Proc. Natl. Acad. Sci. USA.* 110:13198-13203.

Leino, M.W., J. Hagenblad, J. Edqvist J, and E.-M. Karlsson Strese. 2009. DNA preservation and utility of a historic seed collection. *Seed Sci. Res.* 19:125-135.

Leino, M.W., and J. Hagenblad. 2010. Nineteenth century seeds reveal the population genetics of landrace barley (*Hordeum vulgare*). *Mol. Biol. Evol.* 27:964–973.

Leino, M.W. 2011. Rediscovery of a historic oat collection. (In Swedish.) *Sv. Bot. Tidsk.* 105:27-30.

Lister D.L., M.A. Bower, C.J. Howe, and M.K. Jones. 2009. Latitudinal variation in a photoperiod response gene in European barley: insight into the dynamics of agricultural spread from ‘historic’ specimens. *J. Archae. Sci.* 36:1092-1098.

Lister, D. L., H. Jones, M.K. Jones, D.M. O'Sullivan, and J. Cockram. 2013. Analysis of DNA polymorphism in ancient barley herbarium material: Validation of the KASP SNP genotyping platform. *Taxon* 62:779-789.

Loscos, J., E. Igartua, B. Contreras-Moreira, M.P Gracia and A.M. Casas. 2014. HvFT1 polymorphism and effect—survey of barley germplasm and expression analysis. *Front Plant Sci* 5:251.

Ma, X.F., D. Hall, K.R. St Onge, S. Jansson, and P.K. Ingvarsson. 2010. Genetic differentiation, clinal variation and phenotypic associations with growth cessation across the *Populus tremula* photoperiodic pathway. *Genetics* 186:1033-1044.

Malysheva-Otto, L.V., M.W. Ganai, and M.S. Röder. 2006. Analysis of molecular diversity, population structure and linkage disequilibrium in a worldwide survey of cultivated barley germplasm (*Hordeum vulgare* L.) *BMC Genet.* 7:6.

Muñoz-Amatriaín, M., A. Cuesta-Marcos, J.B. Endelman, J. Comadran, J. M. Bonman, H. E. Bockelman, S. Chao, J. Russell, R. Waugh, P. M. Hayes, and G. J. Muehlbauer. 2014. The USDA barley core collection: genetic diversity, population structure, and potential for genome-wide association studies. *PloS One* 9: e94688.

Nitcher, R., A. Distelfeld, C. Tan, L. Yan and J. Dubcovsky. 2013. Increased copy number at the HvFT1 locus is associated with accelerated flowering time in barley. *Mol. Genet. Genom.* 288: 261-275.

Palmer, S.A., J.D. Moore, A.J. Clapham, P. Rose, and R.G. Allaby. 2009. Archaeogenetic evidence of ancient Nubian barley evolution from six to two-row indicates local adaption. *PLoS One*, 4: e6301.

Pasam, R.K., R. Sharma, A. Walther, H. Özkan, A. Graner, and B. Kilian. 2014. Genetic diversity and population structure in a legacy collection of spring barley landraces adapted to a wider range of climates. *PLoS One* 9: e116164.

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.

R Development Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria.

Rodriguez, M., D. Rau, D. O'Sullivan, A.H. Brown, R. Papa, and G. Attene. 2012. Genetic structure and linkage disequilibrium in landrace populations of barley in Sardinia. *Theor. Appl. Genet.* 125:171-184.

Silvar, C., D. Perovic, A. M. Casas, E. Igartua and F. Ordon. 2011. Development of a cost-effective pyrosequencing approach for SNP genotyping in barley. *Plant breed.* 130: 394-397.

Stracke S., G. Haseneyer, J.B. Veyrieras, H.H. Gieger, S. Sauer, A. Graner, and H.P. Phipo. 2009. Association mapping reveals gene action and interactions in the determination of flowering time in barley. *Theor. Appl. Genet.* 118:259–273.

Sørensen, L., and S. Karg. 2012. The expansion of agrarian societies towards the north—new evidence for agriculture during the Mesolithic/Neolithic transition in Southern Scandinavia. *J. Archae. Sci.* 51:98-114.

Turner A., J. Beales, S. Faure, R.P. Dunford, and D.A. Laurie. 2005. The pseudo-response regulator *Ppd-H1* provides adaption to photoperiod in barley. *Science* 11:1031-1034.

