Ancient mitochondrial DNA from the northern fringe of the Neolithic farming expansion in Europe sheds light on the dispersion process.

Helena Malmström1,†, Anna Linderholm2,3,†, Pontus Skoglund1, Jan Storå4, Per Sjödin1, M. Thomas P. Gilbert5, Gunilla Holmlund6,7, Eske Willerslev5, Mattias Jakobsson1,8, Kerstin Lidén2 and Anders Götherström1,2

1Department of Evolutionary Biology, Uppsala University, Norbyvägen 18D, 752 36 Uppsala, Sweden
2Archaeological Research Laboratory, Department of Archaeology and Classical Studies, Stockholm University, 106 91 Stockholm, Sweden
3Durham Evolution and Ancient DNA, Department of Archaeology, Durham University, South Road, Durham DH1 3LE, UK
4Osteolarchaeological Research Laboratory, Department of Archaeology and Classical Studies, Stockholm University, 106 91 Stockholm, Sweden
5Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5–7, 1350 Copenhagen, Denmark
6Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, and 7Department of Clinical and Experimental Medicine, Linköping University, Artillerigatan 12, 587 58, Linköping, Sweden
8Science for Life Laboratory, Uppsala University, Uppsala, Sweden

The European Neolithization process started around 12 000 years ago in the Near East. The introduction of agriculture spread north and west throughout Europe and a key question has been if this was brought about by migrating individuals, by an exchange of ideas or a by a mixture of these. The earliest farming evidence in Scandinavia is found within the Funnel Beaker Culture complex (Trichterbecherkultur, TRB) which represents the northernmost extension of Neolithic farmers in Europe. The TRB coexisted for almost a millennium with hunter–gatherers of the Pitted Ware Cultural complex (PWC). If migration was a substantial part of the Neolithization, even the northerly TRB community would display a closer genetic affinity to other farmer populations than to hunter–gatherer populations. We deep-sequenced the mitochondrial hypervariable region 1 from seven farmers (six TRB and one Battle Axe complex, BAC) and 13 hunter–gatherers and authenticated the sequences using postmortem DNA damage patterns. A comparison with 124 previously published sequences from prehistoric Europe shows that the TRB individuals share a close affinity to Central European farmer populations, and that they are distinct from hunter–gatherer groups, including the geographically close and partially contemporary PWC that show a close affinity to the European Mesolithic hunter–gatherers.

1. Introduction

The introduction of agriculture to human society brought major changes, as animal husbandry and cultivation of plants led to a more sedentary way of life [1] associated with increasing population sizes and density [2], and increasing social complexity [3]. This process, the Neolithization, started some 12 000 years ago in the Near East [4,5], after which it spread north and west throughout Europe. A long-standing question in European archaeology has been how these new practices spread. Did migrating people bring the new knowledge replacing the hunter–gatherer populations [6,7], or was the new lifestyle adopted by the Mesolithic hunter–gatherer communities through contacts with early farmers [8–10], or did the process vary in different regions [11]? According to the theoretical and ideological frameworks of the early-twentieth century, where changes in...
the material culture often were explained by demographic changes, the transition to the Neolithic lifestyle was first described as a revolution including a major migration wave [4]. However, further studies have shown that this process occurred along several different routes into and across Europe, at multiple time points [5,12,13].

The spread of Neolithic (farming) communities in the Northern parts of Europe seems to have taken place in phases, with one of the last being in Scandinavia. In Scandinavia, Northern Germany and Poland, archaeological evidence indicates that early farmers even lived side by side to hunter–gatherer groups [14,15]. These hunting groups appear to have adopted pottery, although not farming and animal husbandry practices. Furthermore, some of the hunter–gatherers (such as the Pitted Ware communities, PWCs) coexisted with farming communities for up to a millennium [16], indeed, sufficiently long that the PWC partly overlaps with two distinct agricultural complexes, the Funnel Beaker Culture (Trichterbecherkultur, TRB), and the Battle Axe complex (BAC) a variant of the Corded Ware Culture [17].

