Steppe Ancestry in western Eurasia and the spread of the Germanic Languages

Genetics Supplementary Material

S2. Data Generation.

Charleen Gaunitz, Lasse Vinner

All aDNA laboratory procedures on non-amplified DNA were conducted in dedicated aDNA clean lab facilities at the Lundbeck Centre for GeoGenetics, University of Copenhagen, according to strict guidelines (Willerslev and Cooper 2005; Gilbert et al. 2005). Post amplification procedures were conducted in post-PCR laboratories physically separated from the clean lab facility.

During the project a semi-automated data generation pipeline was introduced over a two-year period, hence the samples included in the present study were processed in an increasingly automated manner as procedures were automated for extraction, library preparation, PCR setup, post-amplification purification, and library pooling.

Curation of included samples

In order to improve data quality and minimize destruction of unproductive samples, we implemented a thorough quality control check to make sure that all the samples analyzed, teeth and petrous bones, were of great quality.

Macroscopic study of the surface of the teeth using a magnifier and an indirect source of light was performed. We only considered teeth that were fully formed, erupted, and showing complete roots with the apex intact. We only retained teeth with roots of yellow to brownish color. Teeth with roots of chalky or white colors were excluded.

The petrous bones were considered when the portion containing the cochlea was present. The petrous bones that were cut off too short from the temporal bone were excluded as well as those that were showing a too chalky color aspect and a very light weight.

Drilling

Drilling was performed manually. If possible, right or left petrous bone, one tooth and associated calculus were subsampled from one individual. Upto150 mg of crushed subsample were stored in a 2 ml LVL barcoded tube (XLX2000, DSC-X20-BL-NS-SLP-S). While sampling for DNA analysis, a small piece of bone (up to 1 gram) was collected for radiocarbon dating. Sample status was documented photographically before and after the destructive sampling. Sub-samples were stored at -20°C until demineralization.

Demineralization

Sample demineralization was carried out in two steps. First, an initial short incubation step was performed at 37°C for 30 min to increase the recovery rate for endogenous DNA, using 1.0ml of incomplete digestion buffer consisting of 0.5M EDTA and 30% N-Lauroylsarcosin (Damgaard et al. 2015). The buffer was replaced with 1.8ml of freshly prepared digestion buffer (0.5M EDTA, Proteinase K and 30% N-Lauroylsarcosin) and incubated for 48 to 72 hours at 37°C on a rocking table. Demineralized samples (lysates) were stored at 4°C for no longer than 14 days, otherwise at -20°C for long term storage.

Ancient DNA Extraction

All aDNA extractions for this project were carried out according to (Rohland et al. 2018) using a modified version of the Qiagen PB buffer (Qiagen, cat: 19066). In total, two extractions were performed per lysate for initial library sequencing. One extraction per lysate underwent treatment with USER enzyme mix (NEB, cat: M5505L) the other remained untreated.

Manual extraction setup

For manual extractions 1.8 ml of lysate was combined with 18 ml of binding buffer (modified Qiagen PB binding buffer: 500 ml Qiagen PB, 15 ml Sodium acetate 3 M, 1.25 ml 5M NaCl, phenol red, adjusted to pH=5). The mix was passed through a large volume silica column (Roche, cat: 05114403001) by centrifugation for 4 min at 1500 rpm. The flow through was discarded and the step above repeated 2 times for the full volume. The column was washed twice using 750 µL 80% ethanol + 20% 10mM Tris-HCl (Qiagen, cat: 19065) and a centrifugation step of 30 seconds at 6-8 K·g, followed by a dry spin at full speed, according to kit protocol. The column was then transferred to a fresh 1.5 ml LoBind DNA Eppendorf tube (Eppendorf, cat: 0030108051) before final elution with 65 µl of 10 mM Tris-HCl + 0.05% Tween-20 incubated for 5 min at room temperature, followed by a centrifugation step for 1 min at 11.5 K·g. Extractions were stored at 4°C until library preparation, no more than 2 days or longer at -20°C.

Automated aDNA extraction setup

The aDNA extraction procedure was automated on the Biomek i5 Automated Workstation (Beckman Coulter) with the following adaptations: For each subsample 150 µl of demineralized lysate were combined with 1560 µl of binding buffer and 10 µl of magnetic

126 silica beads (G-Bioscience, cat: 786-915). The mixture was incubated for 15 min and tip-mixed
127 every 5 min. Pelleted beads were washed twice in 450µl and 100 µl of 80% ethanol + 20% 10
128 mM Tris-HCl, respectively (Qiagen PE, cat: 19065). The final product was eluted in 35 µl of
129 10 mM Tris-HCl + 0.05% Tween-20. Extracted DNA was stored in 96 well non-skirted plates
130 (VWR) at 4°C until library preparation for no more than 2 days otherwise at -20°C. All
131 extraction runs included a positive control sample and ≥1 negative control (buffer only). Robot
132 deck specifications, method (.bmf format) and plastic/labware are provided on request.
133 To verify successful DNA extraction, 2 µl of the positive and negative control extracts were
134 measured on QuBit (dsDNA HS Assay Kit, cat: Q32851). The QuBit threshold value for the
135 positive extraction control was >0.1ng/µl for approval, whereas the negative was below the
136 lower detection limit.

137 USER-treatment

138 For the USER-treatment 10 µl of USER Enzyme and 35 µl of DNA extract were mixed and
139 incubated for 3 h at 37°C in a thermal cycler (SimpliAmp™ ThermoFischer). The treatment
140 was performed in a 96 well format. After performing tests with reduced USER enzyme volume,
141 the reaction volume of the USER enzyme was reduced to 2.5 µl. From May 2022 the USER
142 treatment protocol reads: 2.5 µl USER Enzyme + 7.5 µl H2O + 35 µl aDNA, incubated for 3 h
143 at 37°C.

144 NGS Library preparation

145 Manual preparation

146 Double-stranded DNA libraries were prepared according to (Meyer and Kircher 2010) from
147 the USER- and one nonUSER-treated extracts, using input template volumes of 42.5 µl (final
148 reaction volume 50 µl) and 21.25 µl (final reaction volume 2 µl), respectively. Clean-up steps
149 were performed after the end-repair and adapter-ligation step, using 10 volumes of modified
150 Qiagen binding buffer (as above) and MinElute columns (Qiagen, cat: 28004) otherwise
151 according to the MinElute kit protocol.

152 Automated NGS Library preparation

153 The NGS library preparation for USER-treated and non-USER-treated libraries was also
154 adapted to the Biomek i5 Automated workstation (Beckman Coulter). Two different methods
155 were written to accommodate the different input volumes of USER and non-USER extracts.

Clean-up steps were performed after the end-repair and adapter-ligation step, using 10 volumes of modified Qiagen binding buffer (as above) and 10 µl of magnetic silica beads (G-bioscience, cat: 786-915). Pelleted beads were washed twice in 450 µl and 100 µl of 80% ethanol + 20% 10 mM Tris-HCl, respectively (Qiagen PE buffer). The purified products were eluted in 10 mM Tris-HCl + 0.05% Tween-20. Robot deck specifications, method (.bmf format) and plastic/labware are provided on request.

Real-time PCR

We used real-time PCR for determining the number of PCR cycles required in the indexing PCR, for each batch of libraries. Excluding the control libraries, c_t-values for the sample libraries were converted to a consensus amplification cycle number, considering amplification curves, c_t-values and melting curves. The qPCR with a final reaction volume of 20µl was set up manually in a 96 well format (Roche, cat: 5102413001), using 1 µl of input material, 7 µl of H_2O, 500 nM forward- and 500 nM reverse primer (10 mM) and 10 µl of 2x Lightcycler 480 SYBR green I MasterMix (Roche, cat: 4707516001). The qPCR amplification was run in the POST PCR lab facilities using the Roche LightCycler 480 Real-time PCR system (Roche).

Primer ID	Primer Sequence (5'-3')
qPCR CH_P5_	CTACTGACTTTCAGTGAGTGCAACCCACGACGCTCTTCCGATCT
qPCR CH_P7_	CTCTCACATTGAATCCGACTAGGATACGTGTGCTCTTCCGATCT

qPCR conditions

Temp	Time	Cycles
95 °C	10 min	1
95 °C 55 °C 72 °C	30 sec. 30 sec. 30 sec.	 30 Detection
95 °C 55 °C	30 sec. 30 sec.	1

95 °C Slow ramp	30 sec.	Detectio n

Indexing PCR

The indexing PCR was set up either in a small PCR reaction volume of 50 µl for non-USER-treated libraries or in a big PCR reaction volume of 100 µl for USER-treated libraries. For the small PCR reaction setup 25 µl of KAPA HiFi HotStart Uracil+ ReadyMix (Roche, cat: 07959079001) was mixed with 4 µl of UDI (8-bp index) or UDP (10-bp index) Illumina primer pairs (400 nM final conc. each) and 21 µl of template DNA. For the large PCR reaction setup 50 µl of KAPA HiFi HotStart Uracil+ ReadyMix (Roche, cat: 07959079001) was mixed with 8 µl of UDI or UDP Illumina primer pairs (400 nM final conc. each) and 42 µl of DNA.

PCR amplification conditions:

Temp.	**Time**	**Cycles**
98	45 sec	1
98	15 sec	12-18
65	30 sec	
72	30 sec	
72	min	1

PCR amplification was performed in the post-PCR lab facilities. PCR amplification cycles were determined as described above. Indexing PCR setup was semi-automated on the Biomeki5 Automated Working station (Beckman Coulter). The KAPA HiFi HotStart Uracil+ ReadyMix (Roche, cat: 07959079001) was dispensed manually to a non-skirted 96 well PCR plate (VWR) using a multi-dispenser pipette and then placed on the 96-well cooling element on the Biomeki5. Index primer pairs and template DNA were then transferred by the multichannel robot head to the plate. Robot deck specifications, method (.bmf format) and plastic/labware are provided on request.

Purification of amplified libraries

Manual procedure

For non-USER- and USER-treated libraries, respectively, 80 µl or 160 µl of MagBio High-Prep PCR beads per library were transferred into a 96-deep well plate (e.g., ThermoFisher cat: 267245). Subsequently 50 µl or 100 µl, respectively, of amplified library were added to the beads, thoroughly mixed and incubated for 5 min at room temperature. The plate was placed onto a magnetic plate and incubated until the supernatant was clear. While still on the magnet the beads were washed twice with 200 µl of 80% freshly made ethanol. After discarding the last washing buffer, the plate was left to dry for 2-5 min. The plate was removed from the magnet. The beads were resuspended in 35 µl of 10 mM tris-HCl pH= 8.5 (EB Buffer, Qiagen) and incubated for 2 min, before clearing the supernatant on the magnet (Alpaqua Magnum FLX). Once clear, the supernatant was transferred to a storage tube (LVL technologies, 2DSC-X03-BL-NS-SLC-S).

Automated purification

The purification was set up in an automated 96-well format using the CyBioFelix liquid handler (Analytik Jena), following the conditions described above. Two protocols are currently available. One using the 96R Head allowing the purification of 96 samples at once with a run time of 47 min., and one using the Choice Head (8-Channel). For the latter the number of samples to purify can be chosen. Minimum is 8 the maximum is 96. The final purified DNA library (35µl) is stored in SX 300 1D and 2D barcoded tubes (LVL technologies, 2DSC-X03-BL-NS-SLC-S) at -20°C. The rack is scanned prior purification using the Ziath express scanner. Labware and method are available upon request.

Quality Control (Fragment Analyzer)

Library concentration and fragment length distribution was determined by using the 5300 Fragment Analyzer System (Agilent) with the HS NGS Fragment Kit (1-6000bp), 500 (cat: DNF-474-0500, Agilent). The standard protocol provided by Agilent is followed to run the HS NGS Fragment Kit, using 2 µl of library input with 22 µl Diluent Marker (total 24 µl) in a semi-skirted Eppendorf twin.tec PCR Plate 96. For data analysis the ProSize (Agilent) software was used.

Pooling of libraries

Manual procedure

Each sequencing pool consisted of a maximum of 96 libraries, with each library distinguishable by a unique index pair (Truseq UDP or UDI, Illumina). For initial sequencing the libraries were pooled equimolarly.

Automated procedure

The preparation of sequencing pools was automated on the CyBioFelix liquid handler (Analytik Jena), with the possibility to pool up to 192 uniquely indexed libraries in one run. Input pooling volumes, rack position and optional dilution requirement were loaded via a .csv file.

The pooling method itself is divided into 2 parts, which can be run separately. In the first step, a dilution of all libraries is prepared. The dilution factor is the same for all samples in the plate (typically 5 or 10-fold) and created to ensure no pipetting of ≤ 1µl. In the second part of the method, the pooling of the diluted and undiluted libraries was carried out. Samples are transferred into a 1.5 ml Lobind Eppendorf tube, starting from the largest volume to the smallest volume. The final sequencing pool was purified manually using 1.6x volumes of paramagnetic beads (Magbio High prep Beads, cat: AC-60500), and washed with 80% freshly prepared ethanol (500µl) removing residual adapters and indexing primer. To compensate for possible loss during purification, pools were eluted in 2/3 of their original volumes. CyBioFelix method- and labware files are available on request.

Sequencing

The molar concentration of sequencing pools was determined using qPCR (KAPA Library Quantification Kit, KAPA Biosystems, KR0405) according to kit protocol. All sequencing was done on the Illumina NovaSeq6000 platform at the GeoGenetics Sequencing Core. Sequencing chemistry version 1.0 was used in the beginning of the project but later switched to the latest version 1.5 in mid 2021 due to updated kit version. In-house comparison between kit versions revealed no biases in subsequent analysis pipelines. For the USER-treated libraries, pools were sequenced on the Illumina S4 flow cell, 200 cycles, generating ≥10G paired-end reads. For the nonUSER-treated libraries, pools were sequenced on one lane of an Illumina S4 flowcell, 2x100 cycles, using Illumina's XP-kit and -workflow, generating 2.5G sequencing reads of 100bp paired-end. Loading concentrations for ancient DNA libraries deviate from Illumina

recommendations for loading concentrations, and differ between the two types of runs, i.e., full flow cells and XP-workflows. For a full flow cell run, aDNA library loading concentration is optimal at 700 pM and for the XP-runs, loading concentration seems optimal at 550-600 pM. Each pool is quality checked before sequencing, using Fragment Analyzer or Bioanalyzer to determine the average bp size and profile of the library, and a qPCR to determine molarity of the pool. The yields from full flow cells seem more stable >10G reads, than output from XP-workflows ≥2.5G reads. Demultiplexing of sequencing output was performed on BaseSpace Sequence Hub (Illumina).