Trevaskis B., M.N. Hemming, E.S. Dennis, and W.J. Peacock. 2007. The molecular basis of vernalization-induced flowering in cereals. *Trends Plant Sci.* 12:353–357.

Yan L., D. Fu, C. Li, A. Blechl, G. Tranquilli, M. Bonafede M, A. Sanchez, M. Valarik, and J. Dubcovsky. 2006. The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proc. Natl. Acad. Sci. USA.* 103:19581-19586.

von Zitzewitz J., P. Szucs, J. Dubcovsky, L. Yan, E. Francia, N. Pecchioni, A. Casas, T.H.H. Chen, P.M. Hayes, and J.S. Skinner. 2005. Molecular and structural characterization of barley vernalization genes. *Plant Mol. Biol.* 59:449–467.

Yahiaoui, S., E. Igartua, M. Moralejo, L. Ramsay, J.L. Molina-Cano, F.J. Ciudad, J.M. Lasa, M.P. Gracia, and A.M. Casas. 2008. Patterns of genetic and eco-geographical diversity in Spanish barleys. *Theor. Appl. Genet.* 116:271–282.

Table 1: Details of the historic barley accessions analysed in this study. Accession name is given as written on the accession labels. Latitude and longitude are given at exact position for some cultivars, while for others only intervals were available. Country of origin of the samples and their harvest years are also indicated.

Accession	Accession name	Latitude	Longitude	Country of origin	Harvest year
NM264	Mattila	61°04'42''N	25°37'44''E	Finland	1882
NM270	Piikis	60°25'29''N	22°30'58''E	Finland	1882
NM599	Matarengi	66°23'13''N	23°39'15''E	Sweden	1896
NM613	Sundby	64°17'39''N	21°13'22''E	Sweden	1896
NM625	Toppmyra	59°44'51''N	17°33'11''E	Sweden	1896
NM633	Pajala	67°12'45''N	23°22'0''E	Sweden	1896
NM646	Ramvik	62°49'9''N	17°51'23''E	Sweden	1896
NM662	Kengis	67°10'60''N	23°30'0''E	Sweden	1896
NM667	Nya Skottorp	56°27'2''N	13°0'22''E	Sweden	1896
NM668	Kurrokveik	66°3'5''N	17°53'11''E	Sweden	1896
NM669	Vuollerim	66°25'44''N	20°37'23''E	Sweden	1896
NM671	Hylkebo	56°35'17''N	15°51'34''E	Sweden	1896
NM705	Brattbäcka	64°14'36''N	15°52'7''E	Sweden	1896
NM727	Sandön	65°32'50''N	22°24'2''E	Sweden	1896
NM777	Assmundstorp	57°46'51''N	11°55'51''E	Sweden	1896
NM789	Wuono	65°51'10''N	23°9'18''E	Sweden	1896
NM2152	Oderbrücker	54°-47°N	14°-6°E	Germany	1877
NM2180	Spanien	43°-36°N	7°-3°W	Spain	1877
NM2181	Grekland	41°-37°N	26°-20°E	Greece	1877
NM2182	Turkiet	42°-35°N	44°-29°E	Turkey	1877
NM2191	Skottskt	58°-54°N	5°-1°W	Scotland	1877
NM2199	Tripolis	32°52'34''N	13°11'15''E	Libya	1877
NM2218	Frankrike	46°13'39''N	2°12'49''E	France	1877
ATT9	Södra Schweiz	46°N	9°E	Switzerland	1898
ATT13	Siebenbürgen	47°9'N	24°30'E	Romania	1898
ATT16	Tyrolen	4°16'N	11°24'E	Austria	1898
NM1585	Öfre Egypten	24°-27°N	31°-32°E	Egypt	1877

Table 2. Details of the polymorphisms analyzed.

