*Article*

# Genetic Diversity in 19th Century Barley (*Hordeum vulgare*) Reflects Differing Agricultural Practices and Seed Trade in Jämtland, Sweden

**Martin N. A. Larsson** [1] , **Matti W. Leino** [2,*,†] **and Jenny Hagenblad** [1,†]

[1] Department of Physics, Chemistry and Biology, Linköping University, SE-581 83 Linköping, Sweden; martin.la90@gmail.com (M.N.A.L.); jenny.hagenblad@liu.se (J.H.)

[2] The Archaeological Research Laboratory, Department of Archaeology and Classical Studies, Stockholm University, SE-106 91 Stockholm, Sweden

* Correspondence: matti.leino@arklab.su.se

† These authors contributed equally.

**Abstract:** Landrace crops are important genetic resources, both for plant breeding efforts and for studying agrarian history. The distribution of genetic diversity among landraces can reflect effects of climate, economic structure, and trade also over a limited spatial and temporal scale. In this study, we have SNP genotyped historical barley seed samples from the late 19th century, together with extant barley landrace accessions from Jämtland, Sweden, a county centrally located, situated between Sweden and Norway. We found two main genetic clusters, one associated with the main agricultural district around lake Storsjön and one in the peripheral areas. Data was also compared with genotypes from landraces from across the Scandinavian peninsula. Accessions from the peripheral part of Jämtland show genetic similarity to accessions from a large part of central Scandinavia, while the accessions from the Storsjön district are more differentiated. We suggest that these dissimilarities in genetic diversity distribution are explained by differences in the relative importance of agriculture and trading. We further compared the historical material with ex situ preserved extant landraces from the same region and found that their genetic diversity was not always representative of the given provenience. The historical material, in contrast, proved particularly valuable for assessing how crop genetic diversity has historically been influenced by economic focus.

**Keywords:** landraces; genetic diversity; barley; *Hordeum vulgare*; population structure; seed exchange networks; gene bank conservation; single nucleotide polymorphism (SNP); historical DNA; agrarian history

## 1. Introduction

The concept of crop landraces dates back to the late 19th century [1,2], although landraces have existed for far longer than this. Landraces are populations of cultivated plants that are reproduced by farmers' repeated use of part of the harvest as seed for subsequent sowing [3,4]. Through natural and unintentional artificial selection, but with little or no formal plant improvement (but see Berg [5]), landraces become adapted to both local growth conditions and the farmers' management regime. These differentiating forces, together with the stochastic effects of genetic drift and homogenization through gene-flow caused by, for example, seed exchange and seed replacement, result in plant populations that are, in spite of being more or less genetically diverse, often distinguishable from each other (e.g., [6–8]). The co-development of landrace crops and agriculture allows questions concerning crop evolution, agricultural history and farmers' interactions to be addressed by investigating the distribution of genetic diversity within and among landraces (e.g., [6,9–11]).

Although the value of landraces as a source for plant breeding was discussed already at the turn of the 19th century, few efforts were made to collect, categorize, and catalogue

landraces, as reviewed by [1]. By the time the first gene bank in the Nordic countries (Denmark, Sweden, Norway, Finland, and Iceland) was founded in 1979, most landraces in this area had already been replaced by modern cultivars. What few extant landraces have been preserved may have lost their genetic integrity since the time they were in active cultivation [12–14]. Fortunately, several historical collections exist, containing 19th century seed of landrace cereals from the Scandinavian peninsula (Norway and Sweden) [15–17]. Molecular genetic studies have shown that DNA can be extracted from the grains in these collections [15] and that the samples can be used to infer the role of natural selection for local adaptation [18] as well as the evolutionary consequences of historical events such as trade and crop failure [14,19].

Barley (*Hordeum vulgare* L.) was domesticated 10,000 years ago [20] and is today a major cereal crop. The Food and Agricultural Organization of the United Nations (FAO) estimates that in 2018 142.9 million tons of barley were produced worldwide, accounting for roughly 5.4% of global cereal production [21,22]. Around 70% of barley production is used as feed for animals with the remainder primarily being used in malting for beer production [23]. Being a hardy crop, barley can be grown in areas where climate and soil are less suited for agriculture and where other cereal species cannot thrive [24]. The wide geographical distribution of barley as a crop has resulted in numerous landraces, harboring substantial genetic diversity and adapted to diverse climates [25,26]. Barley landraces are therefore important resources in breeding efforts to meet future challenges of climatic stress [27,28].

On the Scandinavian peninsula, barley has been grown for more than 6000 years [29] and is still a major crop, especially in the northern part [30]. In Scandinavia, climate, light and soil conditions vary, driving adaptation to different growth conditions in barley from different parts of the region [6,18] and differences in the appearance and traits of many landraces [31]. The Scandes mountain range between Norway and Sweden has been a natural barrier for both trade and agriculture. Where the local climate has permitted, agriculture has, however, been practiced in inland Sweden and trade routes are known to have crossed the mountain range [32]. An area of particular interest is the county of Jämtland, situated between the 62nd and 64th parallel and encompassing part of the Scandes. The county is centrally located between Norway, neighbouring Trøndelag county, and Sweden. Historically, the county has successively belonged to either Norway or Sweden, last coming under Swedish rule in 1645. Cultural influences from Trøndelag in the west have existed alongside trade relations with Swedish counties south and east of Jämtland as well as with Norway [33].

Although trade has been a dominating economic force in Jämtland in historic times, agriculture has also been practiced [34]. The oldest finds of barley cultivation are from the Migration period ~400 AD with cultivation expanding substantially during the Viking age ~1000 AD [35]. During the historical period, as well as today, barley has been the major cereal crop in the area [32,36,37]. Due to the short growth season in the region, not all years resulted in a ripe harvest and during the 19th century crop failures are known to have occurred in 1800, 1808, 1821, 1830–1833 and 1867 [32,38]. Such crop failure events have been shown to result in change in the genetic composition of cultivated crops due to an influx of replacement seed [19]. To what extent replacement crop in Jämtland was sourced from Trøndelag, neighbouring Swedish counties, or locally, is not known.

In this study we have investigated population structure in historical landrace barley from Jämtland. The seed samples were collected in 1896, before widespread use of modern cultivars [39], and can therefore be trusted to be bona fide landraces. We have genetically compared the historical seed samples with ex situ preserved landraces from the same region and the neighbouring Trøndelag county, as well as with historical and extant landrace material from other parts of the Scandinavian peninsula. In doing so we could examine the genetic composition of landraces in this isolated area at the climate margin for cereal cultivation. The data also allowed us to analyze to what extent extant landraces from the

region are good representatives of the barley that was cultivated in historical times. Finally, we investigated whether seed exchange occurred, within and from outside the county.

