Archaeological and historical materials as a means to explore Finnish crop history

Authors: Mia Lempiäinen-Avci[1,2,3], Maria Lundström[4], Sanna Huttunen[2], Matti W. Leino[5, 6], Jenny Hagenblad[4]

[1] Department of Biology, University of Turku, FI-20014, Finland

[2] Herbarium, Biodiversity unit, University of Turku, FI-20014, Finland

[3] Archaeological museum, University of Stavanger, NO-4036 Stavanger, Norge

[4] IFM Biology, Linköping University, SE-581 83 Linköping, Sweden

[5] Nordiska Museet, Swedish Museum of Cultural History, Box 27820, SE-115 93 Stockholm, Sweden

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

Abstract

In Northern Europe, barley (*Hordeum vulgare* L.) has been cultivated for almost 6000 years. Thus far, 150-year-old grains from historical collections have been used to investigate the distribution of barley diversity and how the species has spread across the region. Genetic studies of archaeobotanical material from agrarian sites could potentially clarify earlier migration patterns and cast further light on the origin of barley landraces. In this study, we aimed to evaluate different archaeological and historical materials with respect to DNA content, and to explore connections between Late Iron Age and medieval barley populations and historical samples of barley landraces in north-west Europe. The material analysed consisted of archaeological samples of charred barley grains from four sites in southern Finland, and historical material, with 33 samples obtained from two herbaria and the seed collections of the Swedish museum of cultural history.

The DNA concentrations obtained from charred archaeological barley remains were too low for successful KASP genotyping confirming previously reported difficulties in obtaining aDNA from charred remains. Historical samples from herbaria and seed collection confirmed previously shown strong genetic differentiation between two-row and six-row barley. Six-row

barley accessions from northern and southern Finland tended to cluster apart, while no geographical structuring was observed among two-row barley. Genotyping of functional markers revealed that the majority of barley cultivated in Finland in the late 19th and early 20th century was late-flowering under increasing day-length, supporting previous findings from northern European barley.

Keywords: aDNA**,** archaeobotany**,** barley, genetic diversity, *Hordeum vulgare*, KASP, landraces

Introduction

Barley (*Hordeum vulgare* L.) is one of the founder crop species of Neolithic agriculture (Zohary et al. 2012). Its domesticated form was derived from wild populations of *Hordeum vulgare* ssp. *spontaneum* (K. Koch) Asch. & Graebn., a species occurring in North Africa, the Middle East, parts of the Indian subcontinent, and south-west China (e.g. Lister & Jones 2012). Early indications of domesticated barley have been documented from several archaeological sites in South-west Asia, dated to ca. 10,500-9,550 cal BC (Zohary et al. 2012). According to molecular evidence, however, barley seems to have been domesticated more than once, with independent origins in the Fertile Crescent and in Central Asia (Morrell & Clegg 2007). Additional origins have been suggested in Morocco (Igartua et al. 2013) and on the East Asian Tibetan Plateau (Wang et al. 2016). As barley can withstand many different climatic and soil conditions, it subsequently spread over vast areas and became a principal dietary grain. Today it is cultivated worldwide, in temperate areas as a summer crop and in tropical areas as a winter crop (von Bothmer et al. 2003).

Wild *H. vulgare* ssp. *spontaneum* has rachises with triplets composed of one fertile spikelet and two reduced ones, a form that is known as two-rowed barley. While the earliest domesticated populations were two-row forms similar to wild barley, mutations with three fertile spikelets per rachis segment, known as six-row barley and with increased number of seeds on each spike, occurred during the early stages of domestication (Lister & Jones 2012; Zohary et al. 2012). With the cloning and identification of the *Vrs1* gene as being responsible for row-type, it was shown that causative mutations leading to a loss of function of the *Vrs1* gene and a six-row phenotype occurred independently in several populations (Komatsuda et al. 2007). There is further morphological diversity within six-row barley, which is sometimes subdivided into different species, such as the four-rowed form *H. vulgare* ssp. *tetrastichum* with fertile lateral spikelets on lax ears, and the six-rowed form *H. vulgare* ssp. *hexastichum* with dense and short (ca 4.5 cm) ears, which in cross-section is star-like (e.g. in Finnish tähtiohra “star-barley”; Heinonen 2009, also Gadd 1770). However, the morphological delimitation between four-row and six-row barley is unclear, and they are not formally distinguished as separate taxa (Soreng 2003).

In Northern Europe barley has been cultivated for almost 6000 years (Kirleis et al. 2012; Alenius et al. 2017), and it has been a key crop in many parts of Fennoscandia (Finland, Norway and Sweden) (Engelmark 1992). From the Neolithic and Bronze Age, both naked and

hulled barley are recovered from archaeological assemblages, but the naked form is more common (Behre 1983; Jacomet & Kreuz 1999; Jacomet 2006; Kirleis et al. 2012). However, from the end of the Bronze Age hulled barley dominates in the archaeological material (Grabowski 2011; Stika & Heiss 2013).

The first evidence of farming in Finland includes archaeobotanical finds of naked barley dated to 1690-1260 cal BC (radiocarbon date with 1 σ probability; Pihlman & Seppä-Heikka 1985; Vuorela & Lempiäinen 1988; Asplund 2008), and barley grains have also been found at later Iron Age and Medieval sites (e.g. Taivainen 2004, 2007; Vuorinen 2009; Raninen 2013; Lempiäinen-Avci et al. 2017). Climatic conditions allow barley cultivation throughout Finland, and in the north barley was occasionally cultivated on slash-and-burn fields (Grotenfelt 1922). The northernmost recorded barley cultivation in Finland is at Tsuolisjärvi (Suolijärvi) in Inari (69° 28´N) (Elfving 1897).

It is not known whether the early archaeological finds of barley in Finland were of the two-rowed or six-rowed form. Historical records show that both forms were cultivated in parallel from at least AD 1700 onwards (Gadd 1770), although in Sweden two-row barley was described as a novel type in the 17th century (Leino 2017). By the 1920s six-rowed barley had been almost completely replaced by four-row and two-row barley varieties in Southern Finland (Sauli 1927). Today, the majority of the barley cultivated in Finland is of the two-row type favoured in brewing, but cultivars of six-row barley (*H. vulgare* ssp. *hexastichum*) are used successfully for cultivation in the northernmost areas (Heinonen 2009).

According to historical records, grain was imported from abroad during the first decades of the 20th century. For example, grain was imported to Inari from Kirkenes and Reisivuono in the Finnmark area in northern Norway (Sauli 1927; Heinonen 2009). However, the import of cereals has also been strictly regulated in the past in Finland, especially during the 19th century, when Finland was under Russian rule (e.g. Herstad 2000; Lunden 2004). Barley was imported from Sweden and Estonia to southern Finland, when frost had destroyed the crop, e.g. in 1811-1812 and 1867-1868 (Lindström 1905; Åström 1980), but otherwise Finland was dependent on domestic barley production or imports from Russia (Åström 1980). It is not known whether imported barley was used only for food, fodder and brewing or also as seed grain. Possibly as a result of connections to the east, at the beginning of the 20th century

barley varieties in eastern Finland, such as Viborg province, were morphologically similar to northern Russian varieties (Sauli 1927).

Phylogeographic studies have proved useful in exploring past patterns of migration in wild species (Avise 2009), and the spread of cultivation and seed trade in domestic species (e.g. Papa et al. 1998; Jones et al. 2012; Leino et al. 2013; Jones et al. 2013; Roullier et al. 2013). In Fennoscandia, 150-year-old grains from historical collections have been useful in understanding the distribution of barley diversity. A latitudinally-structured genetic diversity, first detected in Sweden (Leino & Hagenblad 2010), was later shown to be shared across all of Fennoscandia (Forsberg et al. 2015). However, the time depth of these phylogeographical patterns, based on material from the late 19th century, is unknown.

Archaeobotanical material from early agrarian sites could potentially clarify early migration patterns and systems of barley trade and exchange, as well as cast light on the origin of barley landraces. While data on geographical location and morphological characteristics of barley based on archaeobotanical findings and historical documents can only give a scattered picture of the origin and spread of barley cultivation, aDNA may cast light on growth traits and genetic relatedness between specimens of cultivated crops (Jones et al. 2008b; Palmer et al. 2009).

Unfortunately, in most cases poor preservation of DNA hampers the use of archaeological grain for genetic studies (Nistelberger et al. 2016). Archaeological grain is most commonly preserved in a charred state. Successful molecular studies of charred plant remains have been reported (Bunning et al. 2012), though concerns have been raised about the feasibility of obtaining indigenous genetic data from such remains (Nistelberger et al. 2016). At the very least the utility of charred grains will depend on extent of charring, which may be impossible to estimate from the appearance of the grain (Bunning et al. 2012). Cereals preserved under waterlogged conditions might be more amenable to aDNA analysis, but such remains are scarce (Palmer et al. 2012). Desiccated archaeobotanical grains are superior for DNA analyses (e.g. Mascher et al. 2016; Hagenblad et al. 2017). At Finnish archaeological sites waterlogged testa of grains from medieval or younger sites have been found, but like in most of Europe cereals preserved in a desiccated state are largely absent.

Extant Finnish landraces of barley have been well studied, e.g. by Elfving (1897), Grotenfelt (1922), Sauli (1927), Ahokas & Poukkula (1999), and Heinonen (2009). The phenological variation for example in maturation period between different landraces suggests that the origin of southern Finnish four-row barley may be different from that of the northern Finnish forms of barley (Heinonen 2009). Landraces with different maturation periods may result from import of grains to Lapland both from north and south, and from subsequent mixing of populations. Clear genetic differences between barley from northern and southern Scandinavia have also been reported (Leino & Hagenblad 2010; Forsberg et al. 2015; Leino 2017).

Our aim in this study was to evaluate different archaeological and historical materials with respect to DNA content, and if possible explore connections between Late Iron Age and Medieval barley populations and historical samples of barley landraces. We combined genetic data from earlier studies (Forsberg et al. 2015) with analysis of novel samples of historical barley from the eastern part of Fennoscandia. The eastern parts of Finland have been relatively poorly sampled in earlier studies, but the area is important for discovering possible eastern introduction routes into Finland.

Materials and methods

Archaeological and historical samples

Archaeological samples of charred grains of *Hordeum vulgare* were derived from four sites in southern Finland: Raisio Mulli, Espoo Mankby, Pirkkala Tursiannotko, and Hattula Retulansaari (Table 1). At the Raisio Mulli site excavations were carried out in 1994–1997. Mulli is a Late Iron Age and Early Medieval village, inhabited during the 10th to late 13th century cal AD (Vuorinen 2009). Espoo Mankby is a medieval village dating to the 12th to mid-16th century cal AD, where excavations were carried out in 2007–2013 (Lempiäinen-Avci et al. 2017). Archaeological studies at Pirkkala Tursiannotko have been conducted since 2010, and the site dates to cal AD 800–1100 (Raninen 2013). The Hattula Retulansaari site was excavated in 2005 and is dated to cal AD 700–900 (Taivainen 2004, 2007). Archaeobotanical samples from the study sites were stored at the Turku University Herbarium. All samples consisted of morphologically intact charred cereal grains.