The earliest evidence of farming communities in southern Scandinavia is associated with the TRB culture, and dates to around 6000 BP [17]. In line with the different hypotheses regarding the spread of agriculture, some Scandinavian archaeologists have favoured the notion of Mesolithic populations as active adopters of new cultural elements, and that the different contemporary farming and hunting groups were the same people expressing different material culture at different sites [15,17–19]. However, the within Sweden geographically distinct TRB and PWC had different economic and different diets as indicated by stable isotope analyses [16,20]. Furthermore, recent genetic analyses have shed new light on the matter. Mitochondrial DNA (mtDNA) datasets generated on Mesolithic hunter–gatherers in Europe [21] and Neolithic PWC hunter–gatherers in Scandinavia [22] have high frequencies of haplogroups that are uncommon in many Neolithic farmer populations. Comparisons with modern populations also point to a population discontinuity between the Mesolithic/Neolithic hunter–gatherers and the modern populations [21–23]. The genetic background of Neolithic farmers in Europe shows a more complex picture [24], where haplogroup compositions in some regions are similar to extant European population frequencies [25,26], but in other cases, there is less support for direct continuity between early farmers and extant populations [27,28].

With the advent of high-throughput sequencing approaches, complete mitochondrial genomes have begun to appear for archaeological human remains. From Europe, data are available from nine Neolithic farmers [23,29,30], six Neolithic PWC hunter–gatherers [23,30], four Mesolithic hunter–gatherers [29–31] and the Tyrolean Iceman, a mummified body dated to the transition from the Neolithic to the Copper Age [32]. These techniques have also rendered possible the recovery of nuclear DNA data, from a range of archaeological materials [33–35]. However, nuclear data are still limited to small sample sizes, and mtDNA provides the opportunity to generate data from a larger number of individuals. The mtDNA variation of Mesolithic hunter–gatherers from central Europe and the last hunter–gatherers in Neolithic Scandinavia differ genetically from farmers across Europe (Linear Pottery Culture (Lineärbandkeramik, LBK) and TRB) [22,23,29,36–38], and this is also evident in genome-wide analyses of the two distinct Scandinavian groups [23,30].

Taken together, these genetic data show that the farmer communities represent a different population that in many areas dramatically impacted local resident groups, possibly even replacing them. To what extent this holds for Scandinavia is still unclear as only a limited number of early farmers have been genetically analysed [22,23,30]. As Scandinavia is a region at the geographical margin of early farming communities, and contained coexisting farmers and late hunter–gatherer communities [17,20,39], it holds a unique position to further highlight this question. We can hypothesis that if the introduction of farming practices was mitigated by migration, then the TRB may be expected to share a closer genetic affinity to other farming populations in Europe than to the contemporaneous PWC hunter–gatherers.

Here, we investigate the dispersal of the farming communities into Europe during the Neolithic, with particular focus on northern areas, by analysing population data from one of the northernmost remote farming communities, the Scandinavian TRB. We deep-sequenced the mtDNA d-loop from Neolithic samples using a genome sequencer (GS) FLX platform, and used postmortem DNA damage patterns to authenticate clone sequences from 13 PWC hunter–gatherers and from seven farmers (six associated with TRB and one with BAC). This data were merged with 124 previously published mtDNA sequences from Neolithic farmers and Mesolithic and Neolithic hunter–gatherers, from three geographical locations (Scandinavia: [22], Central Europe: [21,27] and Iberia: [24–26,31]), and we investigated the genetic diversity between and within groups, as well as possible continuity between either of the Neolithic Scandinavian groups and extant Scandinavians [40].

2. Material and methods

(a) Material

The study presents genetic analyses of 49 Neolithic individuals recovered in different archaeological contexts in present-day Sweden. Ten samples originate from Megalithic passage graves; Gökhem (Frälsegården; n = 7) located in Västergötland, dated to 5500–4500 years cal BP, and Resmo (n = 3) located on the Baltic Sea Island of Öland, dated to 5600–4600 cal BP [16]. The burials are associated with the TRB culture and a Neolithic way of life that included farming and husbandry. However, on Öland, some of the people buried at the passage grave Resmo included a substantial amount of marine proteins in their diet [16].