## 2. Materials and Methods

### 2.1. Study Material

Sixteen accessions of barley originating from the county Jämtland or neighbouring Trøndelag were genotyped for the study (Table 1). The word "accession" will henceforth be used to refer to seed collected at a specific site on a specific occasion. Six individuals per accession were genotyped in twelve historical accessions from the seed collection housed at the Swedish Museum of Cultural History (prefix NM) [15]. These historical accessions were harvest samples from 1896 originally collected for the 1897 Art and Industry fair in Stockholm. All accessions from this fair originating from Jämtland were included in the study. Coordinates for the origin of these samples were obtained from the exhibition catalogue [40], combined with the search tool on the website of the Swedish Local Heritage Federation [41]. In addition, five individuals per accession were genotyped for four extant accessions obtained from the Nordic Genetic Resource Center (Nordgen, prefix NGB). The extant accessions were chosen based on having good provenience information and originating from the region of interest (Table 1).

**Table 1.** Original collection site, country, collection year, coordinates, and within-accession genetic diversity ($H_s$) of the studied accessions.

| Accession | Farm/Village | Country | Collected | Lat | Long | $H_s$ |
|---|---|---|---|---|---|---|
| NGB13670 | Oppdal | NOR | NA | 62.5943 | 9.6912 | 0.000 |
| NGB15157 | Jormvattnet | SWE | 1905 | 64.7189 | 14.0468 | 0.000 |
| NGB13394 | Snåsa | NOR | 1879 | 64.2465 | 12.3822 | 0.000 |
| NM.0406128 | Gevåg, Ragunda | SWE | 1896 | 63.1498 | 16.3310 | 0.074 |
| NM.0406276 | Tullingsås, Ström | SWE | 1896 | 63.8214 | 15.5062 | 0.074 |
| NM.0406277 | Risselås, Ström | SWE | 1896 | 63.8493 | 15.6085 | 0.065 |
| NM.0406291 | Rösta, Täng | SWE | 1896 | 63.2407 | 14.5651 | 0.086 |
| NM.0406122 | Backen, Brunflo | SWE | 1896 | 63.0855 | 14.8181 | 0.054 |
| NM.0406120 | Lungre, Kyrkås | SWE | 1896 | 63.2261 | 14.8659 | 0.067 |
| NM.0406124 | Österåsen, Häggenås | SWE | 1896 | 63.4035 | 14.9474 | 0.066 |
| NM.0406241 | Bjerte, Nälden | SWE | 1896 | 63.3056 | 14.3360 | 0.056 |
| NM.0406247 | Böle, Lit | SWE | 1896 | 63.2841 | 14.9096 | 0.059 |
| NM.0406248 | Silje, Rödön | SWE | 1896 | 63.2411 | 14.4327 | 0.046 |
| NM.0406251 | Kvarnsved, Frösön | SWE | 1896 | 63.1795 | 14.5867 | 0.049 |
| NM.0406254 | Landsom, Täng | SWE | 1896 | 63.2680 | 14.5571 | 0.057 |
| NGB6927 | Jämtland | SWE | NA | NA | NA | 0.044 |

### 2.2. Genetic Analysis

DNA was extracted using either the FastPrep kit (MP Biomedicals, Solon, OH, USA, historical accessions) or the DNeasy Plant Mini kit by Qiagen (Hilden, Germany, extant accessions), according to the manufacturers' instructions. To minimize the risk of contaminating DNA, extractions from historical samples were carried out in a sterile bench where work with extant DNA had never been performed.

Genotyping was carried out by LGC genomics using the KASP method [42]. The degraded nature of the DNA in the historical specimens limits the number of markers that can be genotyped from any single grain and in total 100 SNPs were genotyped. Of these, 91 SNPs were chosen from the 384 C-SNP set developed for European landraces [43], based on genetic diversity found in previous studies [6,44]. Additionally, nine SNPs with associated gene functions were also genotyped (Table 2).

To situate the data produced in this study in a Scandinavian context, already available genotyping data for Swedish and Norwegian landrace barley for the same SNPs [6,19] were re-analyzed. In addition, unpublished data from two already genotyped Norwegian accessions were added, the extant accessions NGB456 ('Dønnes') and NGB2079 ('Skjåk')

and the accession "Sjønstå1" consisting of two grains found during the restoration of the roof to an old building at Sjønstå farm, Fauske, Nordland (Hanne Prytz, pers. com.). In total, this comparative genotyping data comprised of nine extant and 41 historical accessions. For accessions where data from more than six individuals were available, the number of individuals was randomly pruned to six.

**Table 2.** SNPs with associated gene functions genotyped in present study.

| Name | Function | Reference |
| --- | --- | --- |
| *ppd-H1_* SNP48 | Photoperiod response, causative | [45] |
| *Int-c_*SNP124 | Spike morphology, modifier | [46] |
| *HvCEN* SNP9 | Spring or winter growth habit/earliness | [47] |
| *HvFT1* i1_270 | Flowering time | [48] |
| *HvFT1* prom_167 | Flowering time | [48] |
| *HvP5CS1* SNP735 | Drought resistance | [49] |
| *HvP5CS1* SNP846 | Drought resistance | [49] |
| *Vrs1* E152 > F.S. | Spike morphology, causative | [50] |
| *Vrs1* A40 > F.S. | Spike morphology, causative | [50] |

*2.3. Data Analysis*

Initial analysis, data cleaning, principal component analysis (PCA), preparation of data for STRUCTURE, mantel tests and multiple regression analysis were all done in R [51] using RStudio v1.2.1335 [52]. The packages *tidyverse* v1.3.0 [53], *adegenet* v2.1.2 [54], *hierfstat* v0.04-22 [55], *ggfortify* v0.4.10 [56], *openxlsx* v4.1.5 [57], *genepop* v1.1.7 [58] and *ggrepel* v0.8.2 [59] were used. SNPs with a success rate below 80% (SNPs 11_20897 and 11_11059) were discarded before further analysis. After discarding these SNPs, success rate per individual was 93% or higher, and all individuals were kept for further analysis. Expected heterozygosity under Hardy-Weinberg equilibrium, $H_s$ [60], was used as a measure of within-accession diversity and calculated using the *Hs* function in *adegenet*. To further assess genetic diversity Pairwise $F_{ST}$ was calculated according to Nei [60] using the *pairwise.fst* function in *hierfstat*. To find clustering patterns, principal component analysis (PCA) was run using the *prcomp* function, the dataset was treated as haploid (with heterozygotes treated as missing data) and only polymorphic SNPs (34 of 98) were included in the PCA. The frequency of each allele at each locus was used for both individual and accession level PCA.

In addition, to assess clustering, Structure v2.3.4 was run [61–64], with heterozygotes treated as missing data. In each run, the length of burn-in period was set to 20,000 repetitions and number of MCMC repetitions after burn-in to 50,000. Ten simulations were run for each of K = 1 to 16 using the admixture model. To assess the most appropriate number of clusters, to compile separate runs and to visualize the results CLUMPAK [65] was run using the main pipeline with the LargeKGreedy algorithm using random input order and 2000 repeats, as well as the best K function. Average per-accession cluster-identity was extracted from the CLUMPAK pop files and displayed using ArcMap v10.6.0.8321 [66] to visualize geographical clustering.