References

Damgaard, P. B., A. Margaryan, H. Schroeder, L. Orlando, E. Willerslev, and M. E. Allentoft. 2015. 'Improving access to endogenous DNA in ancient bones and teeth', Sci Rep, 5: 11184.

Gilbert, M. Thomas P., Hans-Jürgen Bandelt, Michael Hofreiter, and Ian Barnes. 2005. 'Assessing ancient DNA studies', Trends in Ecology & Evolution, 20: 541-44.

Meyer, M., and M. Kircher. 2010. 'Illumina sequencing library preparation for highly multiplexed target capture and sequencing', Cold Spring Harb Protoc, 2010: pdb.prot5448.

Rohland, N., I. Glocke, A. Aximu-Petri, and M. Meyer. 2018. 'Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing', Nat Protoc, 13: 2447-61.

Willerslev, E., and A. Cooper. 2005. 'Ancient DNA', Proc Biol Sci, 272: 3-16.

S3. Bioinformatics Preprocessing Pipeline

Abigail Ramsøe, Isin Altinkaya, Thorfinn Sand Korneliussen

Demultiplexing

The sequencing data was demultiplexed on the BaseSpace Sequence Hub using BCL Convert (*v4.0.3* BCL Convert Support, Illumina Inc.), allowing for one mismatch in the index. As this strategy allows the inclusion of reads with a single mismatch in the index sequence, it avoids the needless exclusion of reads. Furthermore, as the libraries were sequenced using Unique Dual Indexing (UDI), and the Illumina index sets are designed with a Hamming distance of four, the chances of a sequenced read being misassigned is nil. This generates two fastq files (read 1 and read 2) per library per lane - in total two fastq files per library for XP runs, and eight for full flowcells.

Trimming

AdapterRemoval --threads {threads} --file1 {input.R1} --file2 {input.R2} --minlength 30 --adapter1 {a1} --adapter2 {a2} --collapse-conservatively --basename {params.out} --gzip

As we sequence 100bp reads, but, due to the characteristic short fragments of ancient DNA, we expect to sequence into the adapter sequence at the 3' end of the reads. As such, it is imperative to remove these technical sequences from the reads before downstream mapping and analysis. We use AdapterRemoval (2.3.2, (Schubert et al., 2016)) to remove adapter sequences from the 3' end of the sequenced paired end reads. We specify the adapter sequences to be the ends of the Illumina TruSeq adapters, before the index sequence, as shown in Table S3.1**.** This is to allow variations in index sequence lengths.

Table S.3.1 Adapter sequences used by AdapterRemoval

	Adapter 1	**Adapter 2**
Double-stranded libraries (and Santa Cruz Reaction)	AGATCGGAAGAGCACACGTCTGAACTCCAGTCA	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

Single-stranded (ss2.0) libraries	AGATCGGAAGAGCACA CGTCTGAACTCCAGTCA	GGAAGAGCGTCGTGTA GGGAAAGAGTGT

As reads below 30 base pairs are very problematic to align, they are discarded at this stage using the –min-length parameter. Here, paired end reads with an overlap of 11 base pairs or greater were collapsed into longer consensus sequences using the –collapse-conservatively parameter, which achieves an increase in the length and the quality of the insert sequences. This generates three fastq files per library per lane: pair 1, pair 2, and collapsed reads.

Mapping

```
bwa aln -l 512 -t {threads} {REF} {input.fq} > {output.sai}
```

We then map each of the generated fastq files against the human reference genome build 37 (hs37d5, md5sum 12a0bed94078e2d9e8c00da793bbc84e) using bwa aln (*0.7.17*, (Li & Durbin, 2009)). Here, we disable seeding (-l 512) and instead perform end-to-end alignment, as suggested for ancient samples (Schubert et al., 2012) in order to maximise the number of reads mapping to the human genome. As each bwa aln step is an independent process, it happens in parallel for all trimmed fastqs in a given sequencing run.

```
bwa samse -r \"{params}\" {REF} {input.sai1} {input.fq1} > {output.sam}

bwa sampe -r \"{params}\" {REF} {input.sai1} {input.sai2} {input.fq1} {input.fq2} >
{output.sam}
```

Next, when all alignments are completed, we use bwa sampe and samse (0.7.17, (Li & Durbin, 2009)) on the paired and single end fastq and sai files respectively. At this step we set the RG tags of the resulting sam files. The ID tag is set to the fastq name plus whether the reads were paired or collapsed during adapter trimming, the SM tag is a unique sample identifier, and the LB tag is the library barcode. This allows accurate marking of duplicates during downstream analyses.

```
samtools sort {input.cram} -@{threads} -m4G > {output.bam}
```

This produces sam files, which are then sorted and converted into bam files using samtools sort (*v1.10 (Li et al., 2009)*). Lastly, paired- and single-end reads for each library and lane are merged together, such that there is one resulting bam file per library.

samtools merge {output.merged} {input.bams} -@{threads}

Duplicate Marking

java -Djava.io.tmpdir={params.tmpdir} -XX:ParallelGCThreads={threads} -Xmx2g -jar {picard.jar} MarkDuplicates OPTICAL_DUPLICATE_PIXEL_DISTANCE=12000 I={input} o={output} REMOVE_DUPLICATES=false METRICS_FILE={params.metrics} TAGGING_POLICY=All VALIDATION_STRINGENCY=LENIENT

The library construction and sequencing processes create duplication of reads in libraries. Such duplicate reads add no extra information to the downstream analyses, and as such should be discarded. Duplicates were marked using Picard MarkDuplicates (2.25.0) with a pixel distance of 12000. For each duplicate read, this updates the flag to include the bitwise flag 1024, indicating that the read is a duplicate. These reads are not removed at this stage, but setting this flag prompts downstream software to disregard them.

Library complexity

decluster/decluster {input} -w -p 12000 -@{threads} -o {params.out}

For all screening libraries we extrapolated the possible yield from additional sequencing experiments. The approach is using a mark and recapture technique originally implemented in preseq (Daley et al., 2014) and works by downsampling the histogram of the PCR clonality multiplicities. The NovaSeq 6000 platform is, however, using a patterned flowcell technology, and depending on library load it will generate non-negligible amount of cluster duplicates. The preseq software is therefore not directly applicable to our sequencing data. Existing tools for removing these cluster duplicates from single end data are kmer based (Bushnell, 2014) and very memory intensive. Therefore, we developed a program, decluster, that constructs the PCR duplication table from the mapped data (https://github.com/ANGSD/decluster). Using the output of decluster, and information about how many reads were sequenced, we project how

many more reads would be needed to reach a given level of coverage, and use this information to inform further sequencing.

S4. DNA authentication

Abigail Ramsøe, Thorfinn Sand Korneliussen

In order to ensure the authenticity of the sequenced data, we employed three approaches: mapDamage to quantify the extent of deamination, and two contamination estimates. These contamination estimates were first run on a library-level basis, and any contaminated libraries were excluded. When the different libraries for a sample were merged, the authentication methods were then rerun to ensure that there was still no contamination flagged. This strategy ensures the high-confidence detection of both ancient and modern human contamination.

All analysis was performed on filtered files with reads with a base quality of at least 20, and a mapping quality of at least 30. For ANGSD, we removed the pseudoautosomal regions on the X chromosome.

mapDamage

```
mapDamage -i {input.bam} -r {REF}
```

In order to determine the authenticity of the ancient reads, the extent of the characteristic C->T deamination in libraries was quantified using mapDamage2.0 (Jónsson et al., 2013). Libraries were flagged for manual interpretation if the C->T extent at the 5' end, and/or G->A' extent at the 3' end did not follow expected parameters given the library type. mapDamage was also run on USER treated libraries, where we do not expect to see the ancient damage signal. Here, libraries were flagged if they exhibited the elevated C->T, as this could indicate a sample swap, or contamination from an ancient source.

ContamMix

```
Rscript contamix/estimate.R --samFn ${base}_ra.final.bam --malnFn ${base}_312.aligned.fasta --figure ${base}_contam.pdf --trimBases 7 > ${base}.summary.txt
```

Firstly, we applied ContamMix in order to quantify the fraction of exogenous content in the set of reads mapping to the mitochondrial genome by comparing the mtDNA consensus genome to 311 possible contaminant genomes (Fu et al., 2013). For each libraryt two different versions of the endogenous mitochondrial genome were constructed. The first approach (CONTAMIX_APPROX_1Xdif05) used sites with at least 1x coverage, and at each position a base was only called if it was observed in at least 50% of reads covering the site. The second approach (CONTAMIX_PRECISE_5Xdif07) only considered sites with at least 5x coverage and 70% base concordance.

ANGSD

```
${ANGSD} -i ${bam} -r X:5000000-154900000 -doCounts 1  -iCounts 1 -minMapQ 30 -minQ 20 -out ${base} ;
${ANGSDDIR}/misc/contamination -a ${base}.icnts.gz -h ${ANGSDDIR}/RES/HapMapChrX.gz &>${base}.res
```

ANGSD (0.931, (Korneliussen et al., 2014)) was used in samples identified as males to estimate nuclear contamination by quantifying excess heterozygosity on the X chromosome. As males only have one copy of the X chromosome, any heterozygosity is due to contamination or sequencing error. Heterozygosity due to contamination is likely to be restricted to known polymorphic sites. As such, ANGSD quantifies the heterozygosity in these sites, and compares it to the excess heterozygosity in the flanking regions, which provides the level of background sequencing error. Thus, the extent of contamination is calculated.

References

Bushnell, B. (2014). *BBMap: A fast, accurate, splice-aware aligner* (No. LBNL-7065E). Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States). https://www.osti.gov/biblio/1241166

Daley, T., Helus, V., & Smith, A. (2014). *The preseq Manual*. Retrieved April 20, 2023, from http://smithlabresearch.org/wp-content/uploads/manual.pdf

Fu, Q., Mittnik, A., Johnson, P. L. F., Bos, K., Lari, M., Bollongino, R., Sun, C., Giemsch, L., Schmitz, R., Burger, J., Ronchitelli, A. M., Martini, F., Cremonesi, R. G., Svoboda, J., Bauer, P., Caramelli, D., Castellano, S., Reich, D., Pääbo, S., & Krause, J. (2013). A revised timescale for human evolution based on ancient mitochondrial genomes. *Current*

Biology: CB, *23*(7), 553–559.
Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F., & Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* , *29*(13), 1682–1684.
Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, *15*, 356.
Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* , *25*(14), 1754–1760.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* , *25*(16), 2078–2079.
Schubert, M., Ginolhac, A., Lindgreen, S., Thompson, J. F., Al-Rasheid, K. A. S., Willerslev, E., Krogh, A., & Orlando, L. (2012). Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics*, *13*, 178.
Schubert, M., Lindgreen, S., & Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Research Notes*, *9*, 88.

S5. Imputation

Imputation was carried out on 3929 already published ancient genomes (Rasmussen *et al.*, 2010, 2014; Fu *et al.*, 2014; Gamba *et al.*, 2014; Lazaridis *et al.*, 2014; Olalde *et al.*, 2014; Raghavan *et al.*, 2014; Seguin-Orlando *et al.*, 2014; Skoglund *et al.*, 2014; Rasmussen *et al.*, 2015; Olalde *et al.*, 2015, 2015; Raghavan *et al.*, 2015; Allentoft *et al.*, 2015; Günther *et al.*, 2015; Jones *et al.*, 2015; Llorente *et al.*, 2015; Fu *et al.*, 2016; Broushaki *et al.*, 2016; Cassidy *et al.*, 2016; Gallego-Llorente *et al.*, 2016, 2016; Hofmanová *et al.*, 2016; Kılınç *et al.*, 2016; Martiniano *et al.*, 2016; Schiffels *et al.*, 2016; Skoglund *et al.*, 2017; Jones *et al.*, 2017; Martiniano *et al.*, 2017; González-Fortes *et al.*, 2017; Haber *et al.*, 2017; Hansen *et al.*, 2017; Lipson and Reich, 2017; Rodríguez-Varela *et al.*, 2017; Saag *et al.*, 2017; Schlebusch *et al.*, 2017; Sikora *et al.*, 2017; Unterländer *et al.*, 2017; Olalde *et al.*, 2018; Günther *et al.*, 2018; Amorim *et al.*, 2018, 2018; Damgaard *et al.*, 2018; de Barros Damgaard *et al.*, 2018; Ebenesersdóttir *et al.*, 2018; Fernandes *et al.*, 2018; Fregel *et al.*, 2018, 2018; Krzewińska, Kılınç, *et al.*, 2018; Krzewińska, Kjellström, *et al.*, 2018; Lindo *et al.*, 2018; McColl *et al.*, 2018; Mittnik *et al.*, 2018; Moreno-Mayar, Potter, *et al.*, 2018; Moreno-Mayar, Vinner, *et al.*, 2018; Scheib *et al.*, 2018; Valdiosera *et al.*, 2018; Veeramah *et al.*, 2018; Olalde *et al.*, 2019,

2019; González-Fortes *et al.*, 2019; Haber *et al.*, 2019; Saag *et al.*, 2019; Sikora *et al.*, 2019; Antonio *et al.*, 2019; Brace *et al.*, 2019, 2019; Järve *et al.*, 2019; Jensen *et al.*, 2019; Kanzawa-Kiriyama *et al.*, 2019; Kashuba *et al.*, 2019; Malmström *et al.*, 2019; Narasimhan *et al.*, 2019; Ning *et al.*, 2019; Sánchez-Quinto *et al.*, 2019; Schroeder *et al.*, 2019; Cassidy *et al.*, 2020; Fernandes *et al.*, 2020; Ning *et al.*, 2020; Brunel *et al.*, 2020; Coutinho *et al.*, 2020; Furtwängler *et al.*, 2020; Linderholm *et al.*, 2020; Marcus *et al.*, 2020; Margaryan *et al.*, 2020; Seguin-Orlando *et al.*, 2021; Kılınç *et al.*, 2021; Saag *et al.*, 2021; Clemente *et al.*, 2021; Egfjord *et al.*, 2021, 2021; Papac *et al.*, 2021; Posth *et al.*, 2021; Lazaridis *et al.*, 2022; Fischer *et al.*, 2022; Gretzinger *et al.*, 2022; Maróti *et al.*, 2022; Patterson *et al.*, 2022; Rodríguez-Varela *et al.*, 2023; Moots *et al.*, 2023; Allentoft, Sikora, Refoyo-Martínez, *et al.*, 2024; Barrie *et al.*, 2024) along with our newly generated 578 genomes using s GLIMPSE v1.1.1 (Rubinacci et al., 2021), as described in (Allentoft, Sikora, Refoyo-Martínez, *et al.*, 2024, Mota et al., 2022). As a reference panel, we used the 30X coverage version of the 1000 Genomes v5 phase 3 dataset (Auton *et al.*, 2015) that was phased using using the TOPMed panel (Taliun *et al.*, 2021), before being lifted over to hg19.