The 37 hunter–gatherer individuals analysed originate from PWC contexts on two of the large islands in the Baltic Sea; Köpingsvik (n = 7) on Oland dated to 5200–4850 cal BP, and Äjvide (n = 4), Fridtorp (n = 5), Iré (n = 9) and Visby (n = 10) on Gotland, all dating to approximately 5000–4400 cal BP [41,42]. Two samples were from the mainland site Korssnäs (n = 2) in Sweden and they were dated to 5300–4590 cal BP [43].

The two samples representing the BAC burial, a Scandinavian variant of the Corded Ware complex, are from Linköping, Östergötland (Bergsgraven, grave 185). Bergsgraven is a well-preserved flat earth triple grave containing a male, a female and an infant [44]. The grave was radiocarbon dated to cal 4800–4100 BP. The association between the BAC and TRB and PWC remains not fully understood. The BAC is chronologically younger (4900–4200 BP) than the TRB, but it is contemporaneous with the late phase of the PWC. The BAC group has been viewed by some archaeologists as a continuation and/or a transformation of the TRB or more closely associated with the PWC [17,19,45]. Recent studies of stable isotopes indicate that the
diet of BAC individuals was based on terrestrial proteins, thus, more closely resembling that of individuals in TRB context than individuals found in PWC contexts [46].

(b) DNA extraction, amplification, deep-sequencing and sorting

Bones and teeth from the 10 Neolithic samples from Öland (three TRB and seven PWC) were decontaminated, and DNA was extracted from duplicate samples at independent occasions under stringent clean conditions as in Malmström et al. [22]. The samples were extracted together with 48 non-Neolithic samples that are not included in this study and a total of 52 extraction blanks and 97 PCR blanks. Five short overlapping fragments and two nested fragments (between 85 and 127 bp including primers with unique 4 bp tags) of the mtDNA hypervariable region 1 (HVS1) were amplified, visualized on a gel, quantified, pooled at equimolar concentrations and deep-sequenced on a Roche genome sequencer FLX platform as in [22]. The samples were sequenced together with the non-Neolithic samples and with the sporadic negative controls that yielded bands on the gel. The FLX-generated ‘clone’ sequences were computationally sorted into one alignment file for each of the 14 amplicons based on their unique tags and the primer sequences as in Malmström et al. [47].

The data from the samples from Gotland, Västergötland and Östergötland were taken from a previous study in which the DNA had been extracted at two independent occasions, together with 109 extraction blanks (both non-human archaeological samples and mock extractions) and 16 PCR blanks, quantified with real-time PCR, amplified for the seven HVS1 fragments and deep-sequenced on a Roche genome sequencer FLX sequencer and bioinformatically sorted based on tags and primer sequences as in Malmström et al. [22,47]. In the earlier study, a conservative strategy was adopted to authenticate the sequence reads (all samples were required to contain at least 1000 template molecules, have support from two independent extractions, yield at least 20 ‘clone’ sequences from the independent DNA extracts and exhibit an inverse relationship between fragment length and DNA yield in qPCR analysis), and 22 of the samples were used for further analyses. Here, we reanalyse the raw sequence data from 39 of the samples (seven TRB, two BAC and 30 PWC) from which no sequence data were previously reported and, if possible, we improve the haplogroup resolution for the samples with unresolved haplogroup affiliation from the previously published sequences.

(c) Identification of ancient templates and consensus sequences

We used a statistical approach to identify authentic ancient sequence motifs in each amplicon in a similar way as in [47]. In short, the program PHYLOCNET (Agnar Helgason 2010, unpublished data) uses the c-statistic [48] to investigate postmortem nucleotide damage pattern (cytosine deamination) in clone sequences to identify the original sequence that gave rise to a damaged template. The sequence motif among the GS FLX clones from each amplicon that was most probably authentic (with best Cmax-value and p < 0.05) was selected. The amplicon sequences (up to 14 per sample) were aligned in BioEdit to create 341 bp HVS1 consensus sequences. Possible discrepancies between replicate amplicons were manually checked using PHYLONet. All negative controls that yielded amplicon data were also investigated using PHYLONet.