To test for isolation by distance, a mantel test was run comparing the genetic and geographic Euclidian distances using the *mantel.randtest* function. Genetic distance was calculated using Edwards' distance in the function *dist.genpop*. Geographic distance was calculated with the function *dist* using coordinates. Both the full set of accessions from Jämtland and Trøndelag and a reduced dataset with only the historical accessions were analyzed using 9999 repetitions each.

To investigate whether latitude, longitude, altitude, or length of growth season were associated with the clustering seen in the historical data, these parameters were run as independent variables in four separate multiple regression models using the *lm* function. The dependent variables in the four different models were 1: proportion of accession belonging to cluster one in Structure analysis of only historical accessions from Jämtland,

extracted from the CLUMPP output at K = 2 (Figure S1); 2: PC1 from PCA of only historical accessions (Figure S4); 3: PC2 from PCA of only historical accessions (Figure S4); and 4: $H_s$ (Table 1). Both Structure analysis and PCA were run as described above. Altitude was extracted from the "GSD-Altitude-data, grid 50+ NH" dataset from the Swedish mapping, cadastral and land registration authority [67]. Growth season was extracted from a map based on the Swedish Meteorological and Hydrological Institutes open data set which covers 1961–1990 [68].

## 3. Results

### 3.1. Genotyping of Historical and Extant Accessions

We genotyped historical and extant accessions of barley from the counties Jämtland in Sweden and Trøndelag in Norway resulting in a total of 100 SNP genotypes for 92 individuals from 16 accessions. Mean SNP success rate and mean individual success rate were both found to be 98%, with all individuals having a success rate of more than 92%. Out of 98 successfully genotyped SNPs, 34 were found to be polymorphic.

Within-accession genetic diversity ($H_s$) ranged from zero, for the extant accessions NGB13394, NGB13670 and NGB15157, to 0.086 for the historical accession NM.0406291 (Table 1). The extant accession from an unknown farm in Jämtland (NGB6927) had an $H_s$ of 0.044, which was lower than all historical accessions, albeit in the same range (Table 1).

Genotyping of functional markers revealed that almost all the individuals carried the genotypes associated with six-row spike (*vrs1*, *int-c*), no flowering response to photoperiod (*ppd-H1*) and spring growth habit (*HvCEN*, *HvFT*), except for a single individual in NM.0406124 carrying the indel in the *HvFT1* promotor associated with winter growth habit. Of two SNPs in *HvP5CS1* associated with drought resistance, only one individual from each of NM.0406291 and NM.0406128 had the two alleles associated with drought resistance.

### 3.2. Analyses of Population Structure Suggest a Central and a Peripheral Genetic Cluster in Jämtland

Structure analysis was run for the dataset including both historical and extant accessions from Jämtland and Trøndelag (Figure 1; Figure 2). BestK from CLUMPAK suggested that the genetic diversity was best described using either two or five clusters, with K = 2 having the highest H′, and the highest delta using the Evanno method, Ln prob K = k was highest at K = 5 (Table S1). At K = 2, seven historical accessions from the major agricultural district around lake Storsjön and NGB6927, lacking precise geographical coordinates, were assigned to one of the clusters (blue in Figure 1a). The accessions more peripheral to the Storsjön district and the accessions from Trøndelag were assigned to the other cluster (orange in Figure 1a). Only two accessions from the Storsjön district (NM.0406291 and NM.0406122) were assigned to this cluster. At higher values of K, the extant accessions NGB13670 from Trøndelag (purple in Figures 1b–d and 2), NGB15157 from close to the Norwegian border (green in Figures 1c,d and 2) and NGB13394 (maroon in Figures 1d and 2) from Trøndelag, successively formed individual clusters. Although some individuals from some of the historical accessions clustered with these accessions, the majority of the historical individuals retained their clustering from K = 2 and the cluster of accessions from Storsjön district remained intact (Figure 1b–d). Analyzing the dataset using only historical accessions did not reveal any additional clustering (Figure S1).
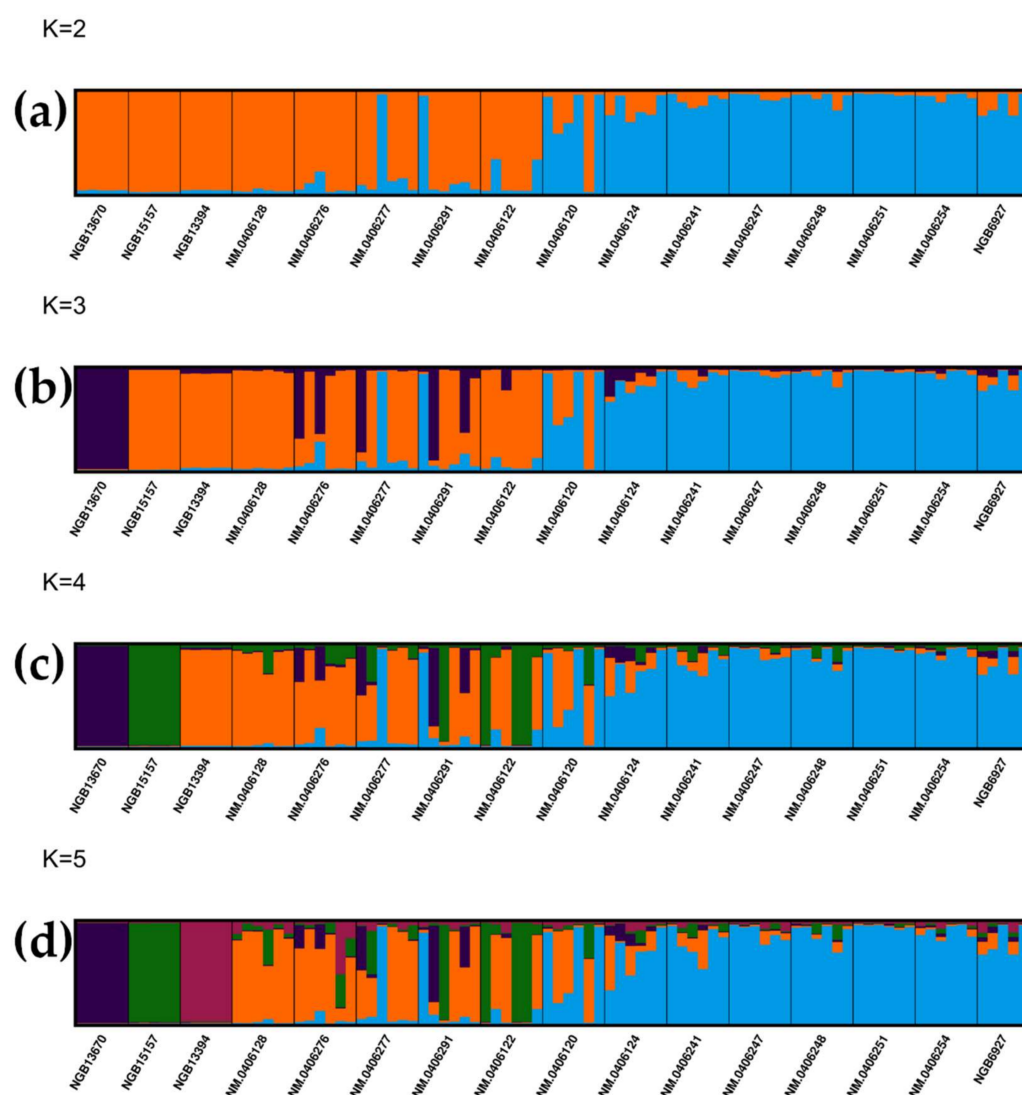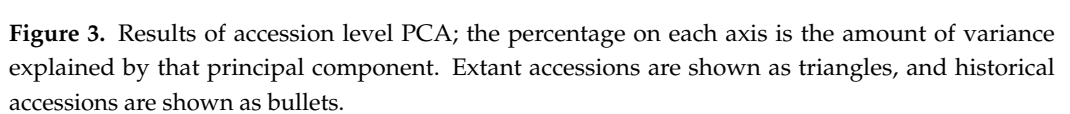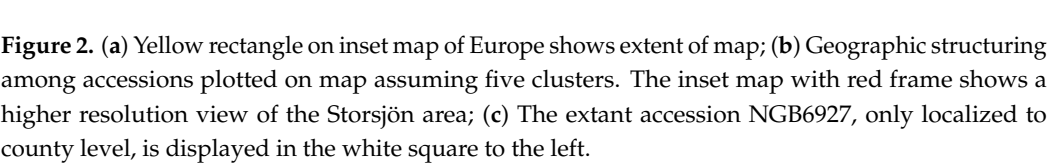
**Figure 1.** Bar plots of Structure analysis (CLUMPP output) assuming two to five clusters (K). Each line represents one individual, and the length of each color section shows the proportion of the individual's genetic diversity belonging to a specific cluster. Results of Structure analysis assuming (**a**) two clusters (**b**) three clusters (**c**) four clusters (**d**) five clusters.