Historical specimens were obtained from three different sources: two herbaria and one seed collection. One grain per spike, in total 14, was sampled from whole plants mounted on

herbarium sheets and stored at the Turku University Herbarium and the Botanical Museum of the Finnish Museum of Natural History, Helsinki (both denoted ML below) (Table 2). The oldest herbarium specimens date from 1870, and only specimens that were collected before the 1920s, i.e. before breeding and cultivation of improved cereal cultivars became widespread, were used in the study. From the seed collection of the Swedish museum of cultural history (Nordiska Museet; NM; Leino et al. 2009) six Finnish specimens were chosen. The specimens, consisting of 2-5 dl of grain, were harvested in 1882 and have been stored in their original glass jars since collection. From each specimen, six individual grains were chosen. Since individual grains from the same jar originate from the same harvested field they were treated as one population, henceforth named an accession. Data on the origin of the historical material was obtained from the herbarium and from seed jar labels (Table 2).

Genotyping data from nine Finnish barley accessions, consisting of six accessions from the seed collections of the Mustiala Agricultural Museum (MU; Leino 2010), two extant accessions from NordGen (NGB) and one accession from the Swedish museum of cultural history (NM), represented by six individual seeds each were obtained from Forsberg et al. (2015). Genotypes from three previously unpublished accessions, the majority of which had been genotyped for five or six individuals, genotyped alongside the accessions from Forsberg et al. (2015), were also included (Table 2).

DNA extraction

DNA extractions of the archaeological samples were carried out on single grains in a dedicated, chambered ancient DNA facility at the University of Warwick, Coventry, UK. No work with modern barley DNA had been performed in the lab, nor any PCRs. Suitable precautions were taken to avoid introducing foreign contaminants. Extractions were performed according to Palmer et al. (2009), but with five days of incubation in the extraction buffer and without the Amicon® concentration step. An additional wash step of the spin columns using acetone was also included. Quantification of DNA was performed using Qubit® dsDNA HS Assay Kit (Life technologies). From historical specimens, DNA was extracted at Linköping University according to Leino et al. (2009).

PCR amplification

Amplification of the P6 loop within the *trnL* locus in the archaeological samples was performed with semi-nested PCR using M13-tagged primers (Taberlet et al. 2007; Willerslev

et al. 2007). Each reaction contained 1 U of DreamTaq DNA polymerase (Thermo Scientific), 1 x DreamTaq Buffer (Thermo Scientific), 0.2 mM of each dNTP (Thermo Scientific), 0.1 µM of forward primer (trnLg with M13F-tag; TGTAAAACGACGGCCAGTGGGCAATCCTGAGCCAA), 0.1 µM of reverse primer (trnLh with M13R-tag; CAGGAAACAGCTATGACCTTGAGTCTCTGCACCTATC) and 1 µl of DNA template. In addition, 2 µg/µl of BSA (Thermo Scientific) was included in the first round of PCR. For the second PCR amplification only primers matching the M13 tag were used (M13F: TGTAAAACGACGGCCAGT and M13R: CAGGAAACAGCTATGACC respectively).

PCR conditions were as follows: initial denaturation for 2.5 min at 94°C; 25 cycles of 94°C for 30 s, 56°C for 30 s and 72°C for 15 s; and a final elongation at 72°C for 8 min. Conditions for the second PCR were: initial denaturation for 2.5 min at 94°C; 8 cycles of 94°C for 30 s, 68°C for 30 s with a decrease of 1°C per cycle, and 72°C for 20 s; 20 cycles of 94°C for 30 s, 60°C for 30 s and 72°C for 20 s; and a final elongation at 72°C for 10 min. The success of the PCRs was assessed on a 3 % agarose gel pre-stained with SYBR Safe (Invitrogen), and samples of sufficient quality were sent to Macrogen Europe, Netherlands, for sequencing.

SNP genotyping

Genotyping of all samples for 97 SNPs was carried out by LGC Genomics, using the KASP assay method (He et al. 2014; Semagn et al. 2014), which has previously been used successfully on desiccated archaeological samples (Hagenblad et al. 2017). All archaeological samples were genotyped four times: duplicates of the undiluted DNA sample, as well as 1:5 and 1:10 dilutions of the sample. Historical samples were genotyped a single time. Genotyping was also attempted on extraction blanks from the respective DNA extractions, and extraction blanks from the archaeological DNA extractions were included in duplicates: undiluted and 1:10 dilution.

Of the 97 SNPs, 90 were derived from the BOPA1 array (Kota et al. 2008). The remaining seven markers were located in causative or associated SNPs of the functional genes *Vrs1* (positions A40>F.S.; F75>L; E152>F.S.) (Komatsuda et al. 2007), *int-c* (nucleotide 124) (Ramsay et al. 2011), *Ppd-H1* (SNP48) (Jones et al. 2008a), *HvNAM-2* (nucleotide 798) (Cai et al. 2013) and *Lhcb1* (nucleotide 907) (Xia et al. 2012).

Data analysis

The 90 SNPs derived from the BOPA1 array were used to assess the samples for genetic structure. All individuals failing genotyping, and individuals with more than 15 % missing genotypes (among the 90 BOPA1 SNPs), were removed from further analysis leaving a final dataset consisting of 114 individuals from 19 accessions and 8 separate herbarium specimens.

The data from the 90 BOPA1 SNPs was analysed for geographic structure using the software *STRUCTURE* and Principal Component Analysis (PCA). *STRUCTURE* (v 2.3.4) was run using the haploid setting, and a model with correlated allele frequencies and admixture. The software was run with a burn-in length of 20,000 iterations followed by 50,000 iterations for estimating the parameters, with 10 repeated runs at each level of predetermined clusters (K) ranging from 1 to 15. The software *CLUMPP* (v 1.1.2) (Jakobsson & Rosenberg 2007) was used to compare the outcome of individual runs with the Greedy algorithm for $4 < K < 6$ and with the LargeKGreedy algorithm for $K \geq 6$. The number of clusters best describing the data was evaluated from the *CLUMPP* H' values and ΔK calculated according to Evanno et al. (2005). Results were visualized using *DISTRUCT* (v 1.1) (Rosenberg 2004). PCA was carried out using the command *prcomp* in the software R (v 2.5.0). In the PCA, the data were analyzed both at the accession and the individual level and the numbers of copies of each allele at each locus were treated as independent variables.

Results

DNA analysis of archaeological samples

DNA quantity, evaluated by a fluorometric assay, suggested DNA concentrations ranging from 0.03-3.22 ng/µl for the archaeological samples (Table 1). Neither of the negative controls extracted in parallel with the grain samples contained detectable levels of DNA, verifying that the DNA detected in the samples originated from the grains.

To determine the source of the DNA, a semi-nested PCR targeting *trnL* was carried out. Amplification was detected in 17 out of 27 samples, but not in any of the extraction or negative PCR controls. Sequencing of the 10 PCR-products of sufficient quality for further analysis (Table 1), followed by MegaBLAST comparison with the full nucleotide sequence database at NCBI, showed that the *trnL* sequences did not match those expected from barley. Some sequences instead generated partial hits against plant genera expected to be found in and around the excavation sites. We therefore concluded that the amplifying DNA originated

from contaminating DNA, quite possibly *in situ* contamination, rather than endogenous barley DNA from the charred samples.

Genotyping success of archaeological and historical samples

Genotyping was carried out on archaeological, herbarium, and seed collection samples. Genotyping of all markers failed for all extraction controls, verifying the absence of contaminating DNA. Genotyping also failed for all SNPs for all archaeological samples. This was the case for both undiluted samples and for the different dilutions. Of the herbarium specimens, genotyping was successful for all but six. Four specimens failed for all SNPs and an additional two specimens had a limited genotyping success rate, failing to genotype in 86 and 41 SNPs respectively (Table 2). Among the samples obtained from the seed collection only a single individual failed genotyping for all SNPs, while two samples failed to genotype 42 and 38 SNPs respectively (Table 2).

Functional markers

At the *PpdH1* only the individuals from the accession NM360 carried the responsive allele resulting in early flowering at increased day-length. At *Lhcb1* the majority of the seed collection individuals (with the exception of all individuals of NM297 and one individual of NM288) carried the markers associated with a high number of grains per spike, while all herbarium specimens except ML24 carried the marker associated with low numbers of grains per spike. At *HvNAM-2*, all individuals except ML2 and the individuals of the accession NM310 carried the marker associated with high grain protein content.

Genotyping known causative mutations in the *Vrs1* and *Int-c* genes, determining row-type, revealed all successfully genotyped samples to be monomorphic at Vrs1 A40F_S_ (deletion) and Vrs1 F75L (C). ML2, ML4, ML8, ML13, ML17 and NM310 all carried the *Vrs1* deletion (Vrs1 E152F_S) known to cause the six-row phenotype. These results are congruent with phenotyping as six-row barley, with exception for ML2 and NM310 which were phenotyped as two-row. In addition, a single individual from the accession NM288 also carried the deletion associated with a six-row phenotype. The Vrs1 E152F_S deletion genotype was completely correlated with the Int-c_SNP124 for which the genotype G co-occurred with the deletion. The genotyping thus suggests the remaining accessions to be two-row barley.

Analysis of genetic structure

In the PCA based on the different accessions (accession-level PCA), treating the single herbarium specimens as separate accessions, two- and six-row barley formed distinct clusters along the first principal component, with six-row accessions to the right and two-row barleys to the left (Figure 1). The accessions ML2 and NM310 carried the deletion in *Vrs1* known to cause the six-row phenotype, but clustered with the two-row barleys along the first principal component. ML2 and NM310 were instead separated from all other accessions along the second principal component (Figure 1).

Individual level PCA showed the same two-row or six-row division along the first principal component, with the second principal component dividing NM310 individuals (light green in Figure 2) and ML2 (purple in Figure 2) from all other two-row individuals. Most accessions showed some degree of clustering with other individuals from the same accession, though none as much as NM310. For the accession NM288 (light green in Figure 2) most individuals clustered among the two-row barleys, except the single individual carrying the *Vrs1* deletion which was instead located among the six-row barleys. A single individual of NGB321 (not genotyped for *Vrs1* or *Int-c*, blue in Figure 2) clustered among the two-row barleys, while the remaining individuals clustered among the six-row barleys.

Analysis of each row-type separated the extant NGB321 from all other six-row barleys, while the extant NGB27 did not deviate from the historical six-row barleys. Accession-level PCA showed a strong clustering according to specimen source (herbarium, extant or the different seed collections) among the six-row barleys (Figure 3A), and to a lesser extent among the two-row barley (Figure 3B). In an individual-level PCA, six-row herbarium barleys clustered apart from seed collection six-row barleys, while the herbarium two-row barley ML24 showed high similarity to seed collection two-row barleys (data not shown). The two-row accessions separated by the second principal component in Figure 1, ML2 and NM310, continued to cluster apart in PCA of two-row barleys only (data not shown). Only the individual-based two-row PCA showed evidence of geographic structure, with PC2 being significantly correlated with latitude after removal of the outlying accession NM279 (Pearson's product-moment correlation PC2 vs latitude $p < 0.05$, for all other comparisons $p > 0.05$). PCA for the six-row barley showed indications of geographic structure. In the accession-based PCA, PC1 was significantly correlated with longitude (Pearson's product-moment correlation $p < 0.05$ with NGB321 included and $p < 0.001$ with NGB321 excluded),

and PC2 was significantly correlated with latitude (Pearson's product-moment correlation $p < 0.01$ both with NGB321 included and excluded). In the individual-level PCA after the exclusion of NGB321, latitude of origin was significantly correlated with PC2 (Pearson's product-moment correlation $p < 0.001$).