Name in this study	Gene	Reference	Name in reference	Accession number	Position in accession sequence
HvCO1_506	<i>HvCO1</i>	Stracke et al., 2009	ps506	AF490467	506
HvCO1_582	<i>HvCO1</i>	Stracke et al., 2009	ps582	AF490467	582
HvCO1_843	<i>HvCO1</i>	Stracke et al., 2009	ps843	AF490467	843
HvCO1_885	<i>HvCO1</i>	Stracke et al., 2009	ps885	AF490467	885
HvCO1_912	<i>HvCO1</i>	Stracke et al., 2009	ps912	AF490467	912
HvCO1_1816	<i>HvCO1</i>	Stracke et al., 2009	ps1816	AF490467	1816
HvFT1_1162	<i>HvFT1</i>	Stracke et al., 2009	ps1162	DQ100327	1162
		Yan et al., 2006	167		
HvFT1_1572	<i>HvFT1</i>	Stracke et al., 2009	ps1572	DQ100327	1572
HvFT1_1599	<i>HvFT1</i>	Stracke et al., 2009	ps1599	DQ100327	1599
HvFT1_1801	<i>HvFT1</i>	Yan et al., 2006	270		1801
		Stracke et al., 2009	ps1805	DQ100327	
HvFT1_1915	<i>HvFT1</i>	Yan et al., 2006	384		1915
		Stracke et al., 2009	ps1919	DQ100327	
VRN-H1_191	<i>HvVRN1</i>	Cockram et al., 2007b	SNP1	AY750993	191
VRN-H1_258	<i>HvVRN1</i>	Cockram et al., 2007b	SNP2	AY750993	258
VRN-H1_16705	<i>HvVRN1</i>	Cockram et al., 2007b	SNP4	AY750993	16705
VRN-H1_16723	<i>HvVRN1</i>	Cockram et al., 2007b	SNP5	AY750993	16723
Ppd-H1_2432	<i>Ppd-H1</i>	Turner et al., 2005	SNP 15	AY970701	2432
		Jones et al., 2008	SNP 48		
Vrs1_915	<i>Vrs1</i>	Komatsuda et al., 2007	A40>F.S.	AB489124	915
Vrs1_1021	<i>Vrs1</i>	Komatsuda et al., 2007	F75>L	AB489124	1021
Vrs1_1353	<i>Vrs1</i>	Komatsuda et al., 2007	E152>F.S.	AB489124	1353

Figures

Figure 1. Map showing Structure clustering A) All accessions at $K = 2$. B) Fennoscandian accessions at $K = 3$.

Figure 2. Geographic distribution of alleles at flowering time genes with missing data indicated in white: A) *HvCO1_885*, B) *HvFT1*, segments indicate specific polymorphic positions on the gene, as shown in the legend, C) *VRN-H1*, segments indicate specific polymorphic positions on the gene, as shown in the legend, D) *Ppd-H1_2432*. The responsive allele is shown in blue and the non-responsive allele is shown in yellow. In A) and D) each circle indicates the genotype of a given accession and in B) and C) each circle is divided in four segments each showing the genotype of one of four polymorphic positions.

Figure 3. Results of *Bayenv* analysis for SNPs associated with: A) Length of growth season and B) Precipitation.

Supplementary materials

Supplementary figure S1. PCA of genotype frequencies of the accessions studied. A) All accessions with Fennoscandian accessions in black, other accessions in grey. B) Fennoscandian accessions only.

Supplementary table S1. Details of the PCR primers used in this study, annealing temperature for each primer sets, and expected fragments sizes are shown. References for published primers or Genbank accession numbers in case new primers were designed are indicated.

Supplementary table S2. Putative SNPs for local adaption.

Supplementary table S3. Support values for K in STRUCTURE analyses.

Supplementary Table S1. Details of the PCR primers used in this study, annealing 1

Primer Name	Sequence	Reference
HvCO_e1_1F	gctgctgctcagcaaagatt	
HvCO_e1_1R	tgcagatgttgttcttgc	
HvCO_i1_1F	ctgctgggctagtgcttac	
HvCO_i1_1R	catatggtcttccacctgaagt	
HvCO_i1_2F	cgttgaactaaattaacccatct	
HvCO_i1_2R	taccgcttccattgagaaa	
HvFT_prom_indel_F	atggacatggaacctgccact	Yan et al 2006
HvFT_prom_indel_R	tggtgatgatgagtgttgccc	Yan et al 2006
HvFT_intron_F2	tgagatgaggaccttctacacg	
HvFT_intron_R2	aagcatcgaaaaggacgta	
HvFT_intron_F	ccattttctgtgctctctgg	
HvFT_intron_R	cctgcaggcagtataaagca	
HvVRN1_SNP1_2_R	aaaagctaagcgcatagaaaatg	
HvVRN1_SNP1_2_L	ttggtttgcattcttgaa	
HvVRN1_SNP4_5_R	cacaaccaagcgaacacat	
HvVRN1_SNP4_5_L	ttggtgtcccttcaatcaa	
PpdF1	aggcagcagctagagccata	
PpdR1	tcgatctcgaccttccagg	
vrs1a1_F1	tgcacatctgcagttcctct	Leino & Hagenblad 2010
vrs1a1_R1	gtgtccgccatccacagat	Leino & Hagenblad 2010
vrs1a2_a3_F1	gggagaccagcaagcagag	Leino & Hagenblad 2010
vrs1a2_a3_R1	ggcggccagatacacctt	Leino & Hagenblad 2010