To further investigate the distribution of genetic diversity among the accessions from Jämtland and Trøndelag, principal component analysis (PCA) was carried out (Figure 3). The first principal component explained 25.71% of the observed diversity and primarily corresponded to the Structure clustering seen at K = 2, with most accessions from the Storsjön district and NGB6927, falling at the positive end of PC1. PC2 explained 17.38% and primarily separated NGB13670 (bottom of Figure 3) from the remaining accessions, but also NM.0406128 (top right corner of Figure 3). Both NGB15157 and NGB13394 clustered among the historical accessions from Jämtland peripheral to the Storsjön district. Individual-level PCA showed similar clustering as the accession-level PCA, but with no separation of NM.0406128 along PC2 (Figure S2). With the exception of NGB13670, all extant individuals fell within the space occupied by the historical individuals along PC1 and PC2.

**Figure 2.** (**a**) Yellow rectangle on inset map of Europe shows extent of map; (**b**) Geographic structuring among accessions plotted on map assuming five clusters. The inset map with red frame shows a higher resolution view of the Storsjön area; (**c**) The extant accession NGB6927, only localized to county level, is displayed in the white square to the left.

**Figure 3.** Results of accession level PCA; the percentage on each axis is the amount of variance explained by that principal component. Extant accessions are shown as triangles, and historical accessions are shown as bullets.

### 3.3. Differentiation of Accessions Is Partly Explained by Isolation by Distance

To quantify the differentiation among accessions, pairwise $F_{ST}$ was calculated for both historical and extant accessions. $F_{ST}$ values ranged between 0.03 and 1 (Table 3). At variable

SNPs, the extant accessions NGB13394, NGB13670 and NGB26268 were all completely fixed for different alleles (resulting in pairwise $F_{ST}$ values of 1), whereas NGB6927 shared some diversity with the other extant accessions and to a larger extent the historical accessions, in particular NM.0406247 ($F_{ST}$ = 0.12). (Table 3). Among the historical accessions the highest $F_{ST}$ value was found between the pair NM.0406251 and NM.0406122 and the pair NM.0406251 and NM.0406128 ($F_{ST}$ = 0.33 for both pairs). In both cases one accession belonged to the Storsjön (blue) cluster and one to the peripheral (orange) cluster (Figure 1a). The lowest $F_{ST}$ value (0.03) was found when comparing NM.0406247 and NM.0406254, both belonging to the Storsjön (blue) cluster.

**Table 3.** Pairwise Fst values for all pairs of accessions.

| Acc\Acc | NGB13670 | NGB15157 | NGB13394 | NM.0406128 | NM.0406276 | NM.0406277 | NM.0406291 | NM.0406122 | NM.0406120 | NM.0406124 | NM.0406241 | NM.0406247 | NM.0406248 | NM.0406251 | NM.0406254 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NGB15157 | 1.000 | | | | | | | | | | | | | | |
| NGB13394 | 1.000 | 1.000 | | | | | | | | | | | | | |
| NM.0406128 | 0.520 | 0.290 | 0.290 | | | | | | | | | | | | |
| NM.0406276 | 0.437 | 0.344 | 0.344 | 0.099 | | | | | | | | | | | |
| NM.0406277 | 0.508 | 0.442 | 0.392 | 0.124 | 0.088 | | | | | | | | | | |
| NM.0406291 | 0.349 | 0.341 | 0.280 | 0.113 | 0.074 | 0.076 | | | | | | | | | |
| NM.0406122 | 0.606 | 0.287 | 0.476 | 0.104 | 0.129 | 0.153 | 0.148 | | | | | | | | |
| NM.0406120 | 0.539 | 0.430 | 0.396 | 0.160 | 0.146 | 0.118 | 0.086 | 0.167 | | | | | | | |
| NM.0406124 | 0.481 | 0.516 | 0.423 | 0.252 | 0.225 | 0.179 | 0.141 | 0.284 | 0.141 | | | | | | |
| NM.0406241 | 0.591 | 0.540 | 0.501 | 0.252 | 0.218 | 0.192 | 0.184 | 0.246 | 0.101 | 0.150 | | | | | |
| NM.0406247 | 0.548 | 0.559 | 0.474 | 0.268 | 0.213 | 0.187 | 0.153 | 0.262 | 0.099 | 0.080 | 0.082 | | | | |
| NM.0406248 | 0.635 | 0.597 | 0.573 | 0.297 | 0.221 | 0.234 | 0.185 | 0.282 | 0.099 | 0.173 | 0.084 | 0.079 | | | |
| NM.0406251 | 0.662 | 0.639 | 0.612 | 0.326 | 0.278 | 0.237 | 0.218 | 0.328 | 0.157 | 0.119 | 0.157 | 0.072 | 0.134 | | |
| NM.0406254 | 0.564 | 0.544 | 0.494 | 0.272 | 0.214 | 0.184 | 0.151 | 0.266 | 0.097 | 0.084 | 0.065 | 0.031 | 0.078 | 0.077 | |
| NM.0406254 | 0.644 | 0.672 | 0.606 | 0.345 | 0.276 | 0.248 | 0.214 | 0.291 | 0.198 | 0.183 | 0.184 | 0.119 | 0.169 | 0.210 | 0.156 |

Correlations between the Euclidian geographic distance and the genetic distance (measured as Edwards' distance) between accession pairs suggested the presence of isolation by distance, both for the full data set (r = 0.723; *p* < 0.001) and for the historical accessions only (r = 0.437; *p* = 0.004). We also tested whether latitude, longitude, altitude or length of growth season could explain the observed clustering by using regression modelling. Of the different parameters explored, only longitude significantly predicted average accession cluster membership ($F_{(1,10)}$ = 8.653, *p* = 0.0147, $r^2$ = 0.464). The remaining independent variables did not significantly predict any of their dependent variables (data not shown).