In *STRUCTURE* analysis of all individuals, both ΔK and CLUMPP H values supported two clusters, with the observed division along the first principal component in the PCA corresponding to the division of two-row and six-row (Figure 4). In accession NM288 the individual clustering among the six-row individuals again clustered among the six-row barleys, and in NGB321 the single individual clustering among the two-row individuals again clustered among the two-row barleys. Apart from this all accessions clustered completely as either two-row or six-row barley.

Among the six-row barley ΔK and CLUMPP H values supported two as the number of clusters best describing the genetic diversity. At this level the herbarium samples and the majority of the accession MU52 formed a cluster together with individuals from NGB27, NM264, NGB321 and MU55 (supplementary file 1). Increasing the number of clusters further (with continued high ΔK and CLUMPP H values) resulted in one cluster for MU52 and NGB27, one for the herbarium samples, one for the remaining MU accessions, and mixed ancestry for NGB321 and NM264 (Figure 5). This resulted in geographic structuring with a northern and a southern cluster and a third cluster occurring primarily in the southeast.

For the two-row barley ΔK and CLUMPP H values suggested four or five as the number of clusters best describing the genetic diversity. At $K = 4$, NM310 formed one cluster. MU6 clustered completely in a second cluster together with the majority of NM297 and NM278, while a third cluster was comprised of NM294 and a major part of NM292 and NM279. NM291 formed a fourth cluster. The herbarium specimens ML2 and ML24 showed mixed ancestry and most accessions consisted of individuals belonging to different clusters and individuals with mixed ancestry. Overall, the two-row accessions showed no evidence of geographic structuring (Figure 6).

Discussion

Genetic analysis of archaeobotanical DNA

Recent advances in methods for DNA extraction and sequencing has allowed the field of aDNA genetics to flourish, especially with regard to studies of human remains, but also concerning archaeological specimens of other animal species (e.g. Slatkin & Racimo 2016; MacHugh et al. 2016). Successful DNA analysis of archaeological plant remains has proved more elusive. The rapid decomposition of most plant materials, in addition to the presence of PCR inhibitors, has meant specimens suitable for aDNA analysis have rarely been recovered (Gugerli et al. 2005; Brown et al. 2015). Studies of barley aDNA have been reported from desiccated samples found at Mediterranean and North African sites (Palmer et al. 2009; Mascher et al. 2016; Hagenblad et al. 2017). However, in Europe, including Fennoscandia, genetic studies of archaeological plant specimens are still largely absent, likely due to the fact that most archaeobotanical grains are preserved in a charred state. Although such specimens are abundant, successful aDNA analysis of charred seeds seems to be highly dependent on the extent of charring (Palmer et al. 2012; Nistelberger et al. 2016).

The DNA concentrations obtained from the charred specimens in this study ranged from 0.03 ng/µl to 3.22 ng/µl. However, *trnL* sequencing could not confirm the presence of barley DNA, and instead, returning BLAST hits from plant species likely to grow at the site suggested contaminating DNA of *in situ* origin. KASP genotyping has previously proved successful for genotyping of desiccated historical and ancient barley seeds (Lister et al. 2013; Hagenblad et al. 2017), but yielded no successful amplification in the charred barley seeds studied here. Previous attempts to genetically analyse charred archaeological barley remains from Kaupang, Norway, with next generation sequencing (NGS) has similarly met with no success (Nistelberger et al. 2016). In fact, Nistelberger et al. (2016) showed that preservation of DNA in charred grains is typically too low for utilizing the samples in aDNA studies, and suggested that earlier reports of endogenous DNA are likely to be false. Seeds preserved in desiccated or waterlogged conditions will be a more promising source of aDNA (Brown et al. 2015) but in Fennoscandia, such materials are scarce. Instances where desiccated materials in particular can be found will therefore be all the more valuable for aDNA studies. As archaeological sites in Europe typically do not contain desiccated grains, this type of material instead has to be searched for in buildings, graves or similar structures where grains have been kept dry and protected from predation.

Herbaria and other historical collections

Genetic analyses of grains from herbaria and seed collections, stored under dry conditions, have proved to be a valuable complement to the analysis of both extant plants and archaeobotanical remains in studies of agrarian history (Leino et al. 2009; Lister et al. 2008; Palmer et al. 2012). The specimens studied here, collected over a period of almost 40 years around the turn of the last century, had in general sufficient DNA quality to permit genetic analysis. Although some specimens had a poor genotyping success rate, the majority of samples had favourable genotyping results regardless of age, confirming the previously found amenability of dry-stored historical samples for genetic analysis. Our study also supports previous studies, finding KASP to be a cost-effective method for genetic analyses of aged DNA (Lister 2013, Hagenblad et al. 2017).

Simulation studies have shown that although 90 markers can be sufficient for detecting differentiation among regions using *STRUCTURE*, uneven sampling can have adverse effects on the detection of differentiation (Nelson & Anderson 2013). Increasing the number of sampled individuals has also been shown to increase the potential to detect population differentiation (Willing et al. 2012). When it comes to genetic studies of historical specimens, a large sample size will only be possible in rare instances, such as the seed collection of the Swedish Museum of Cultural History, where large number of grains from different individuals are available from each accession (Leino et al. 2009).

All of the historical accessions except NM288 contained a single row-type (either two-row or six-row). Characterisation of the *Vrs1* gene has identified three mutations that result in a six-row phenotype (Komatsuda et al. 2007). However, these three mutations will not by themselves explain all variation in row-type: six-row individuals may be wildtype for all three *Vrs1* mutations, and two-row individuals may carry deletions in *Vrs1* (Aslan et al. 2015). We found that the *Vrs1* and *Int-c* genotypes of the accession NM310 and the specimen ML2 indicated that they should be six-row barleys. However, in the STRUCTURE analysis and in PCA along the first principal component they clustered among two-row barleys. Phenotypic analysis of NM310 and ML2 confirmed that they had the two-row phenotype. Like Aslan et al. (2015), we conclude that *Vrs1* and *Int-c* genotyping is unable to reliably predict row-type, and that substantially divergent genotypes exist among two-row barley.

Genotyping of functional markers can provide clues regarding phenotypic traits which are not possible to discern from the appearance of seeds or rachides. For example, we can conclude that the majority of barley cultivated in Finland in the late 19th and early 20th century was late-flowering under increasing day-length, supporting previous findings from northern European barley (Jones et al. 2008a; Lister et al. 2009; Aslan et al. 2015).

Genetic structure in Finnish landrace barley

Our Finnish landrace barley samples confirmed previously shown strong genetic structuring between two-row and six-row barley (e.g. Malysheva-Otto et al. 2006, 2007; Kolodinska Brantestam et al. 2007; Yahiaoui et al. 2008; Leino & Hagenblad 2010). Forsberg et al. (2015), studying six-row barley, found a latitudinal geographic structure across Fennoscandia. Despite using less than a third of the markers used by Forsberg et al. (2015) we could repeat some of the results with significant geographic structure in PCA for six-row barley. In *STRUCTURE* analysis, more northern and southern accessions tended to cluster apart from each other as in Forsberg et al. (2015). The herbarium material formed a separate cluster that did not seem to correspond to any cluster detected in Forsberg et al. (2015). This might either reflect a more recent introduction of six-row barley (the herbarium specimens were mostly 20 years younger than the seed collection specimens) or a previously undetected older type. The herbarium samples are mostly from the south-west corner of Finland, a geographic region not covered by Forsberg et al. (2015).

For two-row barley, evidence of geographic structuring of genetic diversity was largely absent. However, the geographical area sampled for two-row barley was also much smaller than for six-row barley. Detection of geographical structure might also be obscured by the admixture of landraces with introduced cultivars (mass selections) of two-row barley from Central Europe. Such cultivars of two-row barley were commonly marketed in Fennoscandia in the late 19th century (Leino 2017). Although two-row barley has been cultivated in Finland for more than 300 years (Gadd 1770; Onnela 2004), barley cultivation in Finland was mainly based on four- and six-row landraces until the beginning of the 20th century. Since then two-row barley has become the favoured type for cultivation in Finland, not least for use in brewing (Heinonen 2009).

In spite of the absence of geographical structure, our PCA of two-row barley showed a clear separation between NM310 and ML2, and the remaining accessions and samples. Jones et al.

(2011) detected a similar separation among their two-row barley accessions, which seemed to primarily separate spring barley from winter barley. Neither of our historical samples had their growth habit recorded, but it seems likely that the separation along principal component two is the result of differences in growth habit. Future studies genotyping the *VRN* locus, controlling growth habit, with a larger number of accessions and individuals are needed to better evaluate the presence of genetic structuring in two-row Finnish barley.

Conclusions

We confirm previously reported difficulties in obtaining indigenous DNA from charred archaeological remains. Instead, we stress the importance of finding specimens preserved in waterlogged or preferably desiccated conditions. When such archaeological remains are not available, historical samples from herbaria or seed collections may fulfil a useful purpose. Based on such materials, Finnish six-row barley showed strong geographic clustering, likely due to climate adaption over long time. In contrast, in Finnish two-row barley – a younger and more rarely cultivated crop – genetic structuring does not seem to be linked to geography.

Acknowledgements

For Lempiäinen-Avci, aDNA analyses were granted by the Finnish Academy of Science and Letters / Jutikkala Trust and by Kone Foundation. We want to thank Dawn Elise Mooney for revising our English.

Conflict of interest

The authors declare that they have no conflict of interest on the content of manuscript and study undertaken.