temperature for each primer sets, and expected fragments sizes are shown. References for published

Newly designed from sequence	Polymorphic position(s)	Ann. Temp	Fragment Length
based on AF490467	HvCO1 SNP 506 and 582	55°	183 bp
based on AF490467			
based on AF490467	HvCO1 SNP 843, 885 and 912	53°	147 bp
based on AF490467			
based on AF490467	HvCO1 SNP 1816	56°	131 bp
based on AF490467			
	HvFT1 indel 1162	59°	90 bp
based on DQ100327	HvFT1 SNP 1572 and 1599	61°	150 bp
based on DQ100327			
based on DQ100327	HvFT1 SNP 1801 and 1915	56°	206 bp
based on DQ100327			
based on AY750993	VRN-H1 SNP 191 and 258	50°	167 bp
based on AY750993			
based on AY750993	VRN-H1 SNP 16705 and 16723	55°	182 bp
based on AY750993			
based on AY970701	Ppd-H1 SNP 2432	58°	120 bp
based on AY970701			
	Vrs1 SNP 1353	55°	167 bp
	Vrs1 SNP 915 and 1021	56°	200 bp

ed primers or Genbank accession numbers in case new primers were designed are indicated

Supplementary table S2. Putative SNPs for local adaptation. Variable SNPs in flowering time

accession number	Label	Country	HvCO1_506	HvCO1_582	HvCO1_843
NM264	Mattila	Finland	C	C	AT
NM270	Piikis	Finland	NA	NA	AT
NM599	Matarengi	Sweden	C	C	AT
NM613	Sundby	Sweden	C	C	AT
NM625	Toppmyra	Sweden	C	C	AT
NM633	Pajala	Sweden	NA	C	AT
NM646	Ramvik	Sweden	C	NA	AT
NM662	Kengis	Sweden	C	C	AT
NM667	Nya Skottorp	Sweden	C	C	AT
NM668	Kurrokveik	Sweden	C	C	AT
NM669	Vuollerim	Sweden	C	C	AT
NM671	Hylkebo	Sweden	C	C	AT
NM705	Brattbäcken	Sweden	C	C	AT
NM727	Sandön	Sweden	NA	NA	AT
NM777	Assmundstorp	Sweden	C	C	AT
NM789	Wuono	Sweden	NA	NA	AT
NM2152	Oderbrücker	Germany	NA	NA	AT
NM2180	Spanien	Spain	C	C	AT
NM2181	Grekland	Greece	NA	NA	AT
NM2182	Turkiet	Turkey	C	C	AT
NM2191	Skottskt	Scotland	C	C	AT
NM2199	Tripolis	Libya	NA	NA	AT
NM2218	Frankrike	France	NA	NA	AT
ATT9	Södra Schweiz	Schweiz	NA	NA	AT
ATT13	Siebenbürgen	Romania	NA	NA	AT
ATT16	Tyrolen	Austria	NA	NA	NA
NM1585	Öfre Egypten	Egypt	NA	NA	NA

genes are shown in bold.

SNPs in Putative Flowering Time C						
HvCO_885	HvCO_912	HvCO1_1816	HvFT1_1162	HvFT1_1572	HvFT1_1599	HvFT1_1801
del	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
A	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
del	C	C	no_del	T	C	A
del	C	C	no_del	T	T	A
del	C	C	no_del	T	T	T
del	C	C	del	T	C	T
A	C	C	no_del	T	C	T
del	C	C	no_del	T	T	T
A	C	C	no_del	T	C	T
del	C	C	del	T	C	T
A	C	NA	NA	T	C	A
del	C	NA	NA	T	C	T
NA	NA	C	NA	T	C	NA
NA	NA	NA	NA	T	C	T