Whole-genome shotgun sequenced genomes and genome-wide targeted capture ('1240k') enriched for 1.24 million sites (Mathieson *et al.*, 2015) were included in this study. Studies have shown for this capture panel that off-target sites impute poorly even at higher coverages (Mota *et al.*, 2022), so all off-target sites were excluded for all samples. For the 1240k on-target sites, imputation performs similarly for genomes of 0.1X-0.5X for shotgun and 1X for capture. Previous studies have demonstrated shotgun samples at 0.1X are suitable for the same methods of demographic inference, even when applied on (Mesolithic and Neolithic) populations that are much more distantly related to the 1000G reference panel than the more recent samples that are the focus of this study (Allentoft, Sikora, Fischer, *et al.*, 2024; Allentoft, Sikora, Refoyo-Martínez, *et al.*, 2024). As such, coverage cut-offs were set at 0.1X for shotgun sequenced, and 1X for captured genomes. After this initial filtering, for each sample, the average genotype probability was calculated for all captured sites, and samples with a low average genotype probability (<0.90) were excluded (Figure S5.1). The final number of samples was 4,507.

The intersection of the 1.24 million capture sites and imputed 1000G sites left 1,085,103 SNPs. A series of filters were also applied to remove problematic sites. We kept sites with imputation score > 0.5 only covering the highly mappable regions in the genomes defined by the 1000 Genomes strict mappability mask leaving the final number of SNPS at 697,179.

In comparison, a study with a similar processing pipeline that only included shotgun sequenced ancient genomes resulted in a panel with over 10 times as many SNPs: 7,321,965 (Allentoft, Sikora, Refoyo-Martínez, *et al.*, 2024). However, for many regions and time periods of importance to this study, restricting to just capture samples would result in entire regions in time (e.g. Western Europe during the Bronze Age) and countries in general (e.g England and the Czech Republic) being under-represented (Figure S5.2). To address the research questions of the study, it was therefore necessary to restrict to the smaller SNP set.

Metadata for the samples was curated from the respective papers and from (Mallick *et al.*, 2023).

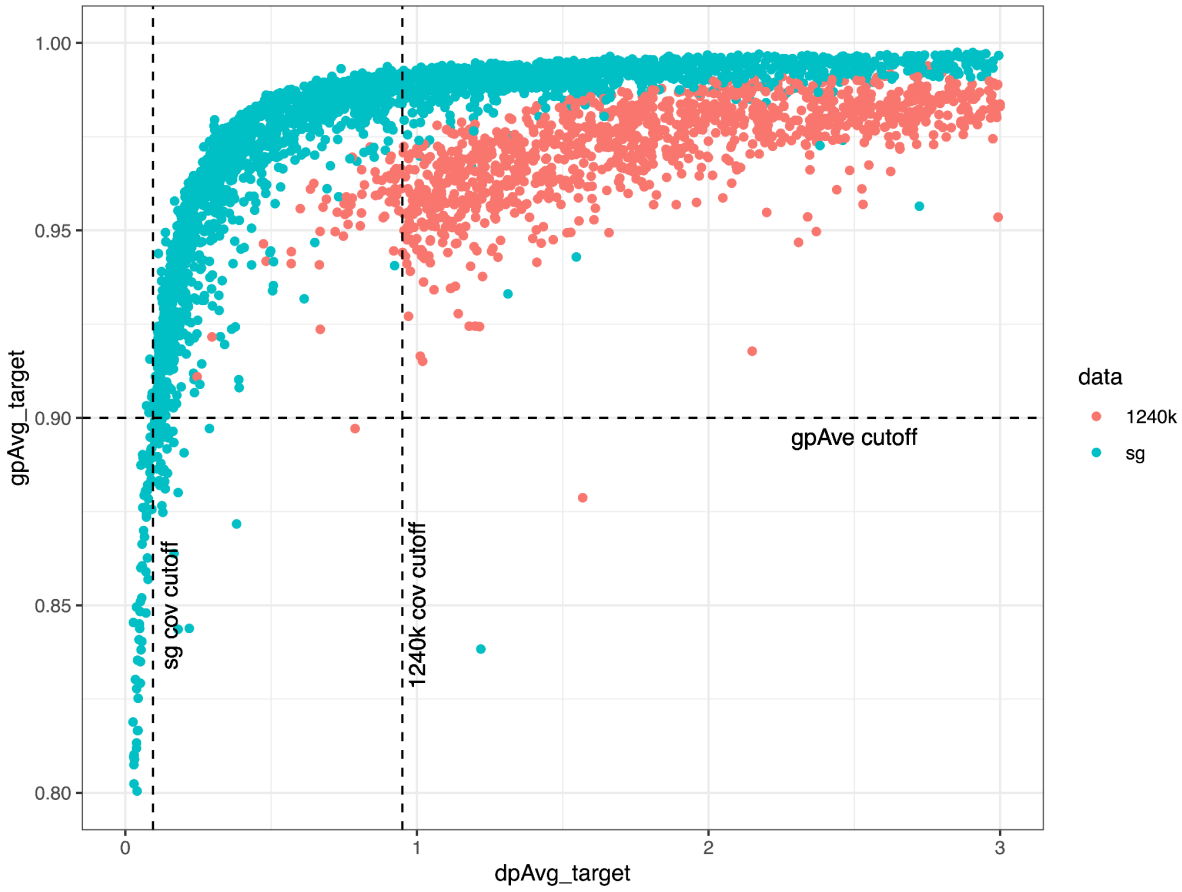


Figure S5.1. Relationship between the average genotype probability (gpAvg_target) and sequencing depth (dpAvg_target) for captured (1240k) and shotgun (sg) genomes, when restricting to target sites.

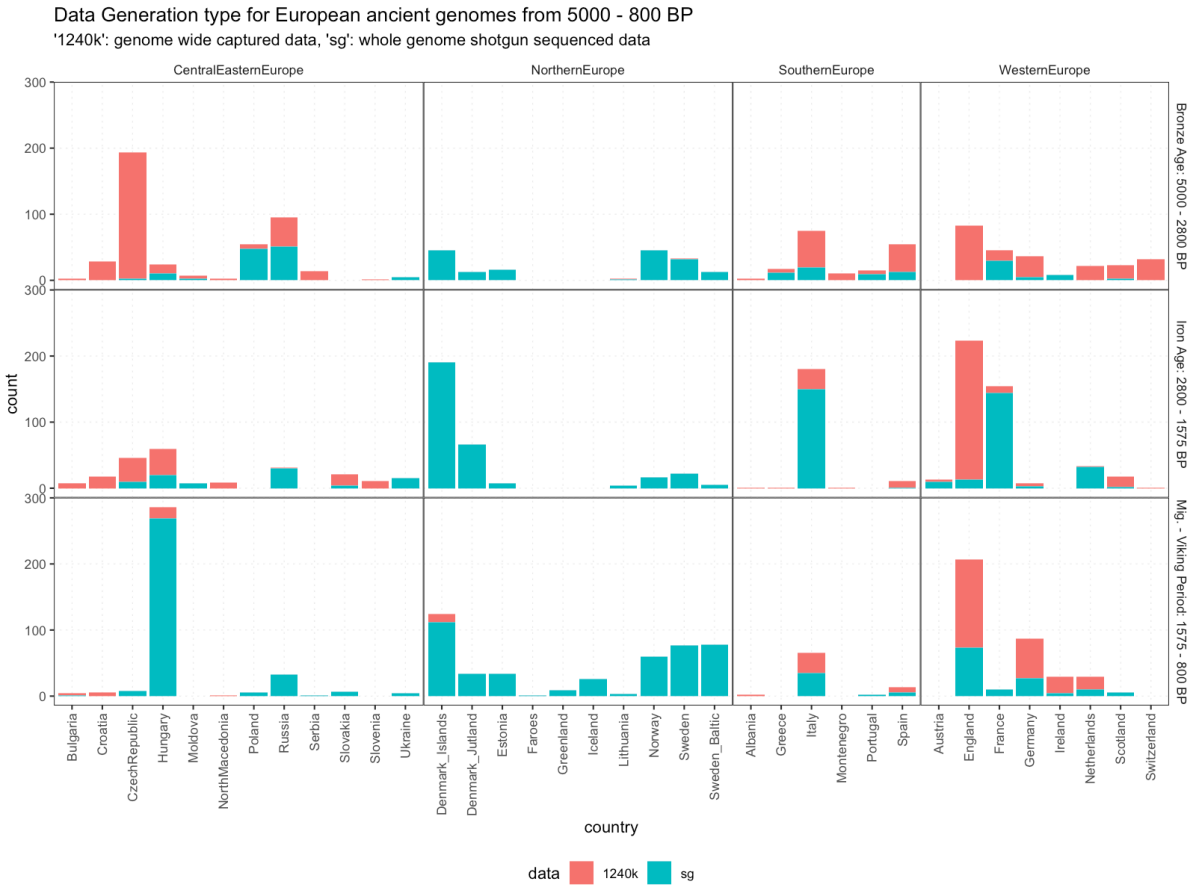


Figure S5.2. Figure showing the geographical and temporal distributions of ancient genomes in the final imputed dataset for each data generation method (1240k capture and shotgun sequencing).

S6. Demographic Inference

S6.1. PCA

To infer basic demographic structure in our data we carried out a principal component analysis (PCA) on the imputed dataset of shotgun genomes (see section S5) using GCTA (v1.94.1). To better visualise IBD clusters and finer scale patterns in the data, we plotted two panels of decreasing size and diversity, the first contains all 'out-of-Africa' populations, (n=4,495) and the second focussing on Western Eurasia (n=3,870). As depicted in Figure S6.1.1, the vast majority of samples from this study (highlighted with a black circle) fall within the European Bronze Age diversity, while a small number of samples display varying levels of Asian admixture.

Colouring the individuals in the PCA by their IBD clusters (Supplementary Note 6.4) shows a correlation between established clines in PCA space and the IBD clustering (Figure S6.1.1). For each of the major four clusters discussed in the main text (Yamnaya, Corded Ware (East), Corded Ware (North), Bell Beaker), a PCA highlighting their subclusters been plotted (Figure S6.1.2-13)

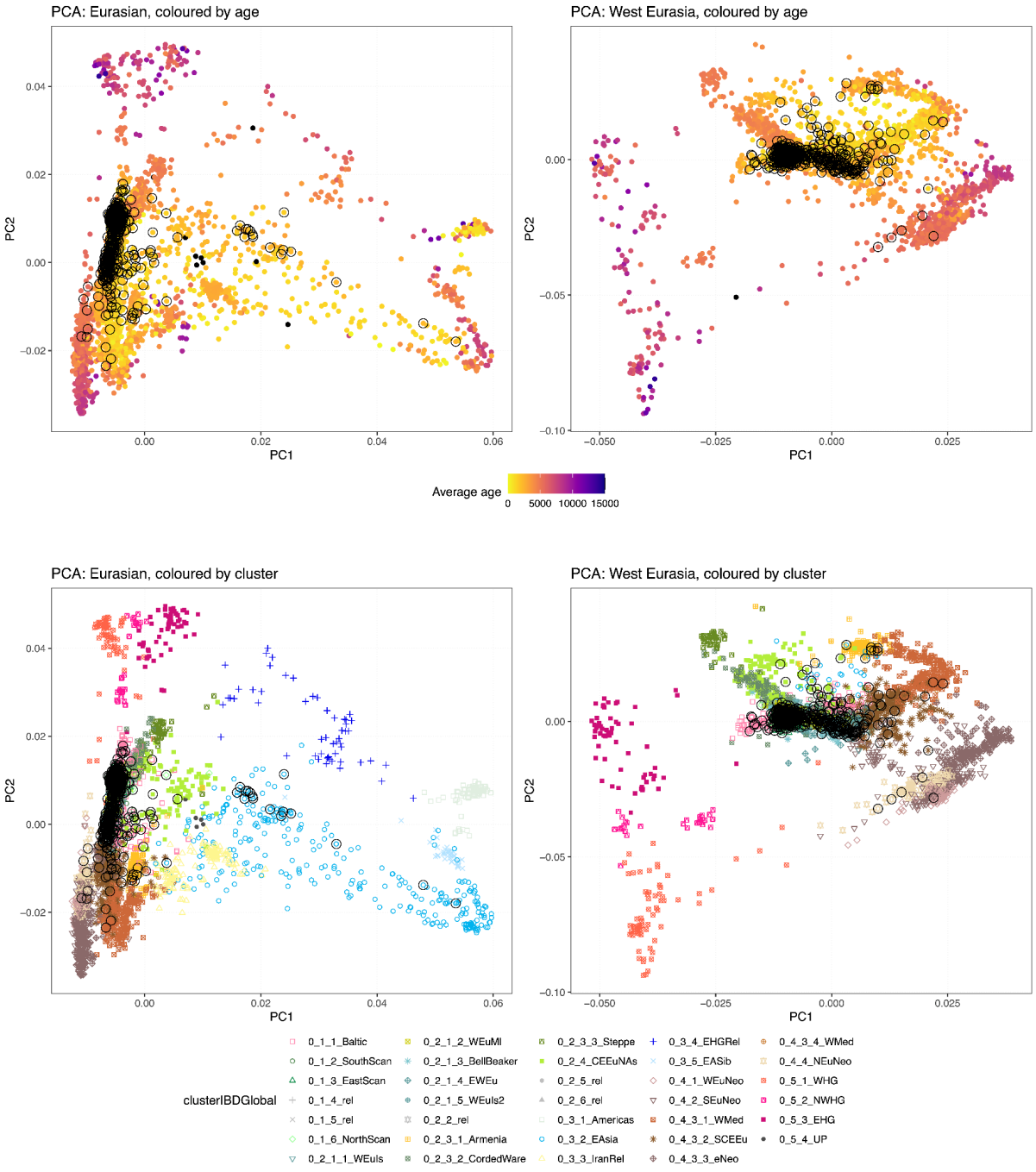


Figure S6.1.1 PCA of Eurasian and Western Eurasian ancient populations, coloured by age and by IBD cluster. Circled individuals are newly generated in this project.