505 **References**

506 Ahokas H, Poukkula M (1999) Malting enzyme activities, grain protein variation and yield
507 potentials in the displaced genetic resources of barley landraces of Finland. Genet Resour and
508 Crop Ev 46(3):251–260
509
510 Alenius T, Mökkönen T, Holmqvist E, Ojala A (2017) Neolithic land use in the northern
511 Boreal zone: high-resolution multiproxy analyses from Lake Huhdasjärvi, south-eastern
512 Finland. Veget Hist Archaeobot 26:469–486 https://doi.org/10.1007/s00334-017-0606-2
513
514 Aslan S, Forsberg NEG, Hagenblad J, Leino MW (2015) Molecular Genotyping of Historical
515 Barley Landraces Reveals Novel Candidate Regions for Local Adaption. Crop Sci 55:66–76
516 https://doi:10.2135/cropsci2015.02.0119
517
518 Asplund H (2008) Kymittæ. Sites, Centrality and Long-Term Settlement Change in the
519 Kemiönsaari Region in SW Finland. Annales Universitatis Turkuensis. Ser. B osa. Tom. 312.
520 Humaniora. Painosalama Oy, Turku. http://urn.fi/URN:ISBN:978-951-29-3628-1
521
522 Avise JC (2009) Phylogeography: retrospect and prospect. J Biogeogr 36:3–15
523 https://doi:10.1111/j.1365-2699.2008.02032.x
524
525 Behre K-E (1983) Ernährung und Umwelt der wikingerzeitlichen Siedlung Haithabu. Die
526 Ergebnisse der Untersuchungen der Pflanzenreste. Karl Wachholtz Verlag, Neumünster
527
528 Brown TA, Cappellini E, Kistler L, Lister DL, Oliveira HR, Wales N, Schlumbaum A (2015)
529 Recent advances in ancient DNA research and their implications for archaeobotany. Veget
530 Hist Archaeobot 24:207–214 https://doi.org/10.1007/s00334-014-0489-4
531
532 Bunning SL, Jones G, Brown TA (2012) Next generation sequencing of DNA in 3300-year-
533 old charred cereal grains. J Archaeol Sci 39:2780–2784
534 http://dx.doi.org/10.1016/j.jas.2012.04.012
535
536 von Bothmer R, van Hintum T, Knüpffer, H, Sato K (2003) Diversity in barley *(Hordeum*
537 *vulgare)* Developments in Plant Genetics and Breeding (Vol. 7). Elsevier
538

Cai S, Yu G, Chen X, Huang Y, Jiang X, Zhang G, Jin X (2013) Grain protein content variation and its association analysis in barley. BMC Plant Biology 13:35 https://doi.org/10.1186/1471-2229-13-35

Engelmark R (1992) A review of the farming economy in South Scania based on botanical evidence. In: Larsson L, Callmer J, Stjernquist B (eds) The archaeology of the cultural landscape. Field work and Research in a South Swedish Rural Region. Acta Archaeologica Lundensia Ser 4:19. Almqvist & Wiksell International, Stockholm, pp 369–374

Elfving F 1897. Anteckningar om Kulturväxterna i Finland. Acta Societas pro Fauna et Flora Fennica 14(2):1–116

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620 https://doi.org/10.1111/j.1365-294X.2005.02553.X

Forsberg NEG, Russell J, Macaulay M, Leino MW and Hagenblad J (2015) Farmers without borders—genetic structuring in century old barley (*Hordeum vulgare*). Heredity 114(2):195–206 https://doi.org/10.1038/hdy.2014.83

Gadd PA (1770) Försök och anmärkningar om utländske sädes-arter i Finska climatet. Första delen, med vederbörandes samtycke vid Kongl. Lärosätet i Åbo. Tryckt hos Johan Christoph Frenckell, Åbo. http://urn.fi/urn:nbn:fi:fv-11937

Grabowski R (2011) Changes in cereal cultivation during the Iron Age in southern Sweden: compilation and interpretation of the archaeobotanical material. Veget Hist Archaeobot 20:479–494 http://dx.doi.org/10.1007/s00334-011-0283-5

Grotenfelt G (1922) Suomalainen peltokasviviljelys. Jälkimmäinen osa. Kustannusosakeyhtiö Otavan kirjapaino, Helsinki

Gugerli F, Parducci L, Petit RJ (2005) Ancient plant DNA: Review and prospects. New Phytol 166:409–418 https://doi:org/10.1111/j.1469-8137.2005.01360.x

Hagenblad J, Morales J, Leino MW, Rodríguez-Rodríguez AC (2017) Farmer fidelity in the Canary Islands revealed by ancient DNA from prehistoric seeds. J Archaeol Sci 78:78–87 https://doi.org/10.1016/j.jas.2016.12.001

He C, Holme J, Anthony J (2014) SNP Genotyping: The KASP Assay. In: Fleury D, WhitfordR (eds) Crop Breeding: Methods and Protocols. Springer New York, New York, pp.75–86

Heinonen M (2009) Maatiaiskasvien ylläpitoviljely Suomessa. Maatalouden Tutkimuskeskus, Viestintä- ja informaatiopalvelut, Jokioinen

Herstad J (2000) I helstatens grep: kornmonopolet 1735–1788. Tano Aschehoug, Oslo

Igartua E, Moralejo M, Casas AM, Torres L, Molina-Cano JL (2013). Whole-genome analysis with SNPs from BOPA1 shows clearly defined groupings of Western Mediterranean, Ethiopian, and Fertile Crescent barleys. Genet Resour Crop Evol 60:251–264 http://dx.doi.org/10.1007/s10722-012-9831-9

Jacomet S, Kreuz A (1999) Archäobotanik. Aufgaben, Methoden und Ergebnisse vegetations- und agrargeschichtlicher Forschung. Verlag Eugen Ulmer, Stuttgart

Jacomet S (2006) Identification of cereal remains from archaeological sites. 2nd edition. Archaeobotany Lab IPAS, Basel University

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806 https://doi.org/10.1093/bioinformatics/btm233

Jones H, Leigh FJ, Mackay I, Bower MA, Smith LM, Charles MP, Jones G, Jones MK, Brown TA, Powell W (2008a) Population-based resequencing reveals that the flowering time adaptation of cultivated barley originated east of the Fertile Crescent. Mol Biol Evol 25:2211–2219 https://doi.org/10.1093/molbev/msn167

606 Jones H, Lister DL, Bower MA, Leigh FJ (2008b) Approached and constraints of using
607 existing landrace and extant plant material to understand agricultural spread in prehistory.
608 Plant Genetic Resources. 6(2):98-112 https://doi.org/10.1017/S1479262108993138
609
610 Jones H, Civan P, Cockram J, Leigh FJ, Smith LM, Jones MK, Charles MP, Molina-Cano JL,
611 Powell W, Jones G, Brown TA (2011) Evolutionary history of barley cultivation in Europe
612 revealed by genetic analysis of extant landraces. BMC Evol Biol 11:320
613 https://doi.org/10.1186/1471-2148-11-320
614
615 Jones G, Jones H, Charles MP, Jones MK, Colledge S, Leigh FJ, Lister DA, Smith LMJ,
616 Powell W, Brown TA (2012) Phylogeographic analysis of barley DNA as evidence for the
617 spread of Neolithic agriculture through Europe. J Archaeol Sci 39: 3230-3238
618 http://dx.doi.org/10.1016/j.jas.2012.05.014
619
620 Jones G, Charles M, Jones M, Colledge S, Leigh F, Lister D, Smith L, Powell W, Brown, TA
621 and Jones H (2013) Barley DNA evidence for the routes of agricultural spread into Europe
622 following multiple domestications in W. Asia. Antiquity 87: 701-13.
623
624 Kirleis W, Klooß S, Kroll H, Müller J (2012) Crop growing and gathering in the northern
625 German Neolithic: a review supplemented by new results. Veget Hist Archaeobot 21:221–242
626 https://doi.org/10.1007/s00334-011-0328-9
627
628 Kolodinska Brantestam A, von Bothmer R, Dayteg C, Rashal I, Tuvesson S, Weibull J (2007)
629 Genetic diversity changes and relationships in spring barley (*Hordeum vulgare* L.) germplasm
630 of Nordic and Baltic areas as shown by SSR markers. Genet Resour Crop Evol 54:749–758
631 https://doi.org/10.1007/s10722-006-9159-4
632
633 Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D, Stein N,
634 Graner A, Wicker T, Tagiri A, Lundqvist U, Fujimura T, Matsuoka M, Matsumoto T, Yano
635 M (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-
636 class homeobox gene. Proc Natl Acad Sci U S A 104(4):1424–1429
637 https://dx.doi.org/10.1073/pnas.0608580104
638

Kota R, Varshney R, Prasad M, Zhang H, Stein N, Graner A (2008) EST-derived single nucleotide polymorphism markers for assembling genetic and physical maps of the barley genome. Funct Integr Genomic 8:223–233 https://doi.org/10.1007/s10142-007-0060-9

Leino MW, Hagenblad J, Edqvist J, Strese EMK (2009) DNA preservation and utility of ahistoric seed collection. Seed Sci Res 19:125–135 https://doi.org/10.1017/S0960258509990055

Leino MW (2010). Frösamlingar på museum. Nord Mus 2010: 96–108 http://dx.doi.org/10.5617/nm.3183

Leino MW, Hagenblad J (2010) Nineteenth century seeds reveal the population genetics oflandrace barley (*Hordeum vulgare*). Mol Biol Evol 27(4):964–973 https://doi.org/10.1093/molbev/msp308

Leino MW, Bostrom E, Hagenblad J (2013) Twentieth-century changes in the genetic composition of Swedish field pea metapopulations. Heredity (Edinb) 110(4):338–346 https://doi.org.10.1038/hdy.2012.93

Leino MW (2017). Spannmål: Svenska lantsorter. Nordiska museets förlag, Stockholm.

Lempiäinen-Avci M, Haggrén G, Rosendahl U, Knuutinen T, Holappa M (2017) Archaeobotanical analysis of radiocarbon-dated plant remains with special attention to rye (*Secale cereale*) cultivation at the medieval village of Mankby in Espoo (Finland). Veget Hist Archaeobot 26(4):435–446 https://doi.org/10.1007/s00334-017-0604-4

Lindström G (1905) Suomen kaupasta Aleksanteri 1. aikana. I Järjestymisvuodet 1808–1808. Osakeyhtiö Weilin & Göös, Helsinki.

Lister DL, Bower MA, Howe CJ, Jones MK (2008) Extraction and amplification of nuclear DNA from herbarium specimens of emmer wheat: a method for assessing DNA preservation by maximum amplicon length recovery. Taxon 57(1):254–258

672 Lister DL, Thaw S, Bower MA, Jones H, Charles MP, Jones G, Smith LMJ, Howe CJ, Brown
673 TA, Jones MK (2009) Latitudinal variation in a photoperiod response gene in European
674 barley: insight into the dynamics of agricultural spread from 'historic' specimens. J Archaeol
675 Sci 36(4):1092–1098 https://doi.org/10.1016/j.jas.2008.12.012
676
677 Lister DL, Jones MK (2012) Is naked barley an eastern or a western crop? The combined
678 evidence of archaeobotany and genetics. Veget Hist Archaeobot: 22:439–46
679 https://doi.org/10.1007/s00334-012-0376-9
680
681 Lister DL, Jones H, Jones MK, O'Sullivan DM, Cockram J (2013) Analysis of DNA
682 polymorphism in ancient barley herbarium material: Validation of the KASP SNP genotyping
683 platform. Taxon 62(4):779–789 https://doi.org/10.12705/624.9
684
685 Lunden K (2004) Recession and new Expansion. In: Almås R (ed) Norwegian Agricultural
686 History. Tapir Academic Press, Trondheim, pp 141–232
687
688 MacHugh DE, Larson G, Orlando L (2016) Taming the Past: Ancient DNA and the Study of
689 Animal Domestication. Annu Rev Anim Biosci 5:329–351 https://doi.org/10.1146/annurev-
690 animal-022516-022747
691
692 Malysheva-Otto LV, Ganal MV, Röder MS (2006) Analysis of molecular diversity,
693 population structure and linkage disequilibrium in a worldwide survey of cultivated barley
694 germplasm (*Hordeum vulgare* L.). BMC Genet 7:6 https://doi.org/10.1186/1471-2156-7-6
695
696 Malysheva-Otto LV, Ganal MV, Law JR, Reeves JC, Röder MS (2007) Temporal trends of
697 genetic diversity in European barley cultivars (*Hordeum vulgare* L.). Mol Breed 20(4):309–
698 322 https://doi.org/10.1007/s11032-007-9093-y
699
700 Morrell PL, Clegg MT (2007) Genetic evidence for a second domestication of barley
701 (*Hordeum vulgare*) east of the Fertile Crescent. Proc Natl Acad Sci U S A 104(9):3289–3294
702 https://dx.doi.org/10.1073/pnas.0611377104
703
704 Mascher M, Schuenemann VJ, Davidovich U, Marom N, Himmelbach A, Hubner S, Korol A,
705 David M, Reiter E, Riehl S (2016) Genomic analysis of 6,000-year-old cultivated grain

illuminates the domestication history of barley. Nat Genet 48(9):1089–1093 https://doi.10.1038/ng.3611