(d) Analyses of sequence data

The haplogroup affinity of each consensus sequence was assessed with HAPLOGREP (http://haplogrep.uibk.ac.at [49]) and with PHYLOTREE BUILD 10 or 15 (http://www.phylotree.org [50]). To investigate whether the Neolithic Scandinavian farmers are more likely to share ancestry with continental farmer populations or with resident hunter–gatherer populations, we merged our consensus sequences with previously published Mesolithic and Neolithic sequences from three different geographical areas. Thus, our data were analysed together with a set of previously published clone-based mtDNA sequences from Scandinavian Neolithic farmers (n = 3) and hunter–gatherers (n = 19) [22], Central European Neolithic farmers (n = 22) [27] and Mesolithic hunter–gatherers (n = 18) [21], and Iberian Neolithic farmers (n = 57) [24–26] and Mesolithic hunter–gatherers (n = 5) [24,31]. The sequences were aligned in BioEdit and after removing regions that were not present in all sequences (304 bp of HVS1 could be used for further analyses).

To investigate genetic differentiation between the six groups (farmers and hunter–gatherers from Scandinavia, Central Europe and Iberia, respectively), we computed FST between each pair of populations [51]. As the time differences between sequences also can give rise to a genetic differentiation, we used a temporal correction to the estimators which adjusts the differences between two sequences by the expected number of differences from the temporal separation [52]. In the case of FST, the mutation rate used for adjusting the number of differences does not affect the inferences. Because FST captures not only between population differentiation, but also within-group variation, and is thus highly sensitive to within-group genetic diversity, we also estimated the average between-group genetic differentiation by calculating pairwise mismatches. Here, each pairwise combination of sequences from the two groups was compared, and the average number of mismatches computed. Temporal correction and confidence intervals (CIs) were computed as above, assuming a mutation rate of 7.71 × 10–8 mutations per year and base pair in the 304 bp region we focused on the analyses. This mutation rate was obtained using the observed average divergence of 0.068992 substitutions per base pair between 613 modern-day Scandinavians to a Neanderthal sequence from Vindija Cave and an assumed Neandertal–modern human mtDNA divergence of 466 000 years [53]. The latter corresponds to 466 000 × 2 = 38 310 = 893 690 years of evolution taking the age of 38 310 years of the Neanderthal individual into account. We constructed neighbour joining trees using the ape R library [54] based on the average number of between-group mismatches both with and without temporal correction.

To evaluate the within-group genetic diversity, we computed the average number of mismatches between two randomly chosen sequences within each population using the same temporal correction as above. For all these analyses, CIs were estimated using 1000 bootstrap replicates over samples.

To assess a maximum possible contribution from an ancient population (TRB or PWC) to a younger population (extant Scandinavians), we used a method that investigates whether a null model incorporating genetic drift and different levels of contribution is consistent with the data [55]. More specifically, the observed frequency differences of the monophyletic mitochondrial U and K clade were used to quantify the maximum level of contribution from TRB farmers as well as from PWC hunter–gatherers to extant Scandinavians (represented by extant Swedes in this case). This assessment is performed in a model-based framework that evaluates the maximum contribution that is consistent with the observed allele frequency change and different amounts of historical genetic drift. The model includes possible gene flow from an unknown ‘ghost’ population, see [55]. The ‘ghost’ population is unspecified in order to set up a conservative null-model as other groups than TRB and PWC probably contributed to the modern-day Scandinavian gene pool. We used an observed frequency of 60 haplogroup U and K in a sample of size 286 from extant Swedes [40]. One of nine TRB/BAC individuals and 25 of 32 PWC individuals carried
the U or K mitochondrial type in the Neolithic samples. We used the age of the samples (in years BP, table 1) divided by 25 (the assumed generation time) to estimate the number of generations that separates the ancient groups and extant Scandinavians. An effective population size of between 2000 and 3000 was assumed for both ancient groups.