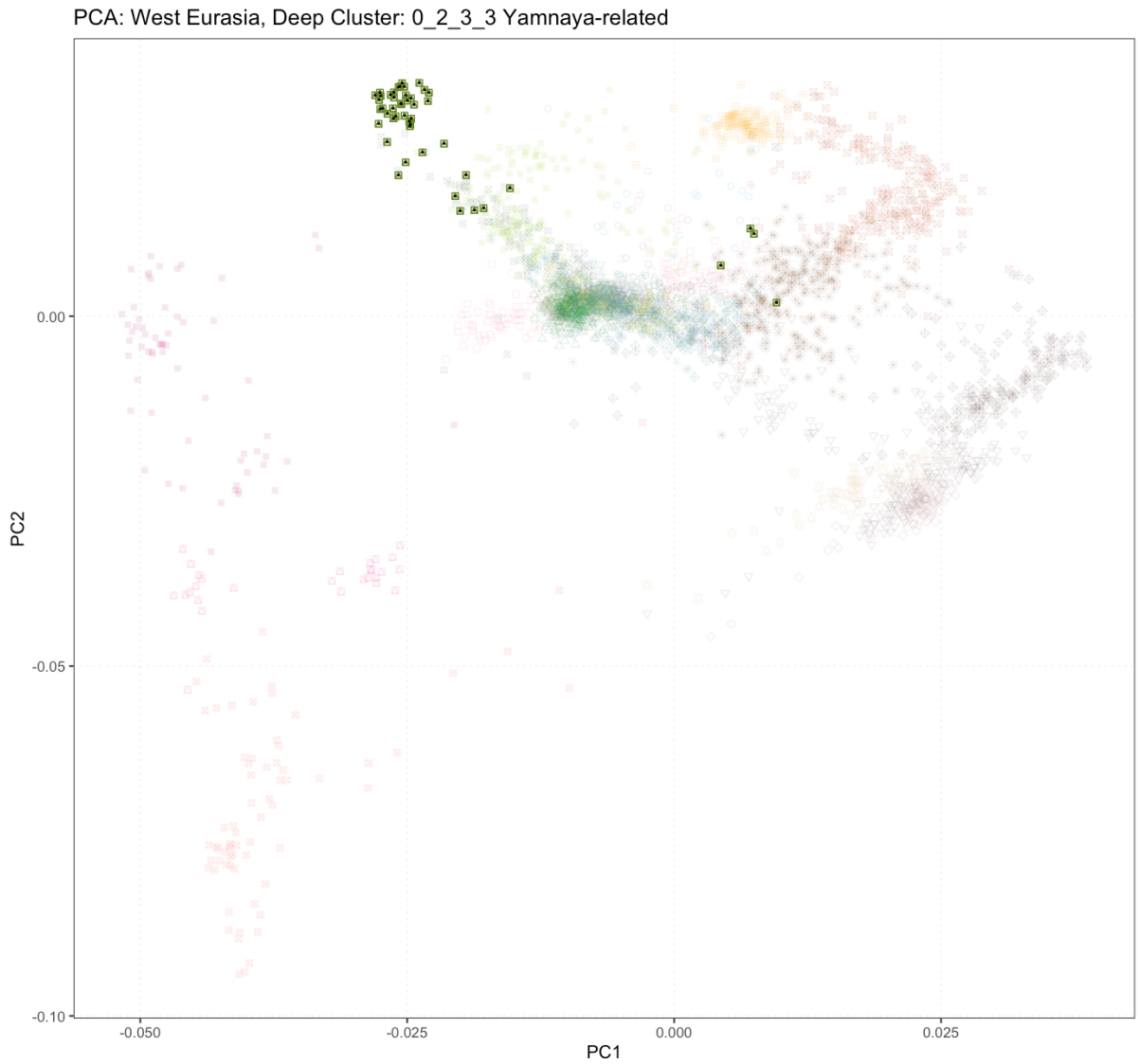
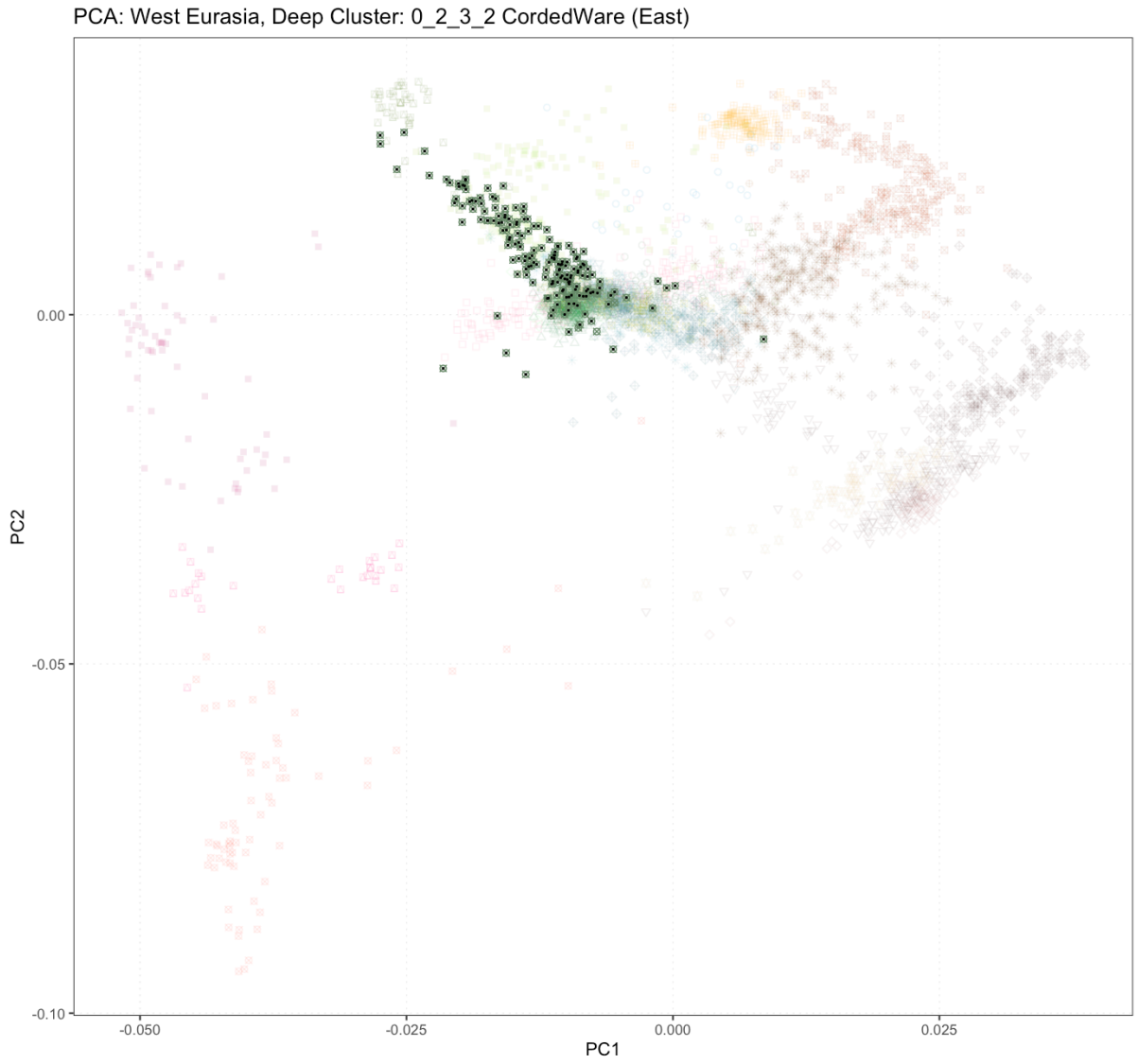


Figure S6.1.2 - Western Eurasian PCA highlighting the Yamnaya-related cluster. Individuals older than 2800 BP are indicated with a '.'



542

543 Figure S6.1.3 - Western Eurasian PCA highlighting the Corded Ware (East)-related ancestry.

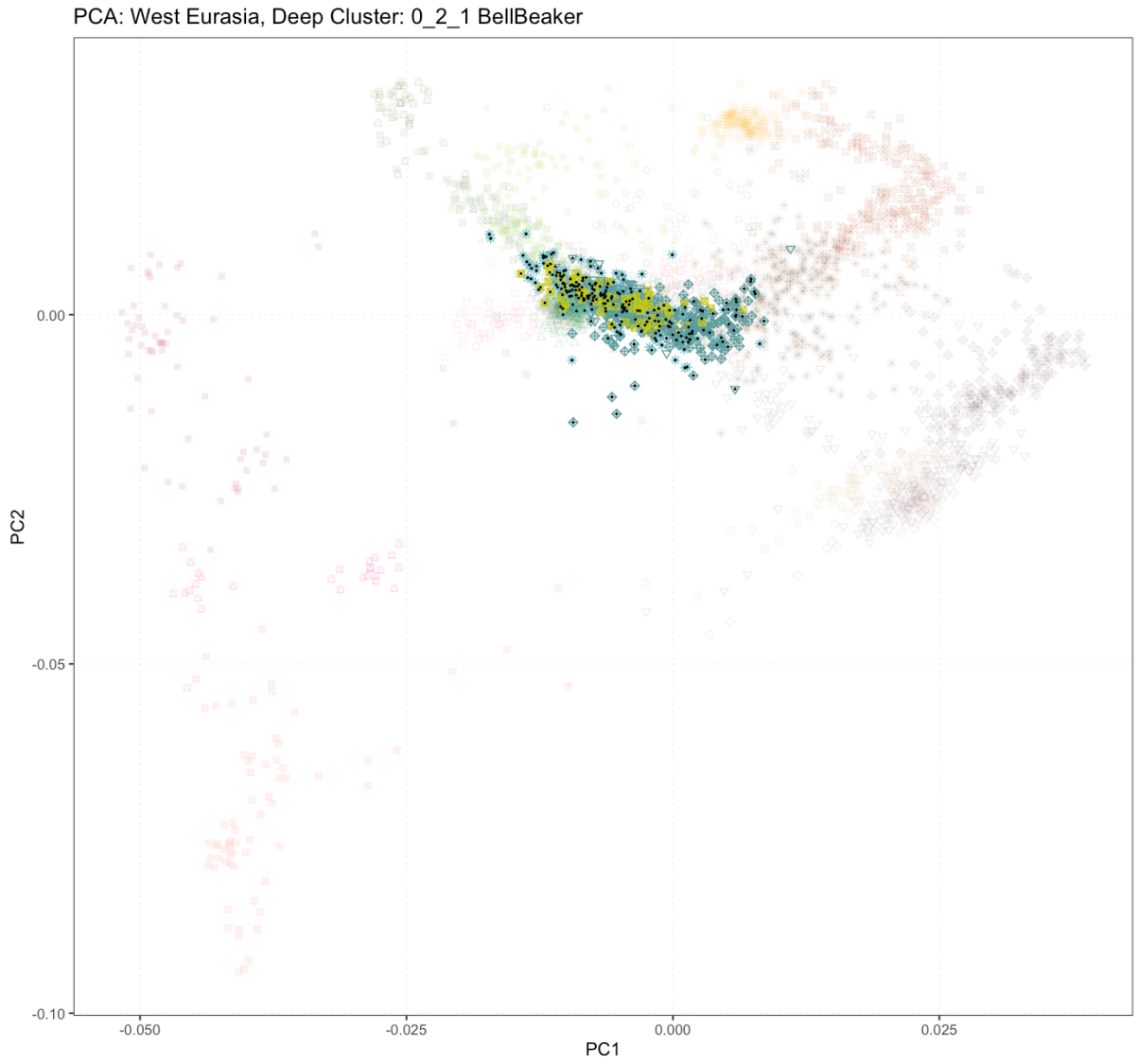
544 Individuals older than 2800 BP are indicated with a '.'