Nelson MF, Anderson NO (2013) How many marker loci are necessary? Analysis of dominant marker data sets using two popular population genetic algorithms. Ecol Evol 3(10):3455–3470 https://dx.doi.org/10.1002/ece3.725

Nistelberger HM, Smith O, Wales N, Star B, Boessenkool S (2016) The efficacy of high-throughput sequencing and target enrichment on charred archaeobotanical remains, Sci Rep 6:37347 https://doi.org/10.1038/srep37347

Onnela J (2004) Vanhakantaisia viljelykasveja Etelä-Suomessa. Kasvijäännetutkimuksia ja kokeellista arkeobotaniikkaa. An unpublished licentiate thesis, University of Turku, Department of Biology.

Palmer SA, Moore JD, Clapham AJ, Rose P, Allaby RG (2009) Archaeogenetic Evidence of Ancient Nubian Barley Evolution from Six to Two-Row Indicates Local Adaptation. PLoS ONE 4(7):e6301 https://doi.org/10.1371/journal.pone.0006301

Palmer SA, Smith O, Allaby RG (2012) The blossoming of plant archaeogenetics. Annals of anatomy - Anatomischer Anzeiger : official organ of the Anatomische Gesellschaft 194 (1):146–156 https://doi.org/10.1016/j.aanat.2011.03.012

Papa R, Attene G, Barcaccia G, Ohgata A, Konishi T (1998). Genetic diversity in landrace populations of *Hordeum vulgare* L. from Sardinia, Italy, as revealed by RAPDs, isozymes and morphophenological traits. Plant Breed 117:523–530 https://doi.org/10.1111/j.1439-0523.1998.tb02201.x

Pihlman S, Seppä-Heikka M (1985). Indication of Late-Neolithic cereal cultivation at the Kotirinne dwelling site at Niuskala, Turku, SW Finland. Memoranda Societatis pro Fauna et Flora Fennica 61:85–88

Ramsay L, Comadran J, Druka A, Marshall DF, Thomas WT, Macaulay M, MacKenzie K, Simpson C, Fuller J, Bonar N, Hayes PM, Lundqvist U, Franckowiak JD, Close TJ,

740 Muehlbauer GJ, Waugh R (2011) INTERMEDIUM-C, a modifier of lateral spikelet fertility
741 in barley, is an ortholog of the maize domestication gene TEOSINTE BRANCHED 1. Nat
742 Genet 43:169–172 https://doi.org/10.1038/ng.745
743
744 Raninen S (2013) Arkeologinen kaivaus Pirkkalan Tursiannotkon rautakautisella
745 asuinpaikalla. Pirkan maan alta 13:6–17
746 http://issuu.com/vapriikki/docs/pirkanmaanalta_13?e=4699319/6432777
747
748 Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population
749 structure. Mol Ecol Notes 4:137–138 https://doi.org/10.1046/j.1471-8286.2003.00566.x
750
751 Roullier C, Benoit L, McKey DB, Lebot V (2013) Historical collections reveal patterns of
752 diffusion of sweet potato in Oceania obscured by modern plant movements and
753 recombination. Proc Natl Acad Sci USA 110: 2205–2210
754 https://doi.org/10.1073/pnas.1211049110
755
756 Sauli JO (1927) Suomen maatiaisohrat ja niiden jalostusarvo. Suomen maataloustieteellisen
757 seuran julkaisuja 161–139
758
759 Semagn K, Babu R, Hearne S, Olsen M (2014) Single nucleotide polymorphism genotyping
760 using Kompetitive Allele Specific PCR (KASP): overview of the technology and its
761 application in crop improvement. Mol Breed 33:1–14 https://doi.org/10.1007/s11032-013-
762 9917-x
763
764 Slatkin M, Racimo F (2016) Ancient DNA and human history. Proc Natl Acad Sci U S A
765 113(23):6380–6387 https://doi.org/10.1073/pnas.1524306113
766
767 Soreng RJ (2003) *Hordeum*. 48: 389–402. In Soreng RJ, Peterson PM, Davidse G, Judziewicz
768 EJ, Zuloaga FO, Filgueiras TS, Morrone O (eds) Catalogue of New World Grasses (Poaceae):
769 IV. Subfamily Pooideae, Contr. U.S. Natl. Herb. Smithsonian Institution, Washington, D.C.
770
771
772
773

774 Stika H-P, Heiss AG (2013) Plant Cultivation in the Bronze Age. In: Fokkens H, Harding A
775 (eds) The Oxford Handbook of the European Bronze Age. Oxfords Handbooks Online,
776 Oxford University Press, Oxford, pp 340–361
777 https://doi.org/10.1093/oxfordhb/9780199572861.001.0001
778
779 Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G,
780 Brochmann C, Willerslev E (2007) Power and limitations of the chloroplast trnL (UAA)
781 intron for plant DNA barcoding. Nucleic Acids Research 35(3):e14
782 https://doi.org/10.1093/nar/gkl938
783
784 Taivainen J (2004) Rautakaudelta keskiajalle - Retulansaaren metsistä, pelloilta ja kylämäeltä.
785 Muinaistutkija 3:7–17
786
787 Taivainen J (2007) Late Iron Age and Medieval Crafts in Rural Environment – with special
788 reference to the Retula Village in Southern Finland. Ruralia 6:179–187
789
790 Vuorela I, Lempiäinen T (1988) Archaeobotany of the site of the oldest cereal grain find in
791 Finland. Annales Botanici Fennici 25:33–45
792
793 Vuorinen J-M (2009) Rakennukset ja rakentajat Raision Ihalassa rautakauden lopulla ja
794 varhaisella keskiajalla. Annales Universitatis Turkuensis. Sarja C osa 281. Scripta Lingua
795 Fennica Edita. Turun yliopisto, Turku.
796
797 Wang Y, Ren X, Sun D, Sun G (2016) Molecular evidence of RNA polymerase II gene
798 reveals the origin of worldwide cultivated barley. Sci Rep 6:36122
799 https://doi.org/10.1038/srep36122
800
801 Willerslev E, Cappellini E, Boomsma W, Nielsen R, Hebsgaard MB, Brand TB, Johnsen S
802 (2007) Ancient biomolecules from deep ice cores reveal a forested southern Greenland.
803 Science 317(5834):111–114 https://doi.org/10.1126/science.1141758
804
805 Willing EM, Dreyer C, van Oosterhout C (2012) Estimates of genetic differentiation
806 measured by F_{ST} do not necessarily require large sample sizes when using many SNP
807 markers. PLoS ONE 7(8):e42649 https://doi:10.1371/journal.pone.0042649

Xia Y, Ning Z, Bai G, Li R, Yan G, Siddique K, Baum M, Guo P (2012) Allelic variations of a light harvesting chlorophyll a/b-binding protein gene (Lhcb1) associated with agronomic traits in barley. PLoS ONE 7(5):e37573 https://doi.org/10.1371/journal.pone.0037573

Yahiaoui S, Igartua E, Moralejo M, Ramsay L, Molina-Cano JL, Ciudad FJ, Lasa JM, Gracia MP, Casas AM (2008) Patterns of genetic and eco-geographical diversity in Spanish barleys. Theor Appl Genet 116(2):271–282 https://doi.org 10.1007/s00122-007-0665-3

Zohary D, Hopf M, Weiss E (2012) Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin. Oxford University Press, Oxford.

Åström S–E (1980) Varhaisteollinen aikakausi. In Jutikkala E, Kaukiainen Y, Åström S–E (eds) Suomen taloushistoria 1. Agraarinen Suomi. Kustannusosakeyhtiö Tammi, Helsinki, pp. 331-494