3. Results

We retrieved HVS1 sequences for 20 of the 49 Neolithic individuals (six TRB, one BAC and 13 PWC; table 1 and figure 1) and we were also able to improve the haplogroup resolution for 10 PWC samples from a previous study [22]. Using the c-statistic to identify authentic sequence motifs among the FLX-generated clones within each amplicon, 204 of the 280 amplicons yielded significant c-scores (electronic supplementary material, table S1). A total of 45 polymorphisms were found within these samples, and 32 of these were supported by significant c-scores from at least two amplicons that were generated from independent DNA extracts. The majority of the polymorphisms that were only supported by one significant amplicon (11 of 13) were consistent with other well-supported polymorphisms and created consistent haplotype types (in Res15, Ajv4, Vis 30, Fri28, Vis7 and Ber2). In some cases, these motifs were also seen in the replicate amplicon with the best Cmax-value, although not reaching significance (electronic supplementary material, table S1). The one sample with significant sequence support from only one amplicon did not have any other polymorphisms in the sequence to help support the haplotype. The average number of amplicon sequences that yielded significant c-scores per individual sample was 10.4 of 14. We also used the c-statistic to investigate for possible contamination in the negative controls (electronic supplementary material, text). No full HVS1 sequence could be retrieved, because results in the controls were highly sporadic. In the rare cases when two or more sequences could be assigned to one control, they consisted of polymorphisms that were contradictory in terms of known haplogroups.

When investigating the haplogroup composition of the HVS1 consensus sequences, the only haplogroup found in more than one of the new TRB individuals was H (table 1 and figure 1). One individual displayed zero polymorphisms and two individuals displayed one polymorphism (either at np 16 293 or at np 16 311) relative to the reference sequence [56]. For the samples that only had the 16311C polymorphism, it is not possible to discriminate between the H- and the R-haplogroup. One TRB individuals carry haplogroup J1d5 and one K1a5. The BAC individual displayed the N1a1a1a haplogroup. It was not possible to assign the Res1 individual to a specific haplogroup based on the three polymorphisms in the sequence. The three previously published TRB individuals also displayed H and J haplogroups in addition to T2b (table 1 and figure 1). Seven of the 13 new PWC individuals carry a haplogroup not found in any of the TRB/BAC samples, haplogroup U (mainly subhaplogroup U4, U5a and U5b; table 1 and figure 1). Another haplogroup unique to PWC was HVS0 which was found in one individual. Three PWC individuals shared the K1a1 haplogroup. One individual had the T2b haplogroup and differed from the TRB T2b that had a back mutation at np 16 296. The last of the new PWC individuals displayed a T-to-C transition at np 16 093 which could place the haplogroup within several subhaplogroups of H and R. The 19 previously published PWC individuals largely share the same haplogroups as the 13 new individuals and different subgroups of U are common (table 1 and figure 1). As the c-stretch between np 16 184–16 193 was removed in the previous study to avoid length heteroplasy difficulties and pyrosequencing misreads, eight individuals could not be resolved between U4 and H1b. After analysing the sequence reads again with the help of the c-statistics, we now conclude that all eight samples belong to U4 (as they do not display the C back mutation at np 16 189 that would be required for H1b). In a similar fashion, the haplogroup of two individuals that were previously assigned to the haplogroup denoted ‘other’ in [22], can now be assigned to haplogroups H/R based on the 16311C polymorphism.

Pairwise FST-estimates show that Scandinavian farmers were most similar to Central European farmers (FST = −0.141 (95% CI: −0.280, −0.002); the negative FST-estimates suggest that there is no differentiation between these groups; figure 1). Both these groups were significantly differentiated from the third group of farmers analysed, the Iberian farmers (FST = 0.292 (95% CI: 0.083, 0.500) and FST = 0.306 (95% CI: 0.181, 0.431) for Scandinavian and Central European farmers, respectively). Note, however, that in the neighbour joining tree based on pairwise genetic divergence between groups, all three farmer groups form a joint cluster compared with all hunter–gatherers (figure 3). The Scandinavian farmers were also distinct from both Neolithic Scandinavian and Mesolithic Central European hunter–gatherers (FST = 0.206 (95% CI: 0.026, 0.385) and FST = 0.222 (95% CI: 0.0856, 0.386)). Small sample size for the Iberian hunter–gatherers hindered direct comparisons with this group. The Scandinavian as well as the Central European farmers display high within-group nucleotide diversity (π = 0.016 (95% CI: 0.008–0.021); π = 0.012 (95% CI: 0.013–0.020); figure 4). In the neighbour joining tree (figure 3), we further show that the Neolithic PWC hunter–gatherers group with the Mesolithic hunter–gatherers of Central Europe and Iberia. Further, all hunter–gatherers show within-group homogeneity relative to the farmers of Scandinavia and Central Europe (π between 0.007 and 0.010; figure 4). Interestingly, the Iberian farmers show a similar within-group homogeneity as the hunter–gatherer groups, in contrast to the other farmers tested (figure 4).