545

546

547

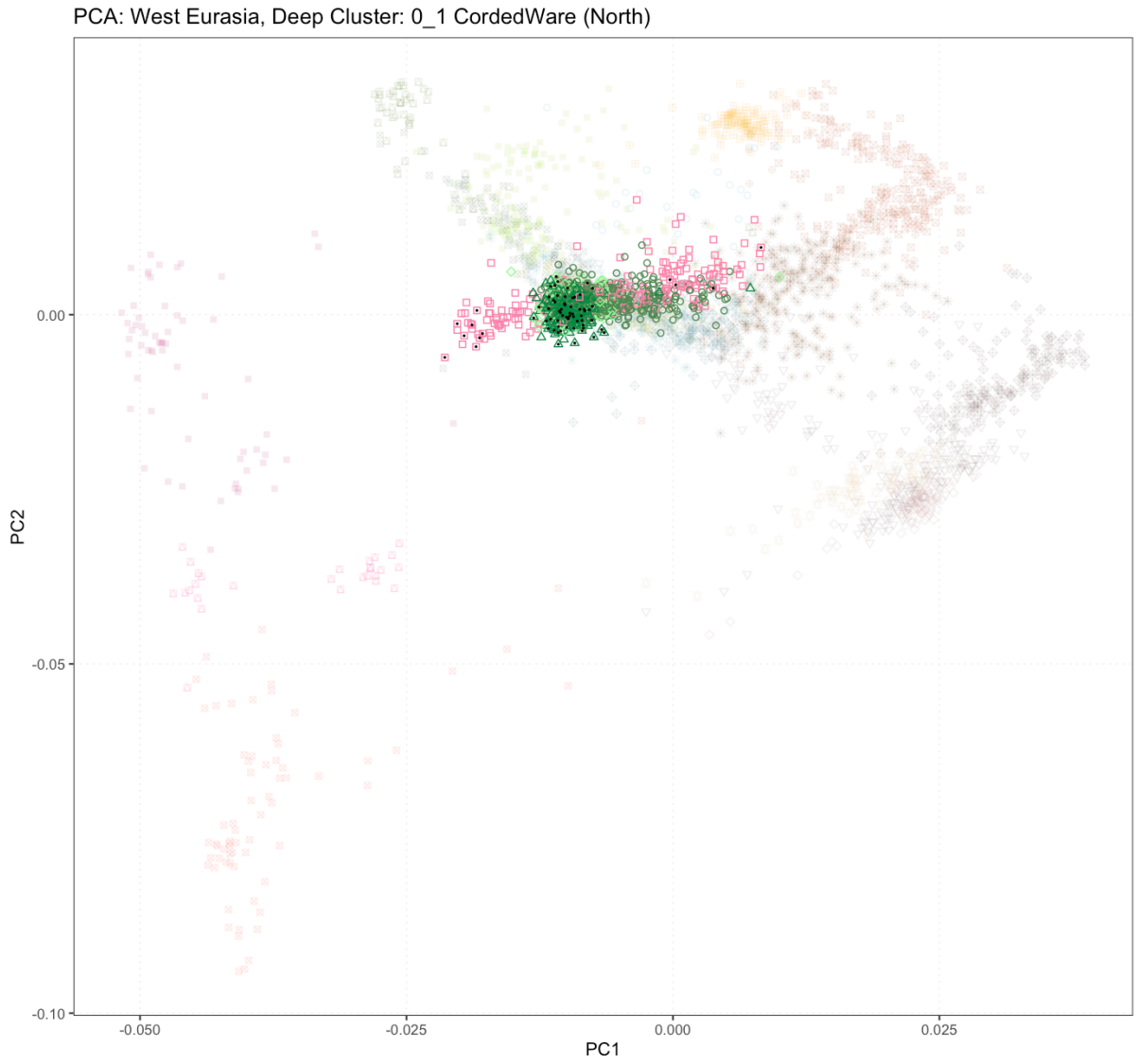
548



549

550 Figure S6.1.4. Western Eurasian PCA highlighting the Bell Beaker-related ancestry.
551 Individuals older than 2800 BP are indicated with a '.'

552



553

554 Figure S6.1.5. Western Eurasian PCA highlighting the Corded Ware (North)-related ancestry.

555 Individuals older than 2800 BP are indicated with a '.'

556

557

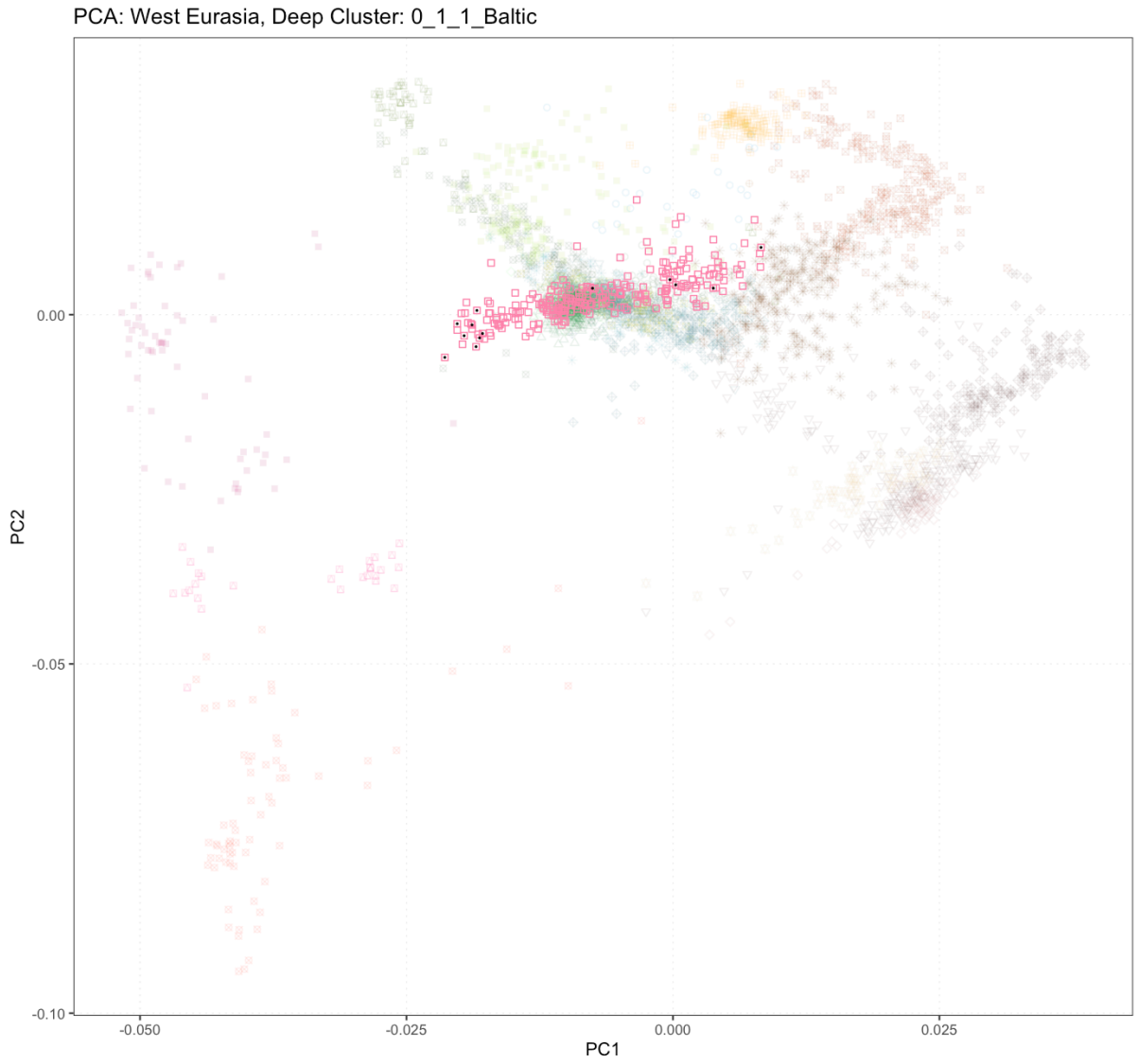


Figure S6.1.6. Western Eurasian PCA highlighting the Baltic subcluster of the Corded Ware (North) cluster. Individuals older than 2800 BP are indicated with a '.'

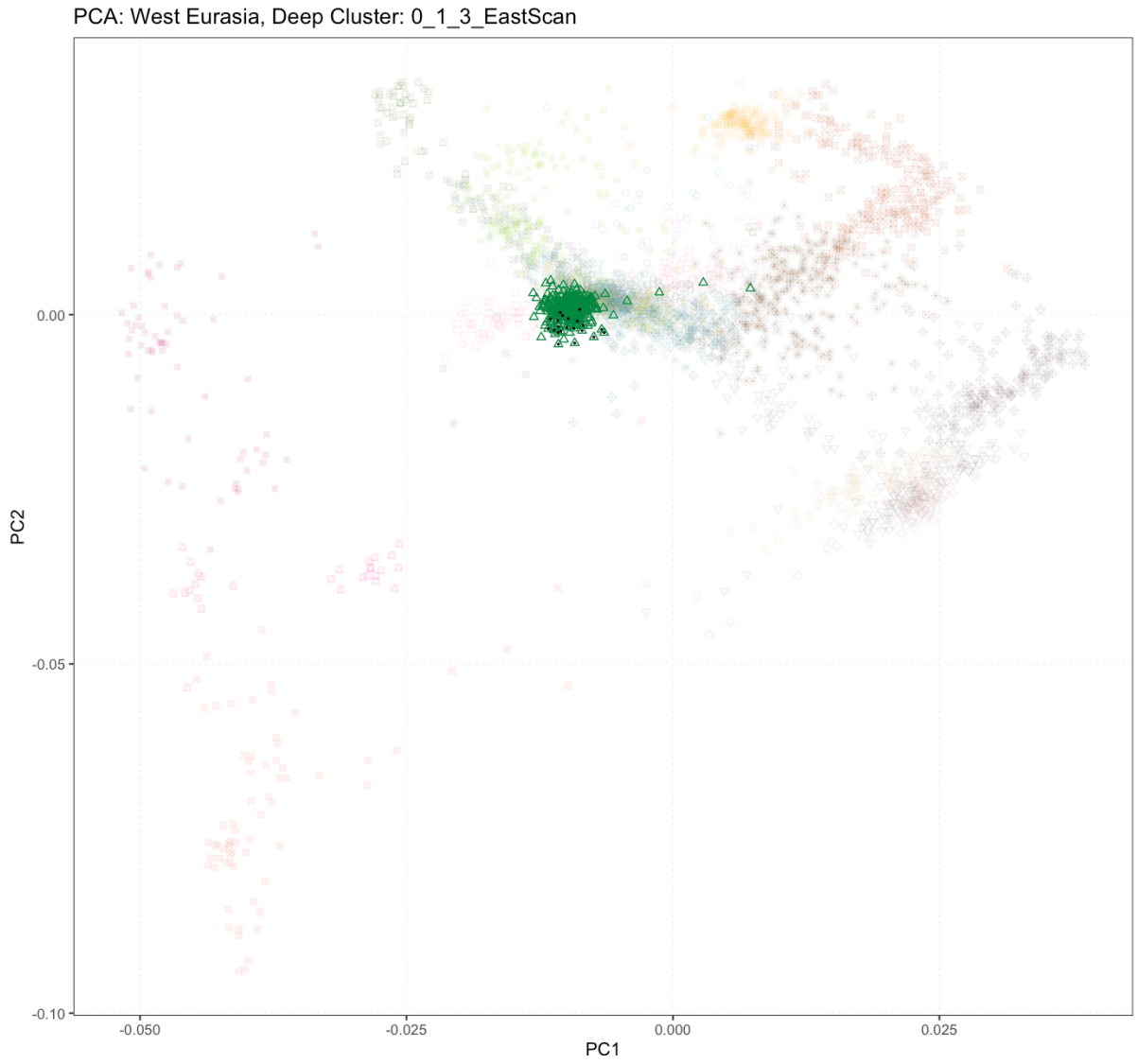
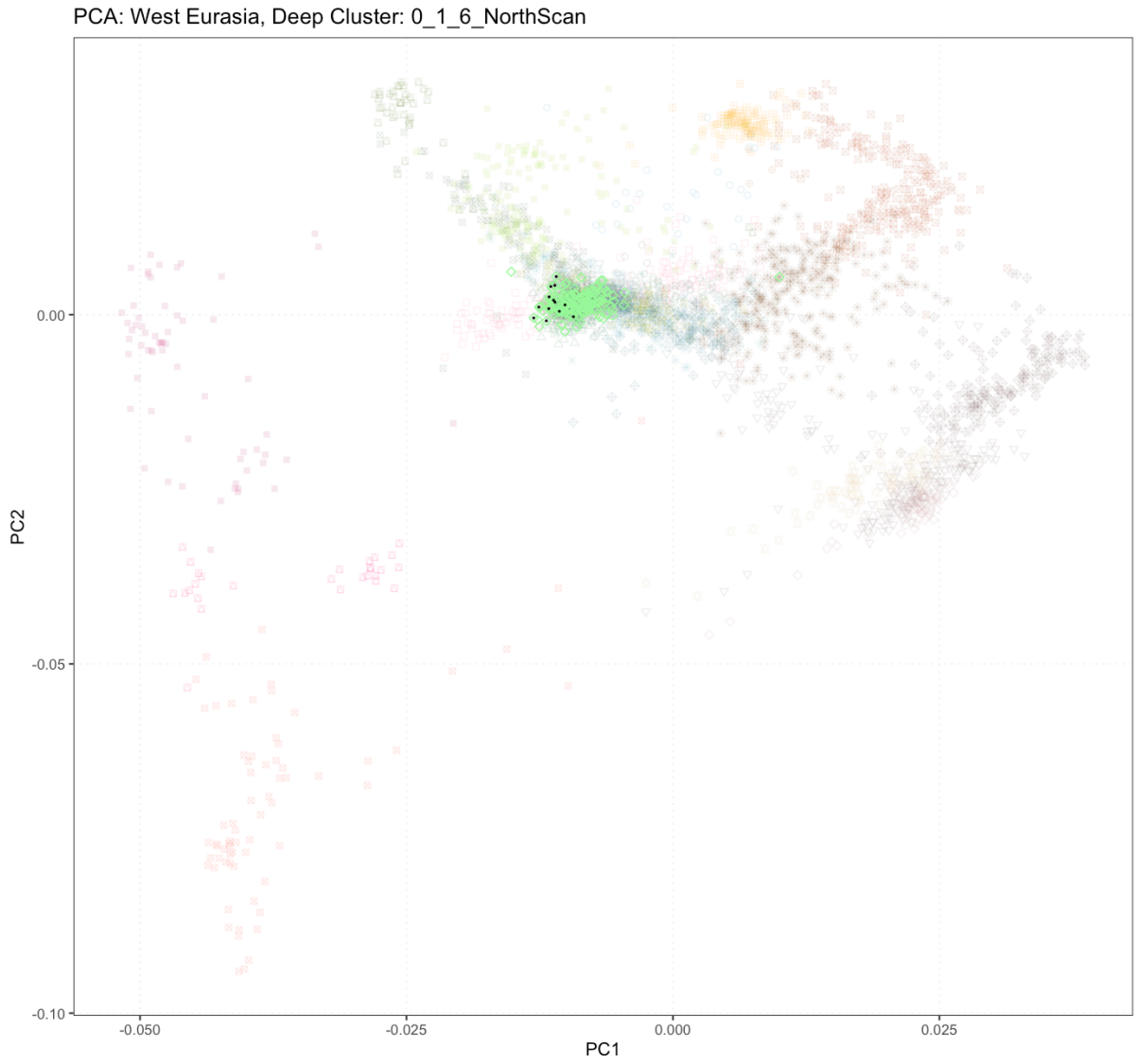


Figure S6.1.7. Western Eurasian PCA highlighting the Eastern Scandinavian subcluster of the Corded Ware (North) cluster. Individuals older than 2800 BP are indicated with a '.'