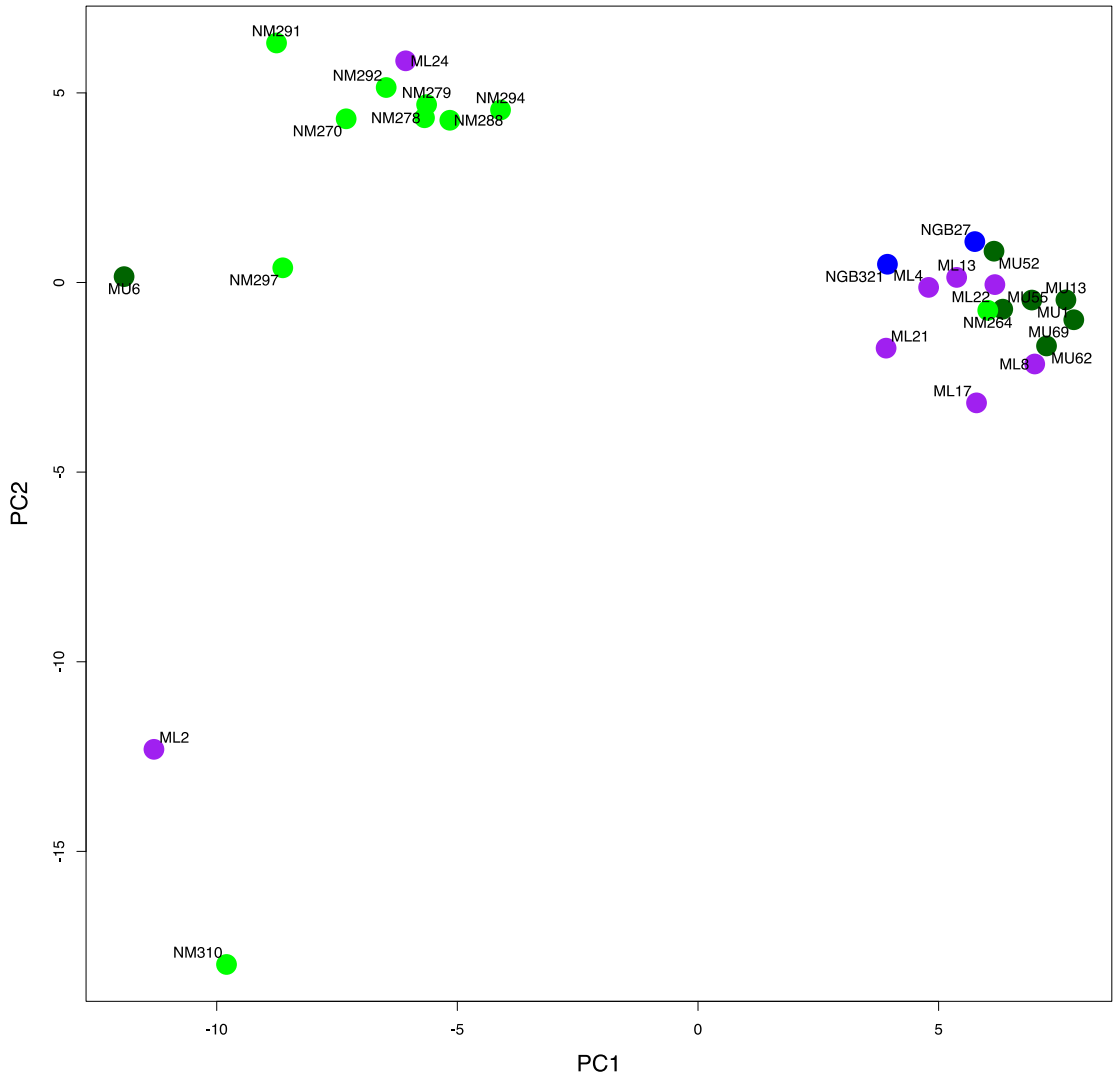


Fig 1. Principal component analysis of allele frequencies for 90 loci from the BOPA1 array, with accessions and individual herbarium specimens treated as separate populations. Each data point represents a separate accession identified by accompanying accession number. Dark green denotes accessions from Mustiala Agricultural College, blue denotes accessions from the Nordic Genetic Resource Center, light green denotes accessions from the Swedish museum of cultural history and purple denotes herbarium specimens. The first principal component explains 27.59 % of the genetic variation and the second principal component 14.53 %.

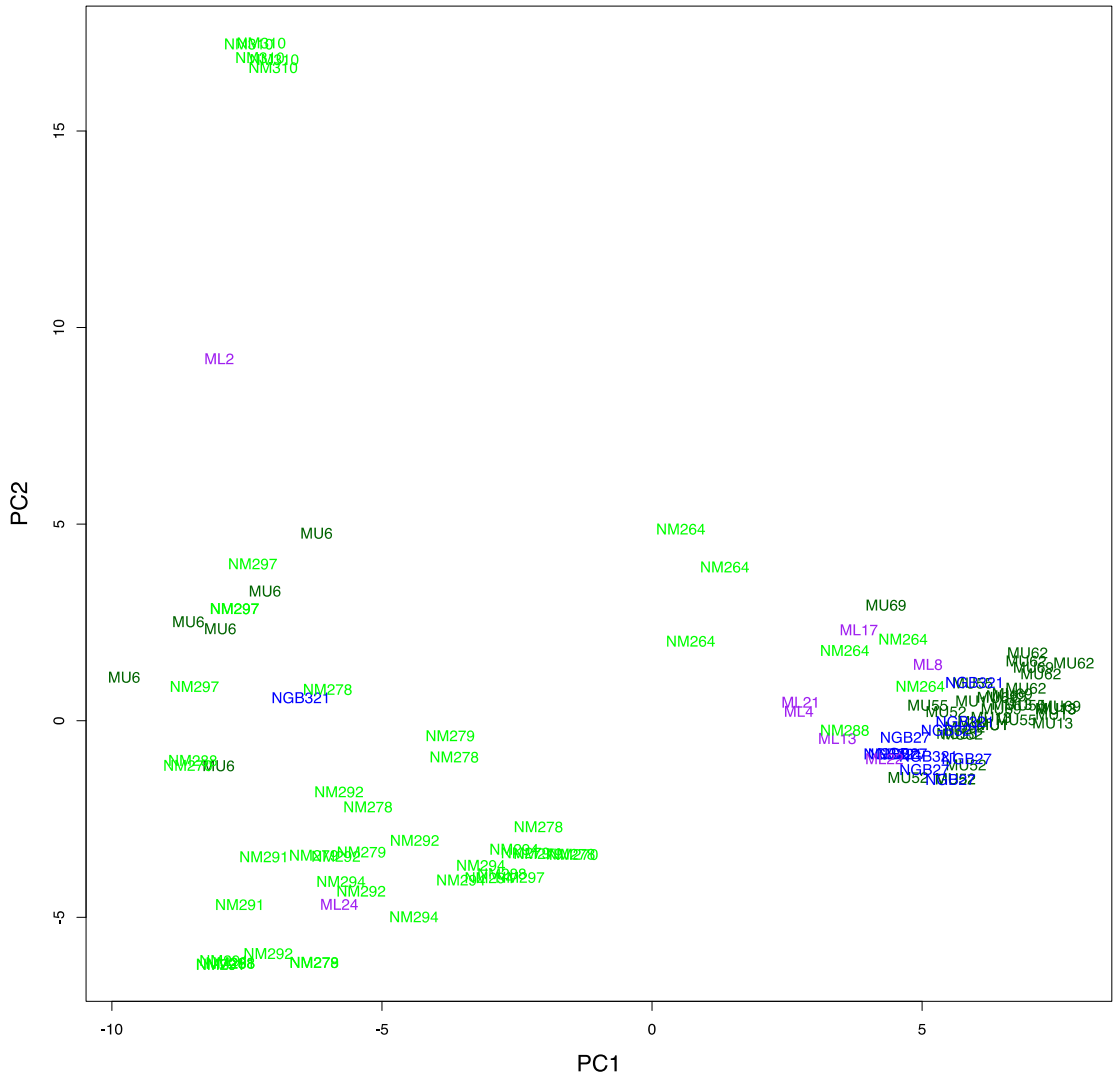


Fig 2. Principal component analysis of the genotypes of the individual specimens for 90 loci from the BOPA1 array, with the accession number given for each individual. Dark green denotes accessions from Mustiala Agricultural College, blue denotes accessions from the Nordic Genetic Resource Center, light green denotes accessions from the Swedish museum of cultural history and purple denotes herbarium specimens. The first principal component explains 19.50 % of the genetic variation and the second principal component 11.97 %.

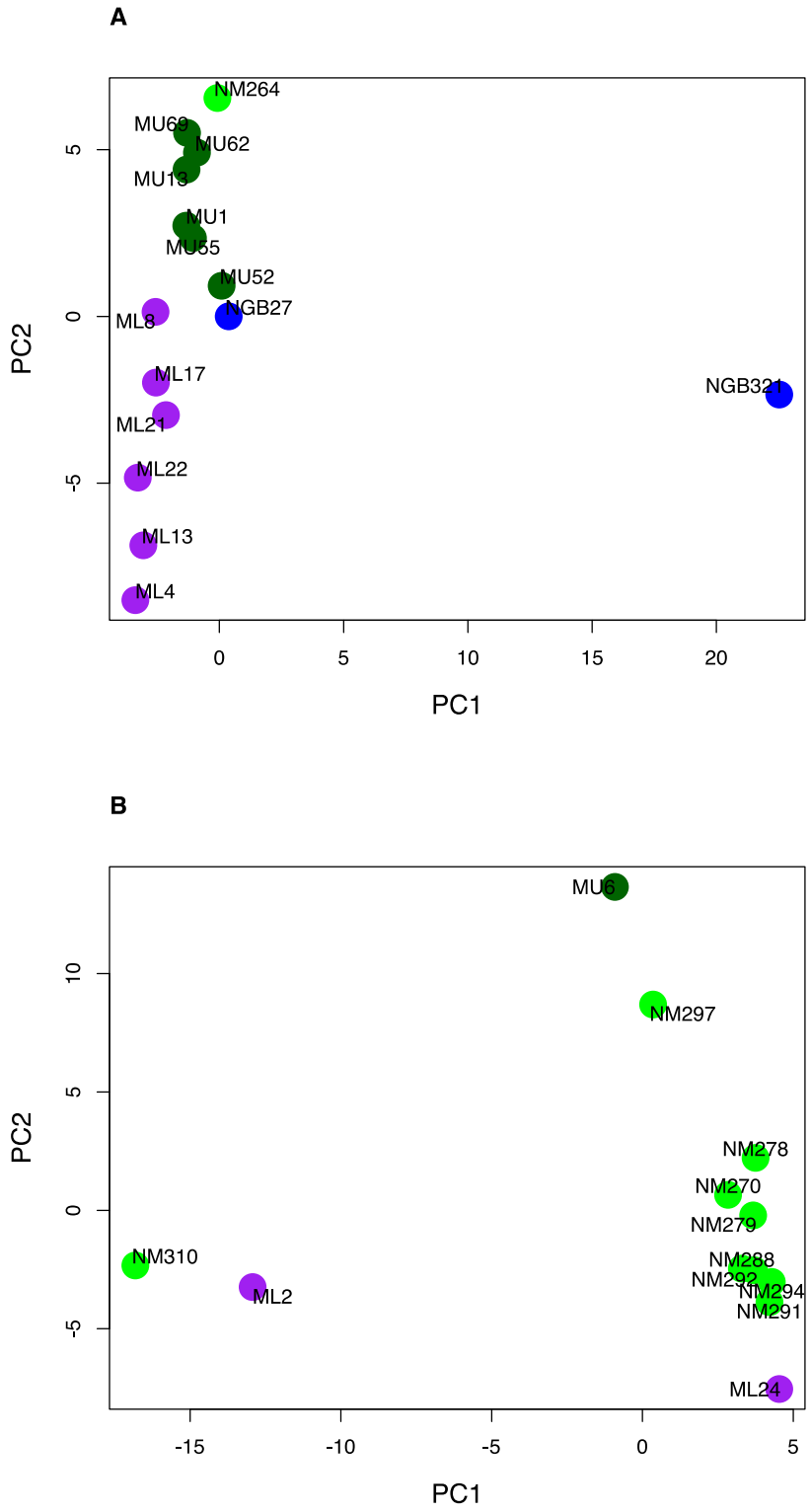


Fig 3. Principal component analysis of allele frequencies for 90 loci from the BOPA1 array, with accessions and individual herbarium specimens treated as separate populations. Dark green denotes accessions from Mustiala Agricultural College, blue denotes accessions from the Nordic Genetic Resource Center, light green denotes accessions from the Swedish museum of cultural history and purple denotes herbarium specimens. A) Six-row barley accessions. The first principal component explains 28.79 % of the genetic variation and the second principal component 14.85 %. B) Two-row barley accessions. The first principal

component explains 28.84 % of the genetic variation and the second principal component 19.15 %.

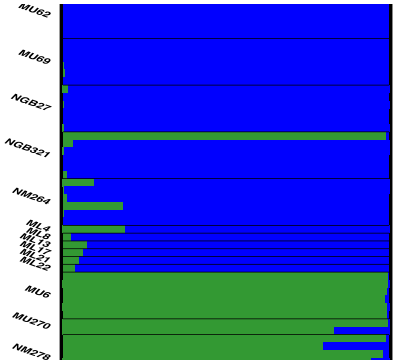


Fig 4. Results of structure analysis of all accessions, regardless of row-type. Each vertical line corresponds to one individual where the proportion of each colour corresponds to the extent to which the individual has been designated to the particular cluster represented by that colour. Six-row barley accessions cluster predominantly in the blue group and two-row accessions in the green group.