We further compared extant Scandinavians to the Neolithic Scandinavian populations to investigate for possible continuity. To assess the maximum demographic contribution from Scandinavian Neolithic populations to extant Scandinavians, we used the following frequencies for the monophyletic U and K clade in our ancient samples: for PWC hunter–gatherers, 25 out of 32 individuals carry a U or K haplogroup, and for TRB/BAC farmers 1 out of 9 carry a U or K haplogroup (table 1). For the extreme upper bound (T = 0.096 units of time, i.e. the maximum number of generations divided by the minimum effective population size assuming 4800 years between the groups, a generation time of 25 years and a maximum effective population size of 2000 in the time between the sampled PWC population and present-day Scandinavians), we can reject a contribution greater than 60% from the PWC to the extant Scandinavian population (figure 5a). We could not reject any level of contribution from the ancient farmers to the extant Scandinavian population (figure 5b), suggesting that ancient farmers could have contributed substantially to extant Scandinavians.
Table 1. MtDNA haplogroup affiliation of new and previously published HVS1 sequences from Neolithic Scandinavian samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Grave</th>
<th>Site</th>
<th>Locality</th>
<th>Calibrated BP</th>
<th>Culture</th>
<th>MtDNA haplogroup</th>
<th>Polymorphisms</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gok5</td>
<td>Ind ID 124273</td>
<td>Gökhem Västergötland</td>
<td>c. 4900</td>
<td>TRB</td>
<td>K1a5&lt;sup&gt;a&lt;/sup&gt;</td>
<td>16224 C, 16311 C, 16362 C</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Gok4</td>
<td>Ind ID 122402</td>
<td>Gökhem Västergötland</td>
<td>c. 4900</td>
<td>TRB</td>
<td>H&lt;sup&gt;f&lt;/sup&gt;</td>
<td>16069 T, 16126 C</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>GE33 (KOP29)</td>
<td>Grave 4</td>
<td>Ajvide Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>T2b</td>
<td>U4a1</td>
<td>16093 C, 16356 C</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>GE76 (KOP27)</td>
<td>Grave 7</td>
<td>Visby Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>U5b1</td>
<td>16192 T, 16270 T</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Ajv29B</td>
<td>Grave 29B</td>
<td>Ajvide Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>U5b1</td>
<td>16192 T, 16270 T</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Vis5B</td>
<td>Grave 30B</td>
<td>Visby Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>U5a</td>
<td>16189 C, 16192 T, 16270 T</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Vis8</td>
<td>Grave 32</td>
<td>Visby Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>U5a</td>
<td>16192 T, 16270 T</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Ajv7</td>
<td>Grave 70</td>
<td>Ajvide Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>U5a</td>
<td>16192 T, 16270 T</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Ajv66</td>
<td>Grave 66</td>
<td>Ajvide Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>U5a</td>
<td>16192 T, 16270 T</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Ire8</td>
<td>Grave 8</td>
<td>Ire Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>U5a</td>
<td>16192 T, 16270 T</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Irel</td>
<td>Grave 9</td>
<td>Ire Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>U5a</td>
<td>16192 T, 16270 T</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Ire6B</td>
<td>Grave 6B</td>
<td>Ire Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>U5a</td>
<td>16192 T, 16270 T</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Ajv9</td>
<td>Grave 19</td>
<td>Ajvide Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>U5a</td>
<td>16192 T, 16270 T</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Ajv13</td>
<td>Grave 13</td>
<td>Ajvide Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>U5a</td>
<td>16192 T, 16270 T</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Ajv52B</td>
<td>Grave 52B</td>
<td>Ajvide Gotland</td>
<td>5000 – 4400 cal BP</td>
<td>PWc</td>
<td>U5a</td>
<td>16192 T, 16270 T</td>
<td>This study&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