570

571 Figure S6.1.8 Western Eurasian PCA highlighting the Western Scandinavian subcluster of the
572 Corded Ware (North) cluster. Individuals older than 2800 BP are indicated with a '.'

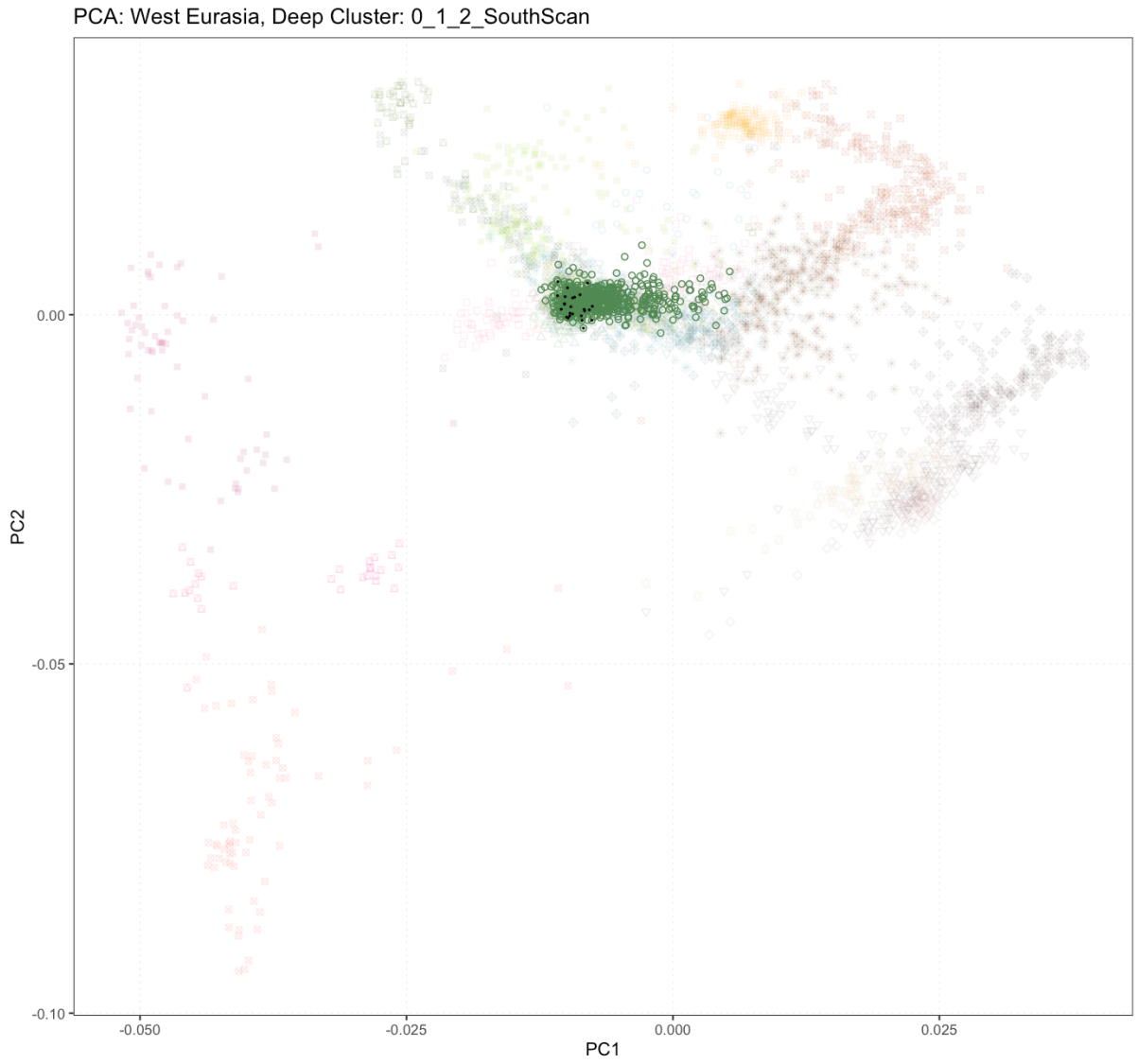
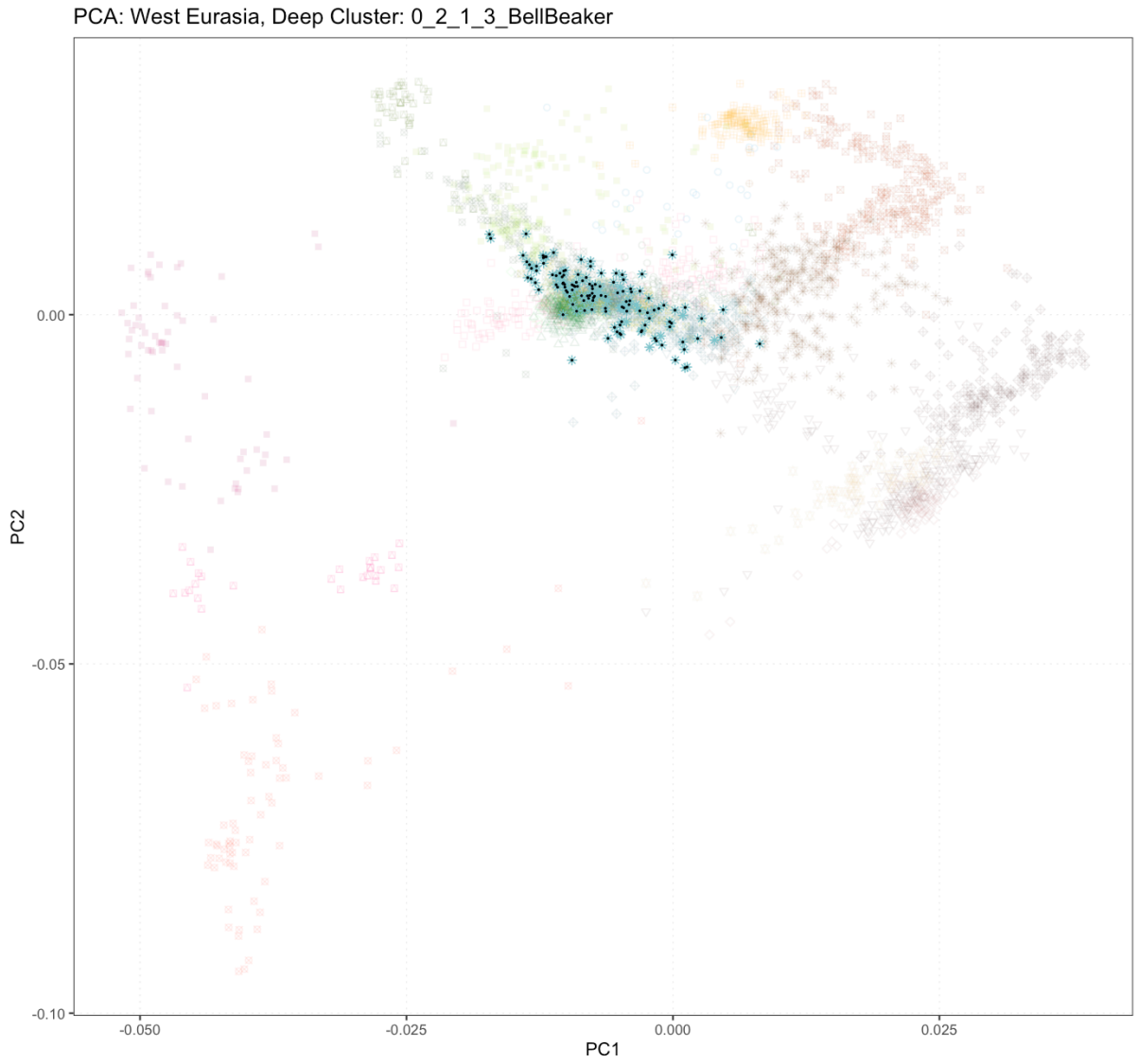


Figure S6.1.9. Western Eurasian PCA highlighting the Southern Scandinavian subcluster of the Corded Ware (North) cluster. Individuals older than 2800 BP are indicated with a '.'

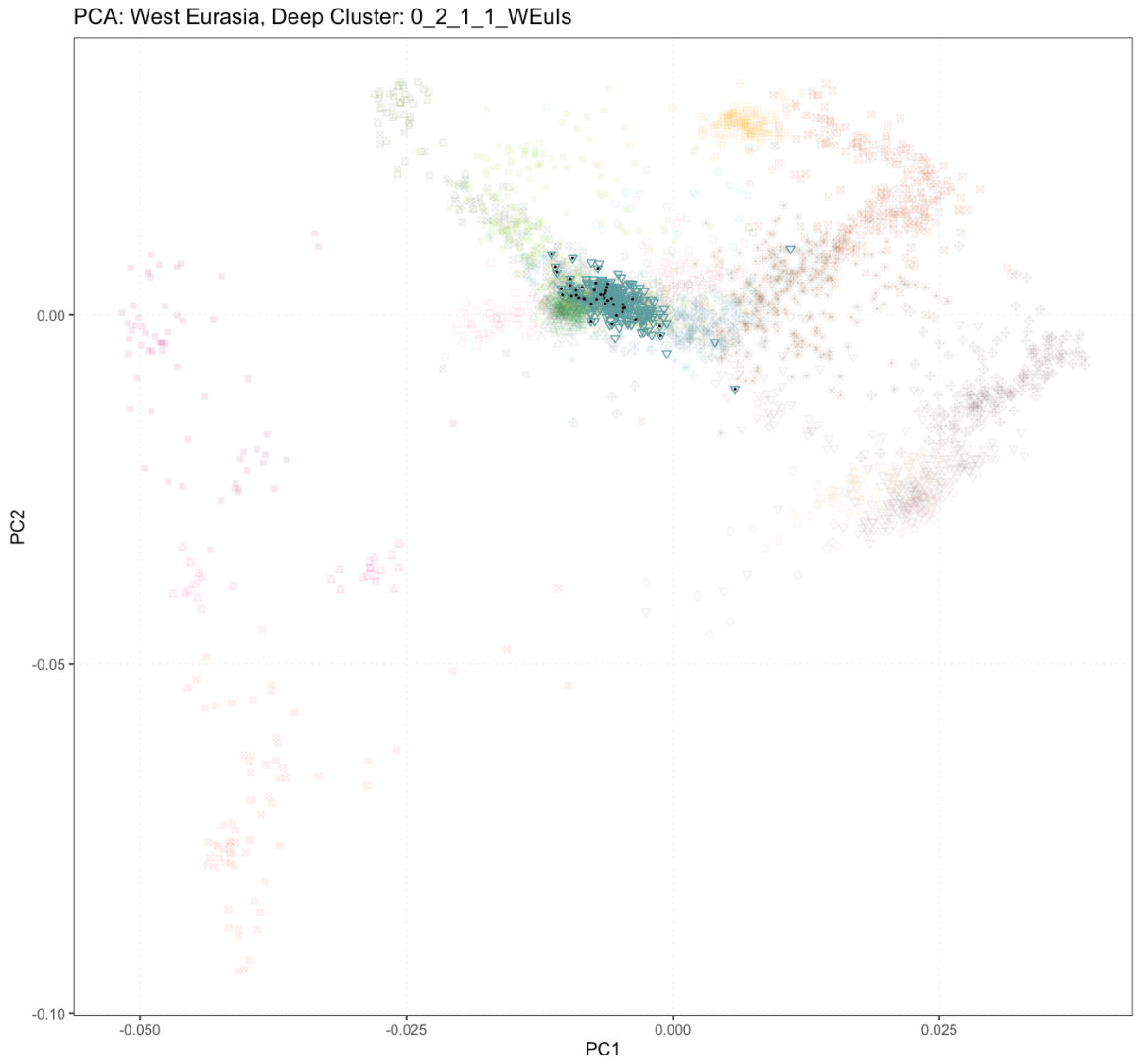


579

580 Figure S6.1.10. Western Eurasian PCA highlighting the Early Bell Beaker subcluster of the

581 Bell Beaker cluster. Individuals older than 2800 BP are indicated with a '.'

582



583

584 Figure S6.1.11. Western Eurasian PCA highlighting the *Western European Insular BellBeakers*
585 subcluster of the Bell Beaker cluster. Individuals older than 2800 BP are indicated with a '.'