### 3.4. Barley from Jämtland Both Belong to and Differ from the Genetic Cluster of Barley in Central Scandinavia

To further investigate the genetic structure of 19th century barley, the data generated in this study was combined with previously published data [6,19] and the two additional Norwegian accessions. This resulted in a dataset of extant and historical accessions from large parts of the Scandinavian peninsula with high resolution sampling from Jämtland and the northernmost parts of Scandinavia. BestK from CLUMPAK and H' suggested the data was best explained by either three or five clusters (Table S2). K = 3 primarily separated the accessions into the southern, central, and northern clusters previously described for the Scandinavian peninsula (Figure 4a, [6]). The accessions from Jämtland and Trøndelag primarily belonged to the central cluster. At K = 4, four extant accessions from Norway (part of the southern cluster at K = 3) formed a new cluster (Figure S3). At K = 5 a cluster consisting of accessions from the Storsjön area emerged as well as a cluster of accessions

from the Norwegian coast (Figure 4b). The accessions from Trøndelag and peripheral parts of Jämtland clustered with accessions from eastern central Sweden.
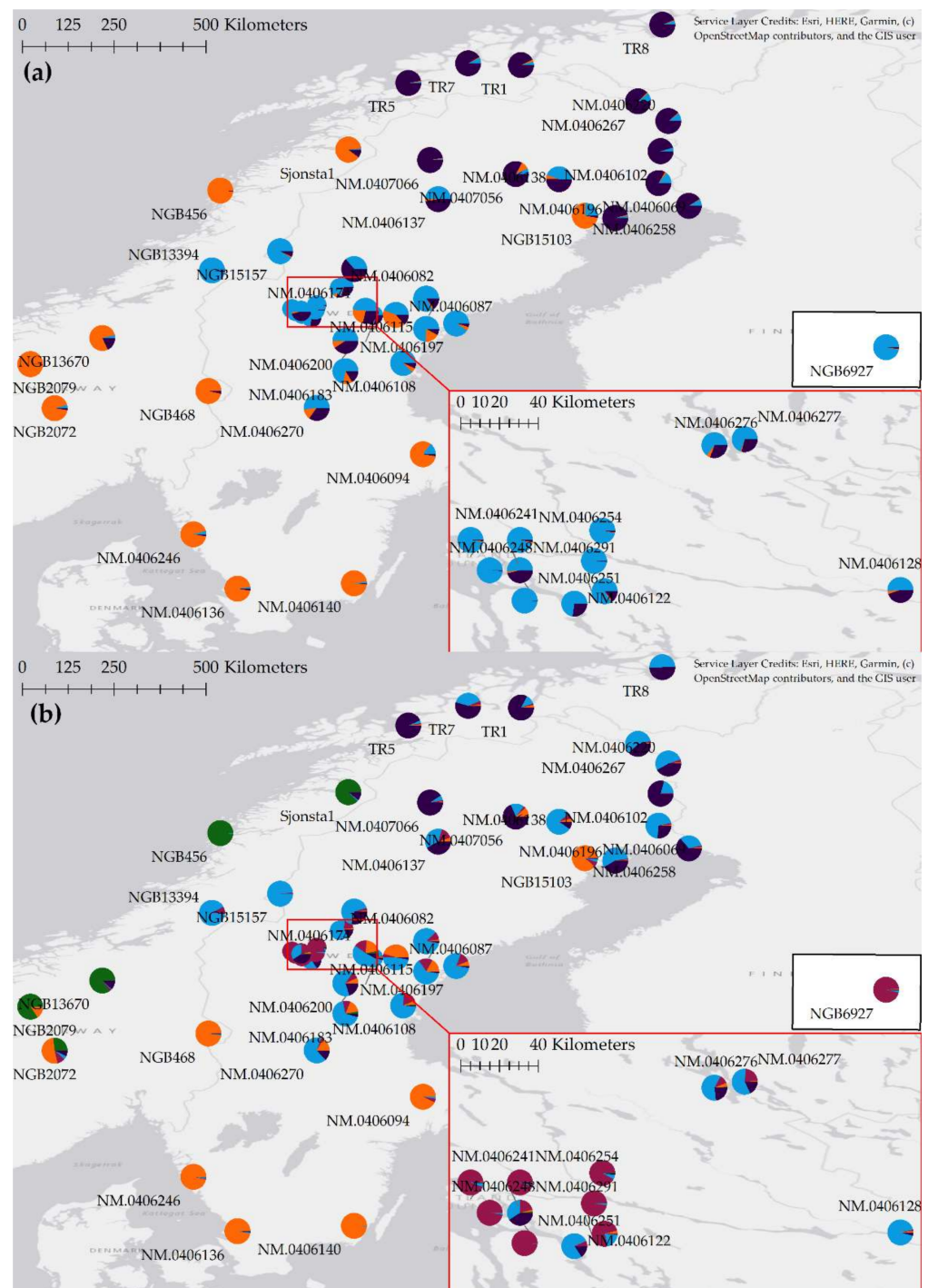


**Figure 4.** Geographic structuring among accessions plotted on map assuming (**a**) three clusters, (**b**) five clusters.

## 4. Discussion

The distribution of genetic diversity detected among the historical accessions showed that during the late 19th century clear geographic structuring of landrace barley existed in the county of Jämtland. Most of the historical accessions from the agricultural district surrounding lake Storsjön belonged to the same genetic cluster while accessions more peripheral to Storsjön belonged to a second cluster. In the Structure analysis of the full

Scandinavian data set, the latter clustered both with accessions in eastern central Sweden and with the accessions from northern Trøndelag (Figure 4b), suggesting seed exchange between these areas.

The county of Jämtland is known to have traded in historical times both to the west, with Trøndelag, and with eastern Swedish regions [33]. Among the peripheral accessions, NM.0406122 and NM.0406291 both originate from along the historical and present-day main road (Jonsson, 2001) linking Östersund, the county town of Jämtland, to the Swedish coast to the east and Norway to the west. The accessions NM.0406276 and NM.0406277 instead originated from along a historical and present-day major north-south inland road [69]. It is reasonable to assume that the areas from which these accessions originated had good trade connections both in the 19th century and before. Several of the peripheral accessions consisted of individuals belonging to different clusters (Figure 1d) and they had high genetic diversity, suggesting that the seed exchange gave rise to seed mixtures being cultivated.