genes

HvFT1_191	VRN-H1_191	VRN-H1_258	VRN-H1_16705	VRN-H1_16723
G	C	G	C	G
NA	T	A	C	G
G	C	G	NA	NA
G	C	G	NA	NA
G	C	G	C	G
G	C	G	C	G
G	C	G	T	G
G	C	G	C	G
G	C	G	T	G
G	C	G	C	G
G	C	A	T	G
G	C	G	C	G
G	C	G	T	G
C	C	G	C	G
G	C	G	C	G
G	C	G	C	G
G	C	G	C	G
C	T	G	C	A
C	T	G	C	A
C	T	A	T	G
C	T	A	T	G
C	T	G	C	A
C	T	G	NA	NA
G	T	A	C	G
G	T	A	T	G
NA	T	A	T	G
NA	T	G	C	A

Growth season associated SNPs

Ppd_2432	11_10030*	11_20449**	11_20379***	11_10023***	11_10814***
T	0,6	1	0,833333333	1	1
T	1	0,5	0,5	0	1
T	0,833333333	1	0	1	1
T	0	1	0	0,833333333	1
T	0	0	1	0	0,166666667
T	0,833333333	1	0	1	1
T	0,833333333	0,5	0,166666667	0,833333333	0,833333333
T	1	1	0	1	1
T	0	0,2	1	0,6	0,4
T	1	1	0,166666667	1	1
T	0,666666667	1	0	1	1
T	0	0	0,833333333	0,833333333	0,166666667
T	0,666666667	0,833333333	0	1	1
T	0,833333333	0,833333333	0	1	1
T	0	0	1	0	0,166666667
T	0,5	0,4	0,166666667	1	1
T	0	0,166666667	0,666666667	0,166666667	0,333333333
T	0,333333333	0,5	0,666666667	0,166666667	1
C	0	0,5	1	0,333333333	1
C	0,333333333	1	1	0,5	0,833333333
C	0,5	1	0,666666667	0,333333333	1
C	0,25	1	0,333333333	0	0,6
C	1	0,5	0,2	0,4	0,8
C	0	1	0,666666667	0,666666667	0,333333333
T	0	0,666666667	1	1	0,666666667
C	0	1	1	0	1
C	0	0	0	0,666666667	1

* A/G polymorphism, ratio of G

** A/T polymorphism, ratio of T

*** C/G polymorphism, ratio of G

Precipitation associated SNPs			Row type SNP	
11_10093*	11_20415*	Vrs1_915	Vrs1_1021	Vrs1_1353
0,333333333		1 del	NA	NA
1		0 del	NA	del
0		1 del	C	NA
0		1 del	C	del
0		1 del	C	del
0		1 del	C	del
0,166666667	0,666666667	del	C	del
0		1 del	C	del
1	0,2	del	C	del
0		1 del	C	del
0		1 del	C	del
0,833333333	0,166666667	del	C	del
0		1 del	C	del
0		1 del	C	del
1		0 del	C	del
0		1 del	C	del
0,666666667	0,333333333	del	C	del
0	0,833333333	del	C	del
0,333333333		1 del	C	del
0	0,5	del	NA	no_del
0	0,333333333	del	C	del
0,5		1 del	C	NA
0		1 del	C	del
0,666666667	0,333333333	NA	NA	no_del
1		0 del	C	del
0		1 del	C	no_del
0		1 del	C	del

ans phenotypes

[illegible]

Supplementary table S3. Support values for K in STRUCTURE analyses.

Clustering, Northern Europe

K	lnl	deltaK	H'
1	-5977,23	NA	NA
2	-5155	97,30341	0,9952
3	-4777,91	6,930749	0,9958
4	-4499,16	60,9436	0,9168
5	-6656,55	0,729853	0,7656
6	-5590,61	0,091937	-
7	-4202,35	7,600109	-
8	-4354,99	0,435937	-
9	-4355,92	0,718245	-
10	-4550,89	NA	-

Clustering, Full set

K	lnl	deltaK	H'
1	-19569,2	NA	NA
2	-14533	494,6345	0,9987
3	-13680,7	0,21799	0,8062
4	-12840,8	1,7776	0,7231
5	-12202,2	0,382436	0,7093
6	-11637,5	0,501233	-
7	-11117,2	29,25501	-
8	-13441	0,169314	-
9	-14804,7	0,077006	-
10	-15485,1	NA	-