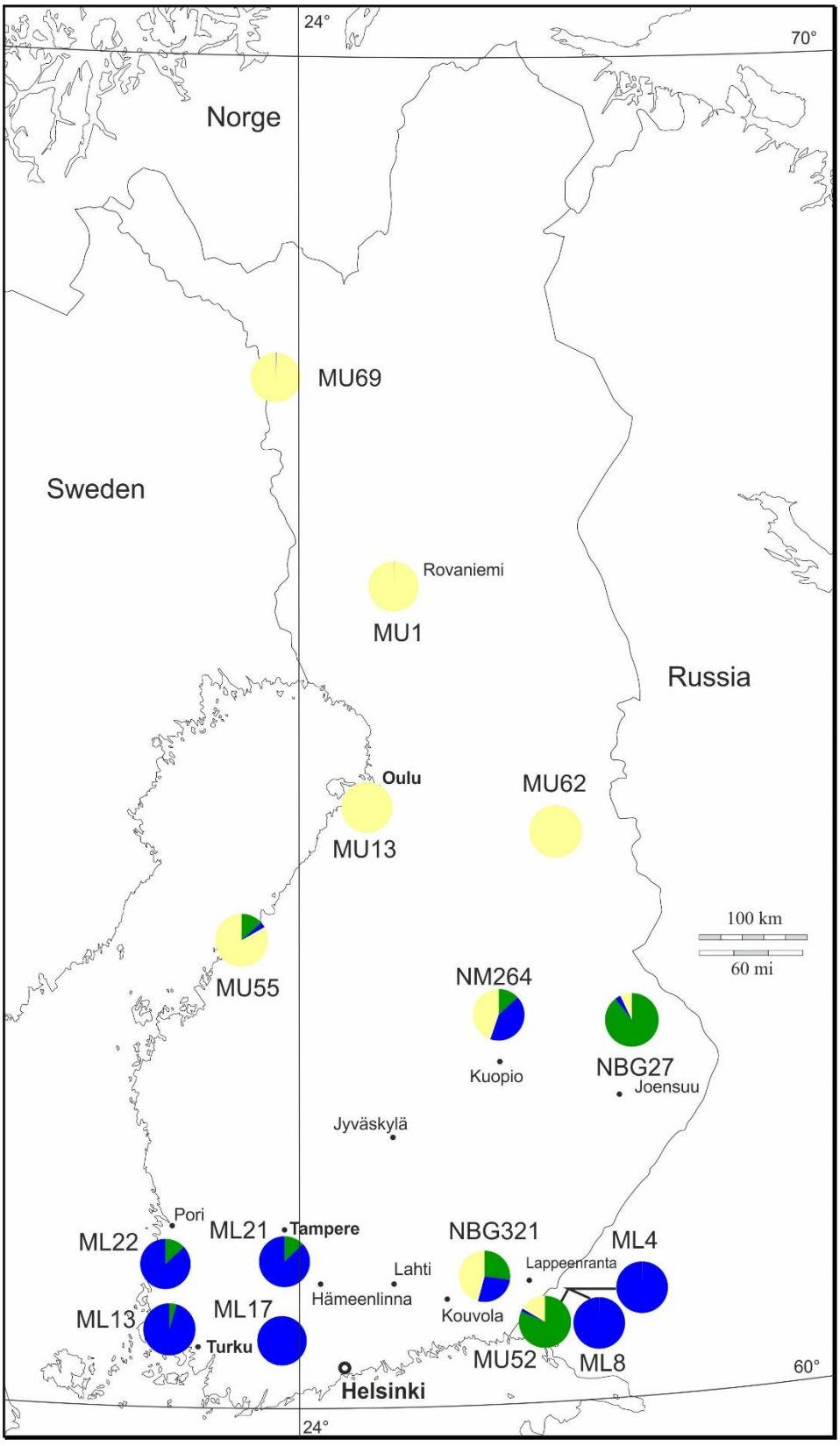


Fig 5. Results of structure analysis of all six-row individuals, assigned to three clusters, represented by different colours in the figure. The proportion of each colour corresponds to the extent to which the individuals from each accession have been designated to the particular cluster represented by that colour.

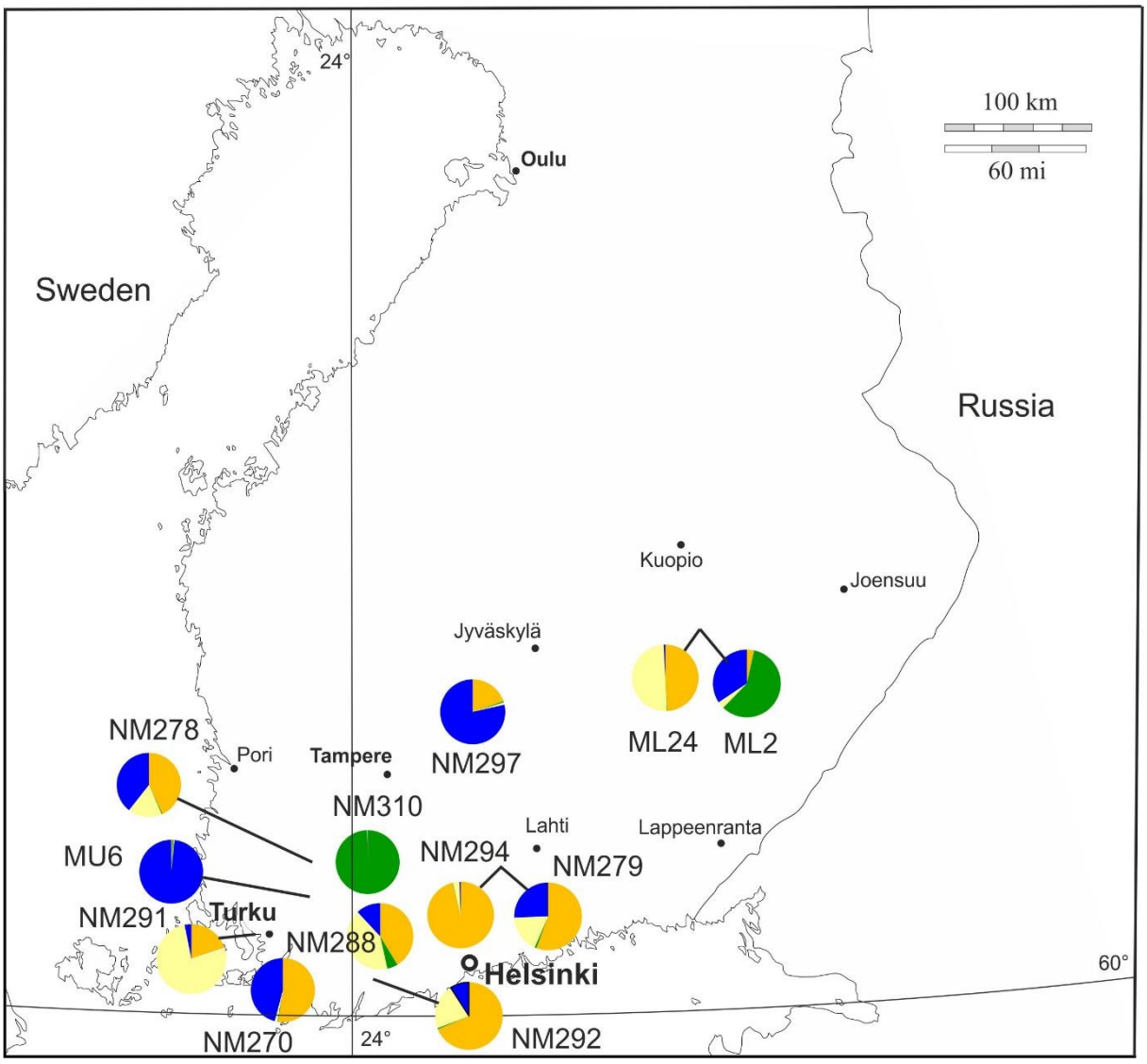


Fig 6. Results of structure analysis of all two-row individuals, assigned to four clusters, represented by different colours in the figure. The proportion of each colour corresponds to the extent to which the individuals from each accession have been designated to the particular cluster represented by that colour.

Table 1. Information about archaeological samples studied

Designation	Location	Latitude	Longitude	Age	DNA concentration (ng/µl)	PCR products
HV10_RM_1	Raisio Mulli	60° 28' 13" N	22° 11' 31" E	AD 980 - 1220	0.53	Yes*
HV10_RM_2	Raisio Mulli	60° 28' 13" N	22° 11' 31" E	AD 980 - 1220	0.30	Faint
HV10_RM_3	Raisio Mulli	60° 28' 13" N	22° 11' 31" E	AD 980 - 1220	0.14	No
HV10_RM_4	Raisio Mulli	60° 28' 13" N	22° 11' 31" E	AD 980 - 1220	0.35	Faint
HV10_RM_5	Raisio Mulli	60° 28' 13" N	22° 11' 31" E	AD 980 - 1220	0.41	No
HV10_RM_6	Raisio Mulli	60° 28' 13" N	22° 11' 31" E	AD 980 - 1220	0.03	Yes*
HV10_RM_7	Raisio Mulli	60° 28' 13" N	22° 11' 31" E	AD 980 - 1220	0.62	No
HV10_RM_8	Raisio Mulli	60° 28' 13" N	22° 11' 31" E	AD 980 - 1220	0.03	Yes*
HV10_RM_9	Raisio Mulli	60° 28' 13" N	22° 11' 31" E	AD 980 - 1220	0.03	No
HV10_RM_10	Raisio Mulli	60° 28' 13" N	22° 11' 31" E	AD 980 - 1220	0.28	Faint
HV24_EM_1	Espoo Mankby	60° 28' 13" N	22° 11' 31" E	AD 1500	0.20	No
HV24_EM_2	Espoo Mankby	60° 28' 13" N	22° 11' 31" E	AD 1500	0.20	Faint
HV24_PT	Pirkkala Tursiannotko	61° 26' 59" N	23° 32' 53" E	AD 800-1000	1.23	Faint
HV23_PT	Pirkkala Tursiannotko	61° 26' 59" N	23° 32' 53 " E	AD 800-1000	2.51	Double bands