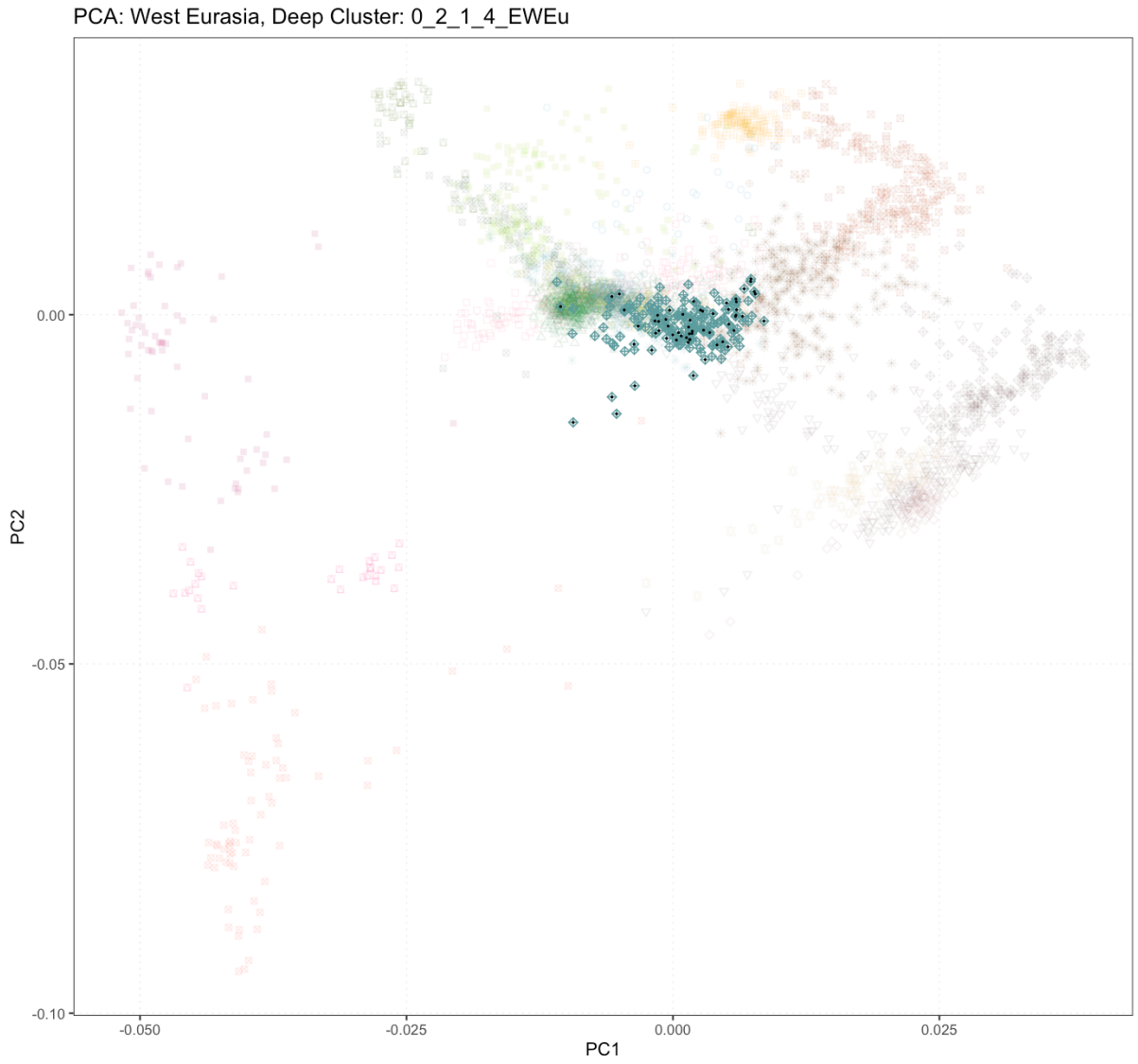


Figure S6.1.12. Western Eurasian PCA highlighting the *Hallstatt / La Tene Bell Beakers* subcluster of the Bell Beaker cluster. Individuals older than 2800 BP are indicated with a '.'

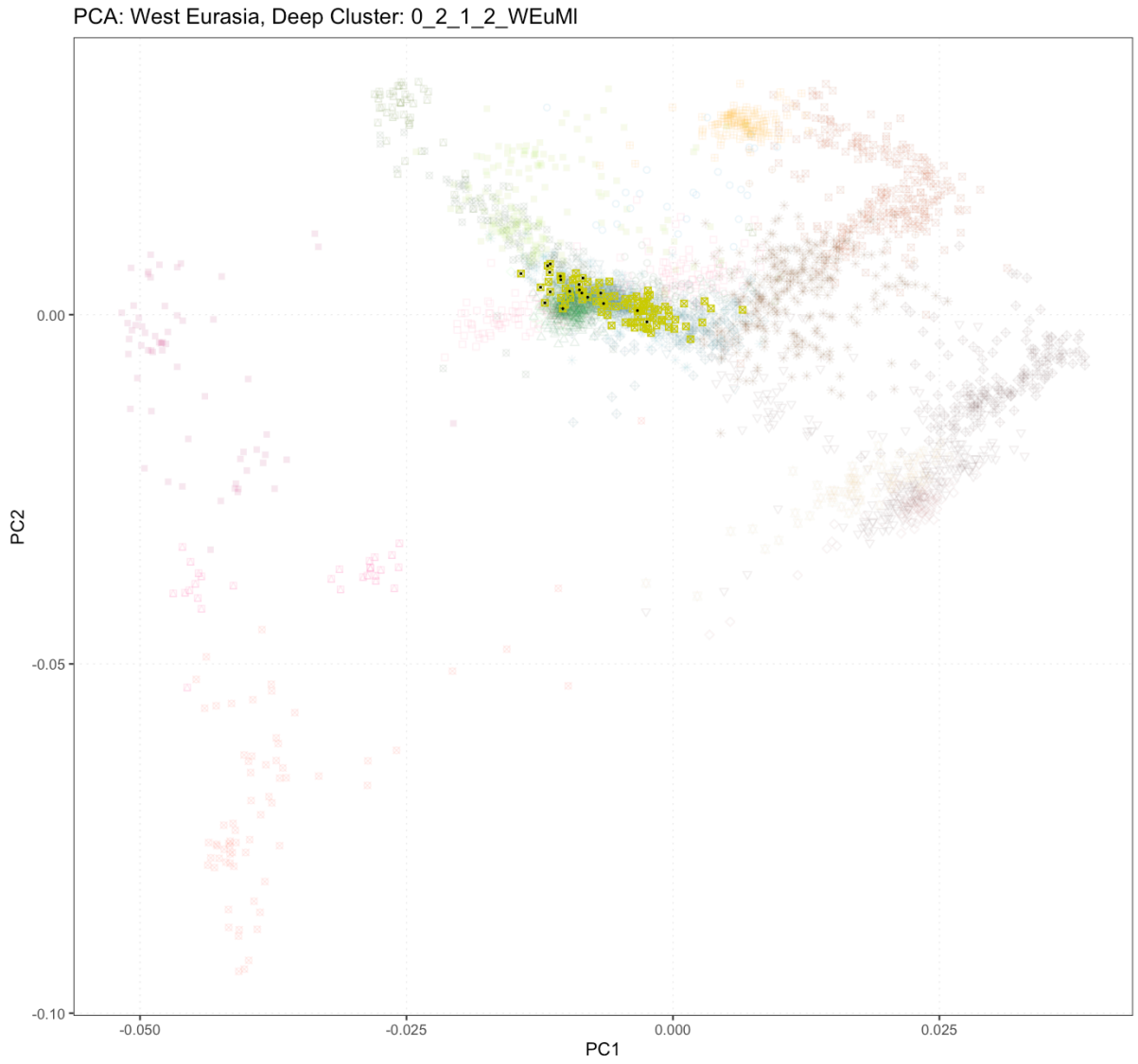


Figure S6.1.13. Western Eurasian PCA highlighting the *Eastern Northern Sea Bell Beakers* subcluster of the Bell Beaker cluster. Individuals older than 2800 BP are indicated with a '.'

S6.2. IBD Clustering

S6.2.1. Whole Datasets Clustering

To identify IBD segments, we ran IBDseq (Browning and Browning, 2013) on the imputed panel. We removed any segments less than 2cM and any hotspot regions with excess sharing across all individuals, as described in (Allentoft, Sikora, Refoyo-Martínez, *et al.*, 2024). For all pairs, the total shared IBD length was summed and number of IBD segments counted. We excluded any pairs sharing less than a total of 5cM, and the lowest coverage of any pair sharing

more than 1500 cM and had less than 150 segments, to avoid the formation of small clusters with only 1st and 2nd degree relatives. We then performed a network-based hierarchical clustering based on total shared IBD lengths (Greenbaum *et al.*, 2019), as described in (Allentoft, Sikora, Refoyo-Martínez, *et al.*, 2024).

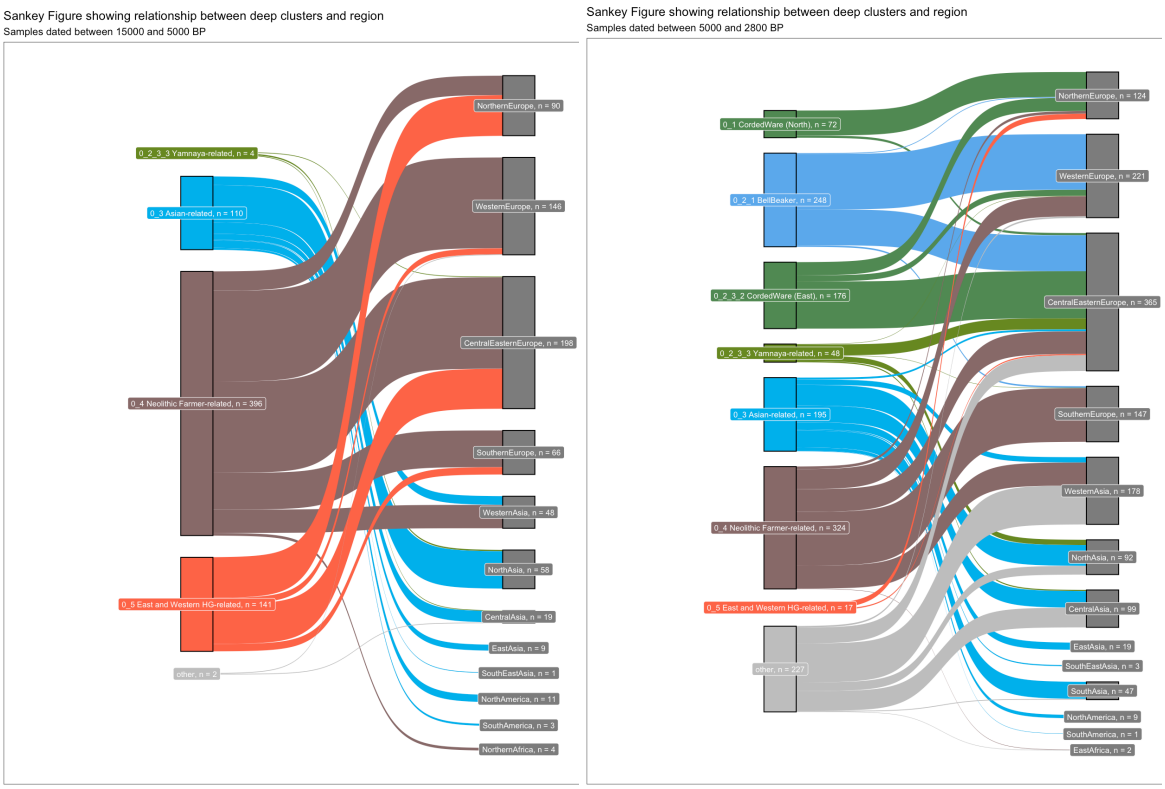
By plotting the deepest clusters on a map (Figure S6.2.1.4) and creating Sankey diagrams showing the relation between clusters and regions (Figure S6.2.1.1, S6.2.1.2, S6.2.1.3), and looking at IBD mixture modelling results faceted by cluster (Figure S6.5.2), clear patterns are apparent.

- Cluster 0_1 and subclusters (indicated by 0_1_x…)
    - high Steppe Ancestry and some Globular Amphora Culture (GAC) ancestry
    - found in Northern Europe
    - associated with the Corded Ware Culture.
- Cluster 0_2_1 and subclusters
    - high Steppe Ancestry and some GAC ancestry and varying amounts of European Farmer ancestry
    - found in Western Europe,
    - associated with Bell Beaker contexts.
- Cluster 0_2_3 and subclusters
    - highest Steppe Ancestry, including Yamnaya, found in Western Eurasia
- Cluster 0_3 and subclusters
    - from across Asia and the Americas
- Cluster 0_4 and subclusters
    - high Neolithic Farming ancestry
- Cluster 0_5 and sub clusters
    - Eastern and Western Hunter gatherer ancestry

From 5000 - 1575 BP, the Northern Corded Ware cluster was restricted primarily to Northern Europe, and the Bell Beaker cluster to Western Europe (Figure S6.2.1.1). From 1575 BP onwards the Northern Corded Ware cluster was more widespread. By using more fine scale clusters (Figure S6.2.1.2), it is apparent this spread is primarily by the Southern Scandinavian cluster. When replacing regions with countries (Figure S6.2.1.3), the widespread appearance

of Southern Scandinavian in England, Germany, and the Netherlands, is apparent. From 5000-2800 BP, individuals from these countries were primarily from Bell Beaker-related clusters.