Accessions from the agricultural district around lake Storsjön showed less evidence of seed exchange along known trade routes. It is known that agriculture in this area began much earlier than in other parts of Jämtland [70]. Agriculture in the area around lake Storsjön also differs from agriculture in other parts of Jämtland, as well as other parts of central Sweden. Among the differences are abiotic factors such as the fertile clay soils around lake Storsjön, agricultural practices such as crop rotation/fallow-system, and societal parameters such as large cohesive villages around the lake [36]. In comparison to the peripheral parts of Jämtland the area around Storsjön provided more advantageous agro-climatical conditions and agriculture was the main livelihood. The close genetic relationship we found among barley from different farms in the Storsjön region suggests a distinct and relatively isolated meta-population of barley was formed, likely through seed exchange. In contrast, Antonson [36] classified the areas of the more peripheral farms in Jämtland into a separate agricultural category. Here, agriculture played a more marginal role alongside trade as climatic conditions were harsher and crop failure was likely more common. We found that seed exchange over a larger geographical region, stretching both westwards and eastwards, resulted in barley genetically distinct from that in central Jämt-land being grown in these areas towards the end of the 19th century.

In the wider phylogeographical analysis (Figure 4) we confirm previously reported major genetic clusters of landrace barley from southern, middle and northern Scandinavia [6]. The peripheral accessions from Jämtland fall into the middle cluster. We further detect a new genetic cluster of accessions from along the Norwegian coast suggesting seed exchange along a costal route. The accessions from lake Storsjön district again form a separate genetic cluster also looking at this wider geographical scale.

Extant landrace crops are important genetic resources for plant improvement, for successfully meeting the challenges of climate change and for developing a more sustainable agriculture. Loss of genetic diversity in gene bank preserved landraces, which could potentially be valuable for plant improvement, has been reported repeatedly [13,71]. Landraces are expected to be adapted to the local conditions at their area of origin and a reliable provenience holds clues to the type of traits that can be expected and potentially utilized from a landrace crop [72]. A correct provenience is thus important for plant breeding schemes. Provenience also becomes critical when landraces are used to study agrarian history, as absent or incorrect passport data can obfuscate evolutionary signals. For example, the separation of *Ppd* alleles across Europe is much clearer when based on genotyping historical specimens compared to gene bank specimens [10,73].

The putative problems with historical analyses using ex situ-maintained accessions became evident in our study. Although the extant accessions NGB13394, NGB13670 and NGB15157 all formed separate clusters in our Structure analysis (Figure 1d), the separate clustering of NGB13394 and NGB15157 seems to result from these accessions being monomorphic. Both accessions fall among the historical Jämtland accessions in the PCA (Figure 3), and single individuals of the historical material match the genotypes of

these accessions well (Figure 1c,d). In addition, both accessions fall among the historical accessions from Jämtland in the full Scandinavian dataset at K = 3 (Figure 4). Taken together this suggests that NGB13394 and NGB15157 have reliable provenience with an origin in landrace barley from the Jämtland - Trøndelag region but represent only fractions of the original diversity. The only extant accession with maintained diversity was NGB6927, which is known to have a provenience in Jämtland, but its exact origin unknown. Our analyses showed that it belonged to the Storsjön cluster, and it had a level of genetic diversity similar to that of the historical accessions from the same region. We consequently consider it to be a good representative of the barley once cultivated in the Storsjön area of Jämtland.

## 5. Conclusions

In conclusion, although a limited number of markers were used, we could detect geographical structuring of the genetic diversity within the Jämtland - Trøndelag region. We suggest that both climate and economic practices have affected the distribution of genetic diversity in landrace populations. In geographic parts of the county where travelling trader-farmers were common [34], genetic diversity gathered from a much larger region was assembled in the farmer´s fields. In the farming-intense area around lake Storsjön, a unique and locally adapted type of barley developed instead.

## References

1. Zeven, A.C. Landraces: A review of definitions and classifications. *Euphytica* **1998**, *104*, 127–139. [CrossRef]
2. Leino, M.W. *"Primitive" Landraces: Swedish Farmers' Perception on Seed Traits before the Era of Crop Improvement, in Seedways. The Circulation, Control and Care of Plants in a Warmer World*; Karlsson, A.R.B., Ed.; Vitterhets Historie och Antikvitets Akademien: Stockholm, Sweden, 2021.

3.  Villa, T.C.; Maxted, N.; Scholten, M.; Ford-Lloyd, B. Defining and identifying crop landraces. *Plant Genet. Resour.* **2005**, *3*, 373–384. [CrossRef]

4.  Harlan, J.R. Our vanishing genetic resources. *Science* **1975**, *188*, 618–621. [CrossRef]

5.  Berg, T. Landraces and folk varieties: A conceptual reappraisal of terminology. *Euphytica* **2009**, *166*, 423–430. [CrossRef]

6.  Forsberg, N.E.; Russell, J.; Macaulay, M.; Leino, M.W.; Hagenblad, J. Farmers without borders—Genetic structuring in century old barley (Hordeum vulgare). *Heredity* **2015**, *114*, 195–206. [CrossRef] [PubMed]

7.  Pusadee, T.; Wongtamee, A.; Rerkasem, B.; Olsen, K.M.; Jamjod, S. Farmers drive genetic diversity of Thai Purple rice (*Oryza sativa* L.) landraces. *Econ. Bot.* **2019**, *73*, 76–85. [CrossRef]

8.  De Luca, D.; Cennamo, P.; Del Guacchio, E.; Di Novella, R.; Caputo, P. Conservation and genetic characterisation of common bean landraces from Cilento region (southern Italy): High differentiation in spite of low genetic diversity. *Genetica* **2018**, *146*, 29–44. [CrossRef]

9.  Hagenblad, J.; Morales, J. An Evolutionary Approach to the History of Barley (Hordeum vulgare) Cultivation in the Canary Islands. *Afr. Archaeol. Rev.* **2020**, *37*, 579–595. [CrossRef] [PubMed]

10. Jones, H.; Lister, D.L.; Bower, M.A.; Leigh, F.J.; Smith, L.M.; Jones, M.K. Approaches and constraints of using existing landrace and extant plant material to understand agricultural spread in prehistory. *Plant Genet. Resour.* **2008**, *6*, 98–112. [CrossRef]

11. Van Heerwaarden, J.; Doebley, J.; Briggs, W.H.; Glaubitz, J.C.; Goodman, M.M.; Gonzalez, J.D.; Ross-Ibarra, J. Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1088–1092. [CrossRef] [PubMed]

12. Chebotar, S.; Roder, M.S.; Korzun, V.I.; Borner, A. Genetic integrity of ex situ genebank collections. *Cell. Mol. Biol. Lett.* **2002**, *7*, 437–444.