(Continued.)
4. Discussion

The understanding of postmortem damage in DNA, and how it applies to ancient DNA [57–59], has enabled the development of methods that statistically identifies ancient DNA sequences [48]. Using this methodology, we assessed the damage distribution within large amounts of ‘clone’ sequences generated by GS FLX-sequencing, in order to compile the sequences that have high probabilities of being ancient into mitochondrial HVS1 consensus sequences. We generated HVS1 sequences from 20 Neolithic farmers and hunter–gatherers from present-day Sweden (six from TRB, one from BAC and 13 from PWC contexts). The human remains came from two sample sets. The first set (the samples from Öland, \( n = 7 \)) was generated for this study. The second set comprised ‘clones’ generated previously (the samples from Västergötland, Östergötland and Gotland, \( n = 13 \) [22]). Here, we analyse this second dataset in detail using the c-statistic framework and present the sequence data from these samples for the first time. With this method, we also improved the haplogroup resolution for 10 previously published PWC individuals [22]. We identified authentic sequence motifs in ‘clones’ from up to 14 amplicons per individual that derived from independent DNA extracts and retrieved coherent haplotypes. We also investigated sporadic human DNA contamination in negative controls. These typically displayed contradictory results and could not be used to generate complete HVS1 sequences (electronic supplementary material, text). These sequence data were analysed together with 124 mtDNA sequences from farmers (Neolithic), and hunter–gatherers (Neolithic and Mesolithic) from three geographical regions (Scandinavia, Central Europe and Iberia) [21,22,24–27,31].

Even though there was some overlap in haplogroup distribution between TRB and PWC (for haplogroups K, T and H), other haplogroups were unique to either TRB (J1d5 and J2b1a), BAC (N1a1a1a) or to PWC (U4, U4a1, U5, U5a and U5b1; table 1 and figure 1). The high frequency of U and K haplotypes in PWC individuals from both Gotland and Öland is in accordance with previous observations from human remains associated with PWC [22,23,30] and similar to a wide range of Mesolithic hunter–gatherers from...
Scandinavia, Central Europe, Iberia and Russia [21,29–31,60]. The U and K haplogroups are rare in Scandinavian TRB and Central European LBK farmers, but more common among individuals from the late Neolithic Treilles culture excavated in France, in Neolithic Iberian farmers and in Neolithic individuals from the Blätterhöhle cave in Germany [24–26,29,61]. The haplogroups seen in TRB (H, T, J and K) are also found in other farmer communities [27,29,38,61]. It has been hypothesized that communal Megalithic burials may represent specific individuals that are related to, or belong to, a certain strata of the community [12]. As eight distinct haplotypes were found among nine individuals from two different Megalithic burial structures, a close maternal kinship among several of the individuals can be ruled out. That these individuals represent more than a close family is further supported by isotopic studies (strontium, Sr) of migration patterns which have shown that as many as 25% of the buried in the megalithic burial structures were of non-local birth [62]. Interestingly, the BAC individual (N1a1a1a) did not share haplogroup affiliation with previously published Corded Ware associated individuals from Central Europe [63] but rather with individuals associated with LBK [64].

The Neolithic farmers analysed in this study differ significantly from the contemporaneous PWC hunter–gatherers, but show a close affinity to LBK individuals (figures 2 and 3). It has recently been shown not only that the TRB individuals

Figure 2. Pairwise $F_{ST}$ estimates with temporal correction and confidence intervals. Black point is value between hunter–gatherer groups (excluding Iberian hunter–gatherers as there were only five samples), grey points are values between farmer groups and white points are values between groups with different subsistence patterns (hunter–gatherers and farmers). HG, hunter-gatherer.