HV22_PT_1	Pirkkala Tursiannotko	61° 26' 59" N	23° 32' 53" E	AD 800-1000	1.87	Faint double bands
HV22_PT_2	Pirkkala Tursiannotko	61° 26' 59" N	23° 32' 53" E	AD 800-1000	0.17	No
HV22_PT_3	Pirkkala Tursiannotko	61° 26' 59" N	23° 32' 53" E	AD 800-1000	0.10	No
HV22_PT_4	Pirkkala Tursiannotko	61° 26' 59" N	23° 32' 53" E	AD 800-1000	0.66	Yes*
HV20_PT_1	Pirkkala Tursiannotko	61° 26' 59" N	23° 32' 53" E	AD 800-1000	1.62	Yes*
HV20_PT_2	Pirkkala Tursiannotko	61° 26' 59" N	23° 32' 53" E	AD 800-1000	3.22	Yes*
HV08_HRY_1	Hattula Retulansaari Ylikartano	61° 10' 26" N	24° 19' 11"	AD 700-900	0.48	Yes*
HV08_HRY_2	Hattula Retulansaari Ylikartano	61° 10' 26" N	24° 19' 11"	AD 700-900	0.35	No
HV08_HRY_3	Hattula Retulansaari Ylikartano	61° 10' 26" N	24° 19' 11"	AD 700-900	0.24	Yes*
HV08_HRY_4	Hattula Retulansaari Ylikartano	61° 10' 26" N	24° 19' 11"	AD 700-900	0.16	Yes*

HV08_HRY_5	Hattula Retulansaari Ylikartano	61° 10' 26" N	24° 19' 11" E	AD 700-900	0.88	No
HV08_HRY_6	Hattula Retulansaari Ylikartano	61° 10' 26" N	24° 19' 11" E	AD 700-900	0.11	No
HV08_HRY_7	Hattula Retulansaari Ylikartano	61° 10' 26" N	24° 19' 11" E	AD 700-900	0.87	Yes*

* PCR product used for sequencing

Table 2. Information about historical samples studied

Designation	Source	Location	Latitude	Longitude	Age	Row type	No individuals	Average success rate (variance) ***
ML1	Herbarium http://mus.utu.fi/TFA.146792	Karelia Ladogensis, Soanlahti, Valjakka *	62° 1' 60" N	31° 3' 0"E	1902	Six-row	1	0
ML2	Herbarium http://mus.utu.fi/TFA.146784	Savonia borealis, Kangaslampi	62° 17' 27" N	28° 15' 3" E	1893	Two-row	1	0.959
ML4	Herbarium http://mus.utu.fi/TFA.146783	Karelia Ladogensis, Jääski, Hyppölä *	61° 2' 00" N	28° 56' 00"E	1913	Four-row	1	0.990
ML8	Herbarium http://mus.utu.fi/TFA.146782	Åland, Lumparland	60° 7' 3" N	20° 15' 39" E	1906	Four-row	1	1
ML13	Herbarium http://mus.utu.fi/TFA.379943	Reso	60° 29' 16" N	22° 10' 49" E	1912	Four-row	1	1
ML16	Herbarium http://mus.utu.fi/TFA.420295	Gamla Karleby	63° 50' 16" N	23° 7' 42" E	1913	Four-row	1	0.113
ML17	Herbarium http://mus.utu.fi/TFA.420294	Somero	60° 37' 3" N	23° 30' 47" E	1920	Four-row	1	0.979
ML18	Herbarium http://mus.utu.fi/TFA.420293	Tavastehus	60° 58' 44" N	24° 30' 12" E	1872	Six-row	1	0
ML19	Herbarium http://mus.utu.fi/TFA.420292	Kangasala	61° 27' 48" N	24° 4' 12" E	1917	Six-row	1	0

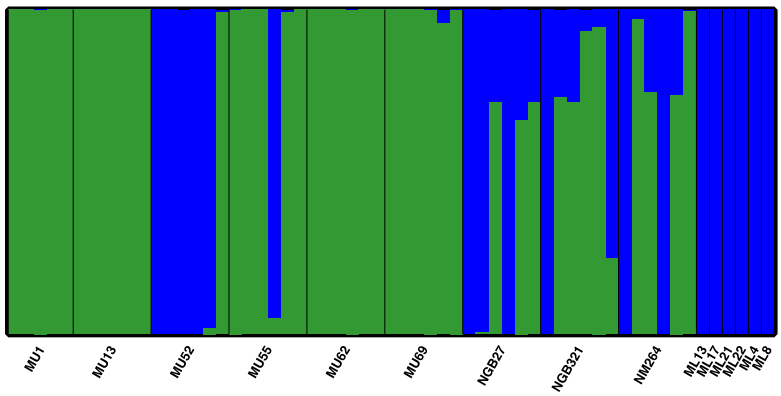
ML20	Herbarium http://mus.utu.fi/TFA.420291	Hausjärvi, Torhola	60° 48' 5" N	25° 3' 49" E	1918	Six-row	1	0.371
ML21	Herbarium http://mus.utu.fi/TFA.420290	Tavastland, Akkas, Toijala	61° 10' 24" N	23° 50' 21" E	1915	Six-row	1	0.918
ML22	Herbarium http://mus.utu.fi/TFA.420289	Satakunta, Wuojoki	61° 13' 7" N	21° 40' 6" E	1896	Four-row	1	1
ML23	Herbarium http://mus.utu.fi/TFA.420288	Borgå, Emsalo	60° 17' 48" N	25° 36' 8" E	1921	Six-row	1	0
ML24	Herbarium http://mus.utu.fi/TFA.420287	Mäntsälä, Nikinoja	60° 48' 29" N	25° 11' 54" E	1913	Two-row	1	0.979
NM264**	Seed collection	Mattila	63° 15’N	27° 28’E	1882	Six-row	6	
NM270	Seed collection	Piikis	60° 16' N	22° 18' 36" E	1882	Two-row	2	
NM278	Seed collection	Kemppi, Alastaro	60° 57' 7" N	22° 51' 42" E	1882	Two-row	6	
NM279	Seed collection	Laitila, Kärkölä	60° 52' 7" N	25° 16' 38" E	1882	Two-row	6	
NM288	Seed collection	Ingeris, St Bertils	60° 26' 59" N	23° 14' 16" E	1882		5	0.919 (0.030
NM291	Seed collection	Artukais, Reso	60° 29' 9" N	22° 10' 8" E	1882	Two-row	6	0.986 (0.000)
NM292	Seed collection	Mättilä, Kisko	60° 15' 53" N	23° 26' 49" E	1882	Two-row	5	0.902 (0.022)
NM294	Seed collection	Laitiala, Kärkölä	60° 52' 7" N	25° 16' 38" E	1882		6	0.988 (0.000)
NM297	Seed collection	Mattila, Längelmäki	61° 43' 45" N	24° 48' 0" E	1882		5	0.814 (0.160)
NM310	Seed collection	Mustiala	60° 49' N	23° 46' 11" E	1882	Two-row	5	0.957 (0.000)
MU1**	Seed collection	Rovaniemi	66° 28' 48" N	25° 43' 12" E	1890s	Six-row	5	
MU6	Seed collection	Koski	60° 39' 7" N	23° 8' 18" E	1890s	Two-row	6	
MU13**	Seed collection	Oulunsalo	64° 55' 48" N	25° 24' E	1890s	Six-row	6	
MU52**	Seed collection	Jääski	61° 1' 48" N	28° 55' 12" E	1890s	Six-row	6	
MU55**	Seed collection	Vielvis, Kelviå	63° 51' N	23° 27' E	1890s	Six-row	6	
MU62**	Seed collection	Hyrynsalmi	64° 43' 48" N	28° 28' 12" E	1890s	Six-row	6	
MU69**	Seed collection	Muonionniska	67° 57' N	23° 39' E	1890s	Six-row	6	

NGB27**	Extant	Sarkalahti	61° 1' 48" N	27° 19' 48" E		Six-row	6	
NGB321**	Extant	Törmälä	63° 10' 48" N	30° 1' 12" E		Six-row	6	

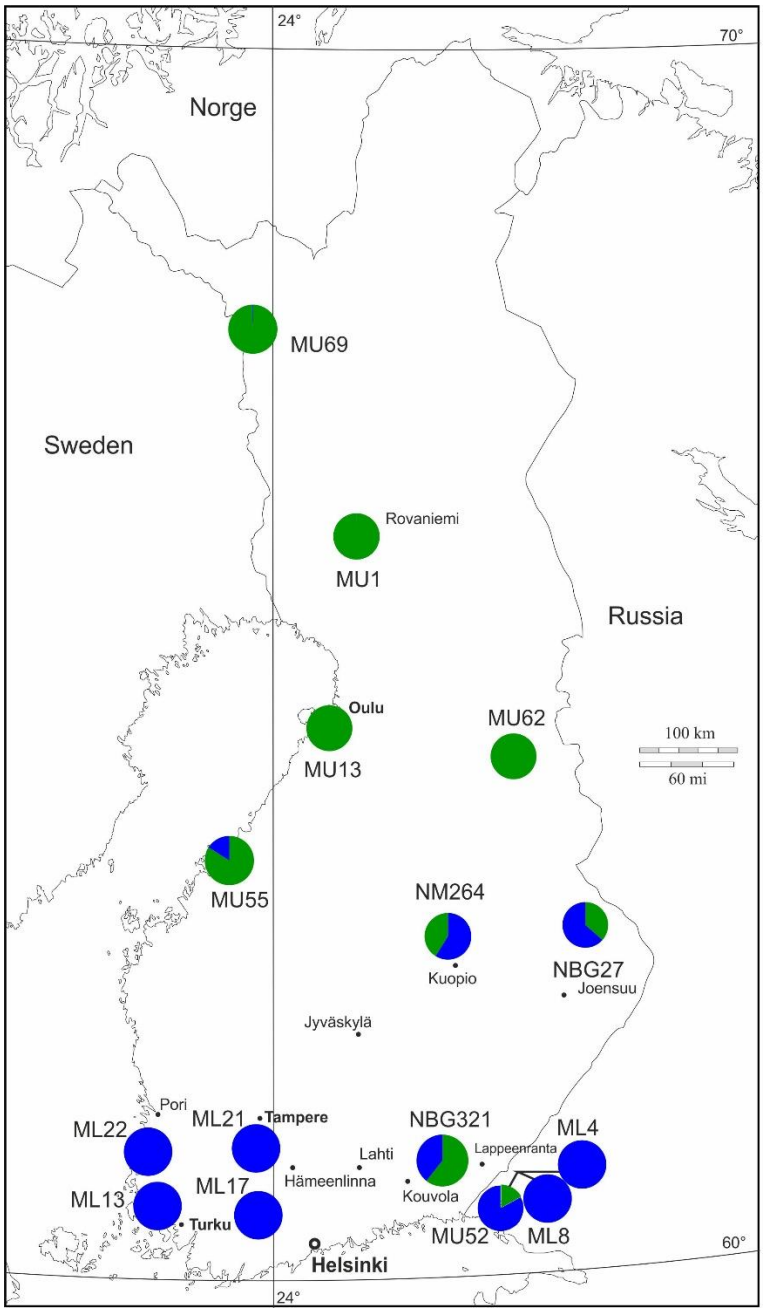
* Nowadays Russia

** Genotyped for Forsberg et al. 2015

*** Data for previously not genotyped samples



A



B

Supplementary file 1. Results of *STRUCTURE* analysis of all six-row individuals, assigned to two clusters. The coloured clusters do not correspond to those depicted in figure 4. A)

Distruct plot with each individual depicted by a vertical line segmented into up to two coloured sections. The length of each section is proportional to the estimated membership coefficient (Q) of the individual accession to each one of the two clusters. Thin black vertical lines separate different accessions. B) Map showing the location and proportional cluster membership of the different accessions.