13. Parzies, H.; Spoor, W.; Ennos, R. Genetic diversity of barley landrace accessions (*Hordeum vulgare* ssp. vulgare) conserved for different lengths of time in ex situ gene banks. *Heredity* **2000**, *84*, 476–486. [CrossRef]

14. Hagenblad, J.; Zie, J.; Leino, M.W. Exploring the population genetics of genebank and historical landrace varieties. *Genet. Resour. Crop Evol.* **2012**, *59*, 1185–1199. [CrossRef]

15. Leino, M.W.; Hagenblad, J.; Edqvist, J.; Strese, E.M. DNA preservation and utility of a historic seed collection. *Seed Sci. Res.* **2009**, *19*, 125–135. [CrossRef]

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

17. Leino, M.W. Frösamlingar på museer–ny teknik gör värdelösa föremål värdefulla igen. *Nord. Museol.* **2010**, 96. [CrossRef]

18. Selçuk, A.; Forsberg, N.; Hagenblad, J.; Leino, M.W. Molecular genotyping of historical barley landraces reveals novel candidate regions for local adaption. *Crop Sci.* **2015**, *55*, 2766–2776.

19. Forsberg, N.E.; Leino, M.W.; Hagenblad, J. Population structure in landrace barley (*Hordeum vulgare* L.) during the late 19th century crop failures in Fennoscandia. *Heredity* **2019**, *123*, 733–745. [CrossRef] [PubMed]

20. Badr, A.; Rabey, H.E.; Effgen, S.; Ibrahim, H.H.; Pozzi, C.; Rohde, W.; Salamini, F. On the origin and domestication history of barley (Hordeum vulgare). *Mol. Biol. Evol.* **2000**, *17*, 499–510. [CrossRef] [PubMed]

21. FAO (Food and Agricultural Organization of the United Nations). FAOSTAT. 2020. Available online: http://www.fao.org/faostat (accessed on 23 June 2020).

22. FAO (Food and Agricultural Organization of the United Nations). Crop Prospects and Food Situation—Quarterly Global Report No. 1. 2020: Rome. Available online: http://www.fao.org/giews/reports/crop-prospects/en/ (accessed on 26 June 2020).

23. Akar, T.; Avci, M.; Dusunceli, F. *Barley: Post Harvest Operations*; Food and Agriculture Organization (FAO) of the United Nations, The Central Research Institute for Field Crops: Ankara, Turkey, 2004; p. 64.

24. Newton, A.C.; Flavell, A.J.; George, T.S.; Leat, P.; Mullholland, B.; Ramsay, L.; Revoredo-Giha, C.; Russell, J.; Steffenson, B.J.; Swanston, J.S.; et al. Crops that feed the world 4. Barley: A resilient crop? Strengths and weaknesses in the context of food security. *Food Secur.* **2011**, *3*, 141–178. [CrossRef]

25. Pasam, R.K.; Sharma, R.; Walther, A.; Özkan, H.; Graner, A.; Kilian, B. Genetic diversity and population structure in a legacy collection of spring barley landraces adapted to a wide range of climates. *PLoS ONE* **2014**, *9*, e116164. [CrossRef]

26. Poets, A.M.; Fang, Z.; Clegg, M.T.; Morrell, P.L. Barley landraces are characterized by geographically heterogeneous genomic origins. *Genome Biol.* **2015**, *16*, 1–11. [CrossRef]

27. Florence, A.; Ennos, R.A.; Hoad, S.P.; Hoebe, P.N. Variation in light interception traits in European spring barley landraces. *Field Crops Res.* **2019**, *241*, 107549. [CrossRef]

28. Kumar, A.; Verma, R.P.; Singh, A.; Sharma, H.K.; Devi, G. Barley landraces: Ecological heritage for edaphic stress adaptations and sustainable production. *Environ. Sustain. Indic.* **2020**, *6*, 100035. [CrossRef]

29. Sørensen, L.; Karg, S. The expansion of agrarian societies towards the north–new evidence for agriculture during the Mesolithic/Neolithic transition in Southern Scandinavia. *J. Archaeol. Sci.* **2014**, *51*, 98–114. [CrossRef]

30. Vorren, K.-D. Farm development at the Arctic cereal limit in northern Norway—Continuity and discontinuities. *Veg. Hist. Archaeobotany* **2005**, *14*, 161–170. [CrossRef]

31. Leino, M.W. *Spannmål: Svenska Lantsorter*; Nordiska Museets Förlag: Stockholm, Sweden, 2017.

32. Hellström, P. *Norrlands Jordbruk*; Almqvist & Wiksell: Uppsala, Sweden, 1917.

33. Brink, S. Marknader och forbönder. In *Perspektiv på Härjedalen*; Sporrong, U., Ed.; Svenska vyer: Sveg, Sweden, 1995; pp. 57–65.