Figure 3. Neighbour joining tree of pairwise genetic differences (with temporal correction) between farmer and hunter–gatherer groups in Neolithic and Mesolithic Europe.

Figure 4. Nucleotide diversity (with temporal correction and confidence intervals (CI)) within each group (hunter–gatherers and farmers from Scandinavia, Central Europe and Iberia, respectively). Note that no CI could be determined for the Iberian hunter–gatherers owing to the small sample size.
investigated in this study, but also other Early and Middle Neolithic cultures in Central Europe, are genetically similar to the LBK [38]. Interestingly, both Scandinavian and Central European farmers are differentiated from the Iberian farmers (figure 2), however, mtDNA indicates that the population history in Iberia is complex during the Neolithic and that the haplogroup composition differ from that found in North/ Central Europe as well as between different areas in Iberia [24]. It has further been suggested that the mode of the Neolithic expansion may have been different in geographically distinct regions [65], although alternative explanations could relate to continuity from Mesolithic populations or admixture with North African groups. However, even though the Iberian individuals span from the early to the late Neolithic, they are much more homogeneous than the other investigated farmer groups, which would contradict admixture (figure 4).

We show that the PWC hunter–gatherers of Neolithic Scandinavia do group with Mesolithic hunter–gatherers (figure 3). Our data further support genetic homogeneity of the hunter–gatherer complexes of Scandinavia and Central Europe (figure 4) which has been previously indicated [21,29–31]. However, the genetic picture of Mesolithic Europe is still incomplete, and open questions specifically related to Scandinavia include the relationship between PWC and the Ertebølle communities. In addition to their apparent mtDNA affinity to each other, the hunter–gatherer groups in both Scandinavia and Central Europe show significantly lower within-group mtDNA diversity than farmer groups in the same regions (figure 4). This could be a consequence of a number of previously hypothesized historical processes, such as reduction in the size of the ancestral hunter–gatherer population in the Late Pleistocene, long-term small effective population sizes in the small and probably dispersed hunter–gatherer bands, or a greater degree of admixture in migrating farming groups [66,67]. To what extent this reduced genetic diversity is also seen in the autosomal genome and the Y-chromosome is an interesting topic that will require further studies.

In a previous mitochondrial study, we found that the observed genetic differentiation between PWC and extant Scandinavians was inconsistent with complete population continuity [22]. To give a more precise upper bound of possible PWC ancestry in extant Scandinavians, we modelled extant Scandinavians as the result of gene flow between the PWC group and an unknown population. We found that the maximum direct demographic contribution consistent with the data is 60% (figure 5b). This is consistent with previous modelling of admixture proportions in Scandinavians using low-coverage genome-wide data [23,30]. However, note that we could not reject any level of contribution from TRB to extant Scandinavians, suggesting that our data are compatible with a scenario spanning from a complete replacement of the PWC to a scenario of admixture up to 60% from PWC into the population that eventually lead to modern-day Scandinavians.

To conclude, the observations that: (i) mtDNA of the farmers of the TRB community at the northern fringe of the Neolithic expansion closely resemble that of early Central European farmers, (ii) do not resemble contemporaneous PWC hunter–gatherers, but (iii) the mtDNA of these PWC hunter–gatherers’ resemble the Mesolithic groups of continental Europe, suggest a key role for migration in association with the Neolithization process. A more detailed and nuanced picture will no doubt be achieved by analysis of nuclear DNA and for a greater number of individuals from a larger timespan in the region and in other parts of western Eurasia.

**Data accessibility.** The Neolithic sequences (including sequences with increased haplogroup resolution compared with Malmström et al. [22]) have been deposited to GenBank (https://www.ncbi.nlm.nih.gov/genbank/) under the accession numbers KJ873248–KJ873275.

**Acknowledgements.** The authors thank the staff at the Danish National High-Throughput DNA Sequencing Centre for assisting in the data generation. We also thank Agnar Helgason at deCODE Genetics in Iceland for letting us use the unpublished program PhyloNet version 5 and Karl-Göran Sjögren, Per Persson, Marietta Douglas and Leena Drenzel for providing samples.