34. Welinder, S. The northern margin of cereal cultivation in Sweden during the Middle Ages. *Fornvännen* **2019**, *114*, 36–42.
35. Wallin, J.-E.; Oskarsson, B. *Odlingslandskapets Framväxt i Storsjöbygden: En Miljöhistorisk Studie i Järnåldersbygd, Baserad på Pollenanalyser och GIS-Baserade Kartanalyser*; Jämtlands Läns Museum: Östersund, Sweden, 2002.
36. Antonson, H. *Jämtlands Äldre Agrarlandskap: Försök Till en Regional Indelning Baserad på Lantmäteriakter för Tiden 1650–1899*; Kulturmiljöenheten, Länsstyr. i Jämtlands län: Östersund, Sweden, 1993.
37. Statistics Sweden, Agricultural Statistics. 2020. Available online: http://www.statistikdatabasen.scb.se/pxweb/en/ssd/ (accessed on 13 July 2020).
38. Dribe, M.; Olsson, M.; Svensson, P. *Famines in the Nordic Countries, AD 536–1875*; Lund University, Department of Economic History: Lund, Sweden, 2015.
39. Osvald, H. *Åkerns Nyttoväxter*; Sv. Litteratur: Stockholm, Sweden, 1959.
40. Anonymous. *Allmänna Konst- och Industriutställningen i Stockholm: Officiel Katalog Öfver Industriafdelningen*; Central-Tryckeriet: Stockholm, Sweden, 1897.
41. Sveriges Hembygdsförbund. Sveriges Hembygdsförbund. 2020. Available online: https://www.hembygd.se/ (accessed on 17 April 2020).
42. He, C.; Holme, J.; Anthony, J. *SNP Genotyping: The KASP Assay, in Crop Breeding*; Springer: New York, NY, USA, 2014; pp. 75–86.
43. Moragues, M.; Comadran, J.; Waugh, R.; Milne, I.; Flavell, A.J.; Russell, J.R. Effects of ascertainment bias and marker number on estimations of barley diversity from high-throughput SNP genotype data. *Theor. Appl. Genet.* **2010**, *120*, 1525–1534. [CrossRef] [PubMed]
44. Hagenblad, J.; Morales, J.; Leino, M.W.; Rodríguez-Rodríguez, A.C. Farmer fidelity in the Canary Islands revealed by ancient DNA from prehistoric seeds. *J. Archaeol. Sci.* **2017**, *78*, 78–87. [CrossRef]
45. Jones, H.; Leigh, F.J.; Mackay, I.; Bower, M.A.; Smith, L.M.; Charles, M.P.; Jones, G.; Jones, M.K.; Brown, T.A.; Powell, W. Population-based resequencing reveals that the flowering time adaptation of cultivated barley originated east of the Fertile Crescent. *Mol. Biol. Evol.* **2008**, *25*, 2211–2219. [CrossRef]
46. Ramsay, L.; Comadran, J.; Druka, A.; Marshall, D.F.; Thomas, W.T.; Macaulay, M.; MacKenzie, K.; Simpson, C.; Fuller, J.; Bonar, N.; et al. INTERMEDIUM-C, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene TEOSINTE BRANCHED 1. *Nat. Genet.* **2011**, *43*, 169–172. [CrossRef]
47. Comadran, J.; Kilian, B.; Russell, J.; Ramsay, L.; Stein, N.; Ganal, M.; Shaw, P.; Bayer, M.; Thomas, W.; Marshall, D.; et al. Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat. Genet.* **2012**, *44*, 1388. [CrossRef]
48. Yan, L.; Fu, D.; Li, C.; Blechl, A.; Tranquilli, G.; Bonafede, M.A.; Sanchez, A.L.; Valarik, M.; Yasuda, S.H.; Dubcovsky, J. The wheat and barley vernalization gene VRN3 is an orthologue of FT. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19581–19586. [CrossRef] [PubMed]
49. Xia, Y.; Li, R.; Bai, G.; Siddique, K.H.; Varshney, R.K.; Baum, M.; Yan, G.; Guo, P. Genetic variations of HvP5CS1 and their association with drought tolerance related traits in barley (*Hordeum vulgare* L.). *Sci. Rep.* **2017**, *7*, 1–10. [CrossRef] [PubMed]
50. Komatsuda, T.; Pourkheirandish, M.; He, C.; Azhaguvel, P.; Kanamori, H.; Perovic, D.; Stein, N.; Graner, A.; Wicker, T.; Tagiri, A.; et al. Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1424–1429. [CrossRef] [PubMed]
51. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. 2019. Available online: https://www.R-project.org/ (accessed on 10 May 2020).
52. RStudio Team. RStudio: Integrated Development for R. 2018. Available online: http://www.rstudio.com/ (accessed on 10 May 2020).
53. Wickham, H.; Averick, M.; Bryan, J.; Chang, W.; McGowan, L.D.; François, R.; Grolemund, G.; Hayes, A.; Henry, L.; Hester, J.; et al. Welcome to the Tidyverse. *J. Open Source Softw.* **2019**, *4*, 1686. [CrossRef]
54. Jombart, T. adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **2008**, *24*, 1403–1405. [CrossRef]
55. Goudet, J.; Jombart, T. Hierfstat: Estimation and Tests of Hierarchical F-statistics. *R Package Version 0.04-22* **2015**, *10*. Available online: https://cran.r-project.org/web/packages/hierfstat/index.html (accessed on 27 June 2020).
56. Tang, Y.; Horikoshi, M.; Li, W. ggfortify: Unified interface to visualize statistical results of popular R packages. *R J.* **2016**, *8*, 478–489. [CrossRef]
57. Schauberger, P.; Walker, A. Openxlsx: Read, Write and Edit XLSX Files (Version 4.1.5). 2020. Available online: https://CRAN.R-project.org/package=openxlsx (accessed on 23 July 2020).
58. Rousset, F. genepop'007: A Complete Re-Implementation of the Genepop Software for Windows and Linux. *Mol. Ecol. Resour.* **2008**, *8*, 103–106. [CrossRef]
59. Slowikowski, K. ggrepel: Automatically Position Non-Overlapping Text Labels with'ggplot2'. R Package Version 0.8.2. 2020. Available online: https://cran.r-project.org/web/packages/ggrepel/index.html (accessed on 27 June 2020).
60. Nei, M. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 3321–3323. [CrossRef]
61. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [CrossRef] [PubMed]
62. Falush, D.; Stephens, M.; Pritchard, J.K. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **2003**, *164*, 1567–1587. [CrossRef] [PubMed]

63. Falush, D.; Stephens, M.; Pritchard, J.K. Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Mol. Ecol. Notes* **2007**, *7*, 574–578. [CrossRef] [PubMed]
64. Hubisz, M.J.; Falush, D.; Stephens, M.; Pritchard, J.K. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* **2009**, *9*, 1322–1332. [CrossRef]
65. Kopelman, N.M.; Mayzel, J.; Jakobsson, M.; Rosenberg, N.A.; Mayrose, I. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **2015**, *15*, 1179–1191. [CrossRef] [PubMed]
66. ESRI. ArcGIS Desktop. 2017. Available online: https://desktop.arcgis.com/en/arcmap/ (accessed on 10 May 2020).
67. Lantmäteriet. Öppna Data. 2020. Available online: https://www.lantmateriet.se/sv/Kartor-och-geografisk-information/oppna-data/ (accessed on 6 August 2020).
68. SMHI. Vegetationsperiod. 2020. Available online: https://www.smhi.se/kunskapsbanken/klimat/fenologi/vegetationsperiod-1.6270 (accessed on 6 August 2020).
69. Jonsson, J. *Historiska Vägar i Jämtlands Län*; Jamtli Förlag: Östersund, Sweden, 2001.
70. Salvesen, H. *Jord i Jemtland: Bosetningshistoriske og Økonomiske Studier i Grenseland ca. 1200–1650*; AB Wilénska Bokhandelns Förlag: Östersund, Sweden, 1979.
71. Cross, R.; Wallace, A. Loss of genetic diversity from heterogeneous self-pollinating genebank accessions. *Theor. Appl. Genet.* **1994**, *88*, 885–890. [CrossRef]
72. Lasky, J.R.; Upadhyaya, H.D.; Ramu, P.; Deshpande, S.; Hash, C.T.; Bonnette, J.; Juenger, T.E.; Hyma, K.; Acharya, C.; Mitchell, S.E.; et al. Genome-environment associations in sorghum landraces predict adaptive traits. *Sci. Adv.* **2015**, *1*, e1400218. [CrossRef] [PubMed]
73. Lister, D.L.; Thaw, S.; Bower, M.A.; Jones, H.; Charles, M.P.; Jones, G.; Smith, L.M.; Howe, C.J.; Brown, T.A.; Jones, M.K. Latitudinal variation in a photoperiod response gene in European barley: Insight into the dynamics of agricultural spread from 'historic' specimens. *J. Archaeol. Sci.* **2009**, *36*, 1092–1098. [CrossRef]