643



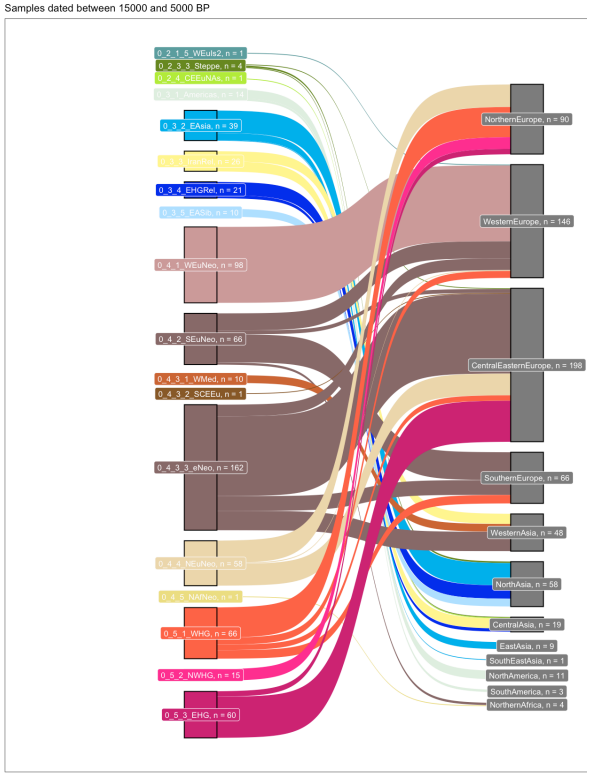
644

645 Figure S6.2.1.1. Sankey diagram showing the relationship between the deepest IBD clusters
646 and geographical regions around the world, in different time bins (15000 - 5000 BP, 5000 -
647 2800 BP, 2800 - 1575 BP, 1575 - 0 BP)

648

Sankey Figure showing relationship between deep clusters and region

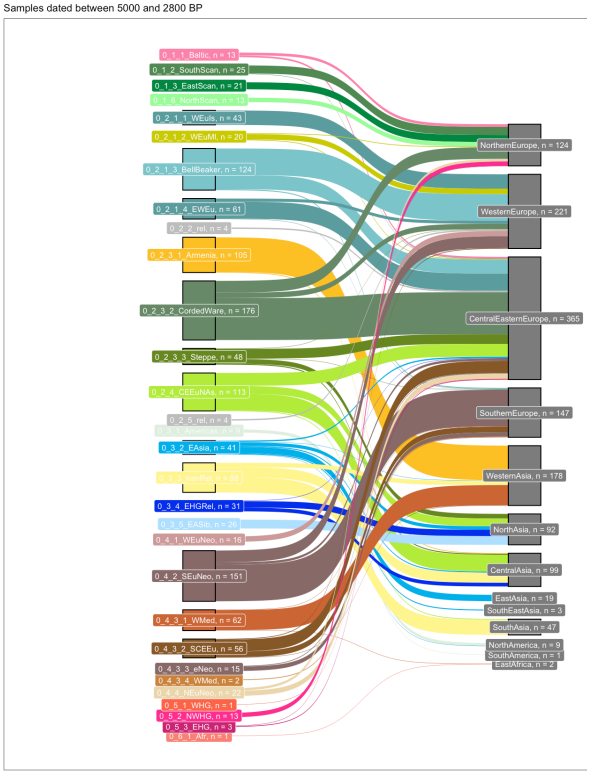
Samples dated between 15000 and 5000 BP



x

Sankey Figure showing relationship between deep clusters and region

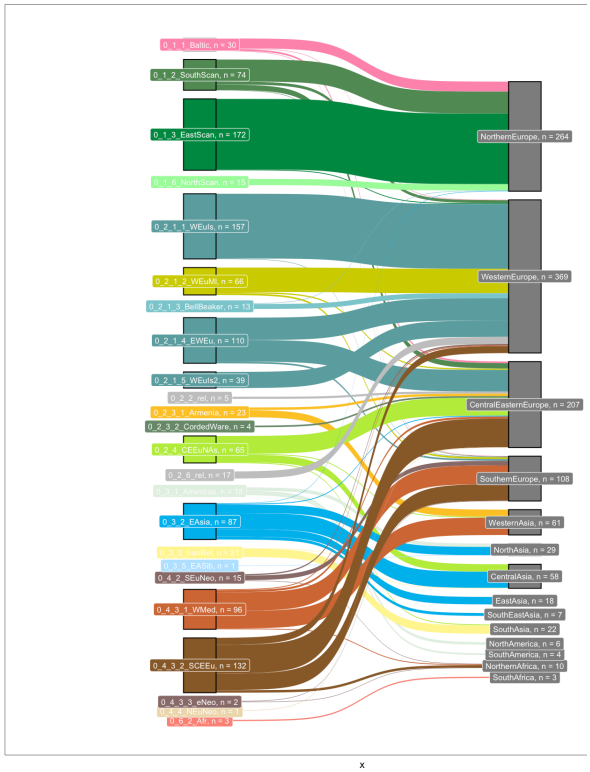
Samples dated between 5000 and 2800 BP



x

Sankey Figure showing relationship between deep clusters and region

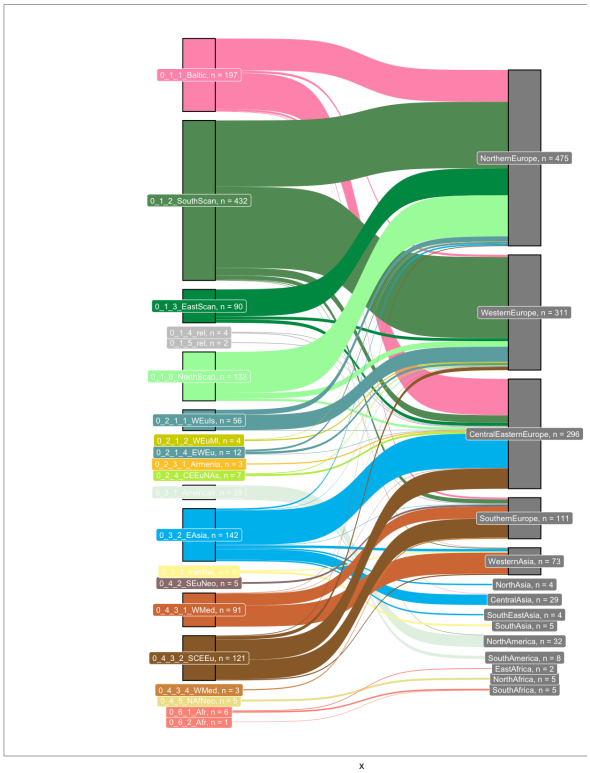
Samples dated between 2800 and 1575 BP



x

Sankey Figure showing relationship between deep clusters and region

Samples dated between 1575 and 0 BP

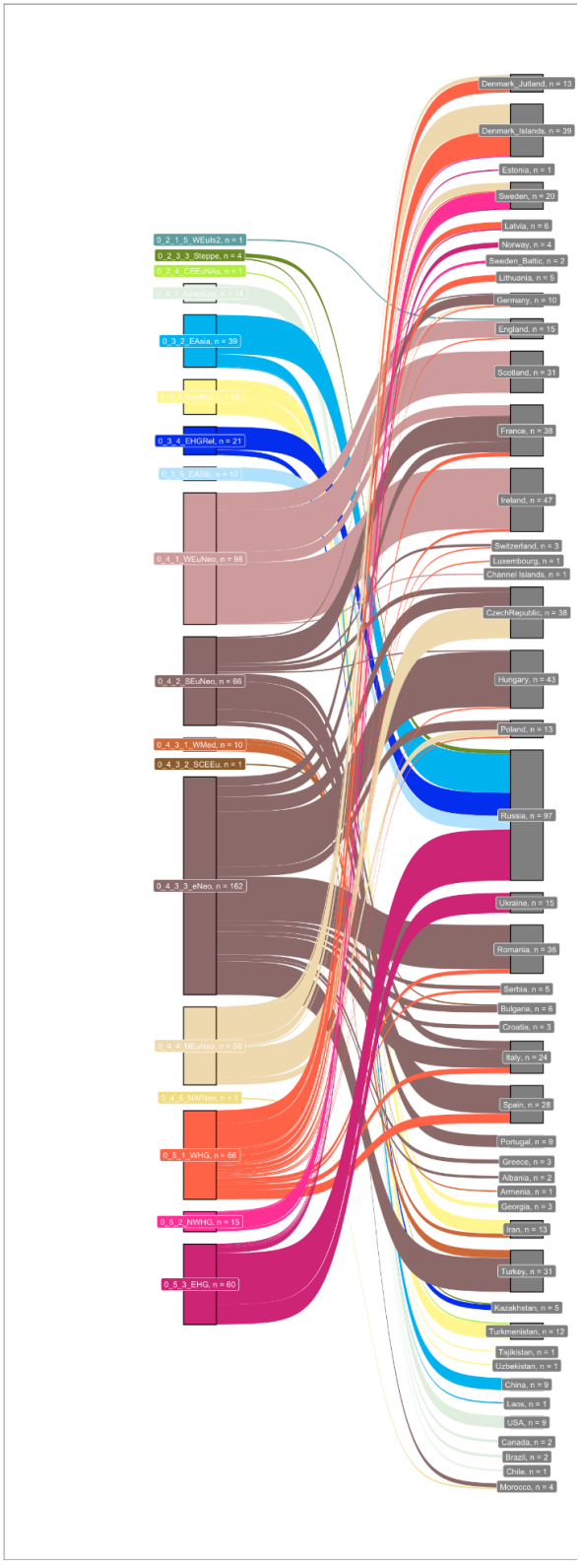


x

Figure S6.2.1.2. Sankey diagram showing the relationship between the deep IBD clusters and geographical regions around the world, in different time bins (15000 - 5000 BP, 5000 - 2800 BP, 2800 - 1575 BP, 1575 - 0 BP)

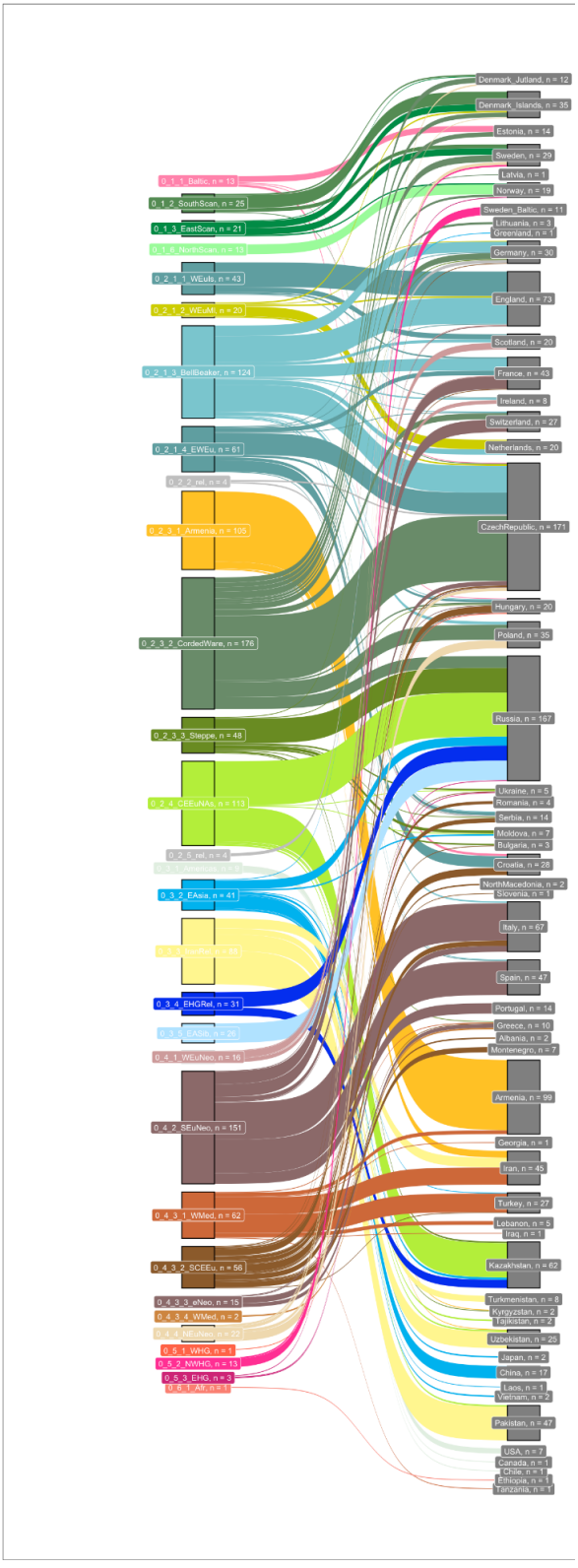
Sankey Figure showing relationship between deep clusters and country

Samples dated between 15000 and 5000 BP



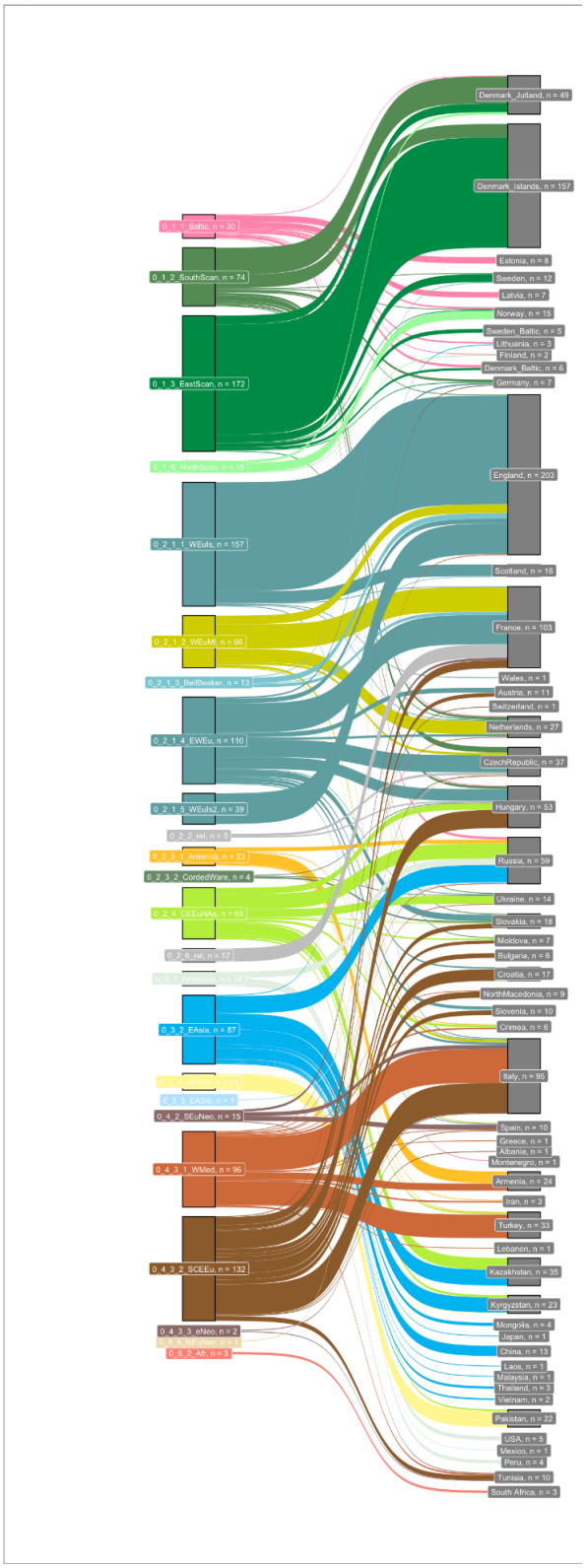
Sankey Figure showing relationship between deep clusters and country

Samples dated between 5000 and 2800 BP



Sankey Figure showing relationship between deep clusters and country

Samples dated between 2800 and 1575 BP



Sankey Figure showing relationship between deep clusters and country

Samples dated between 1575 and 0 BP

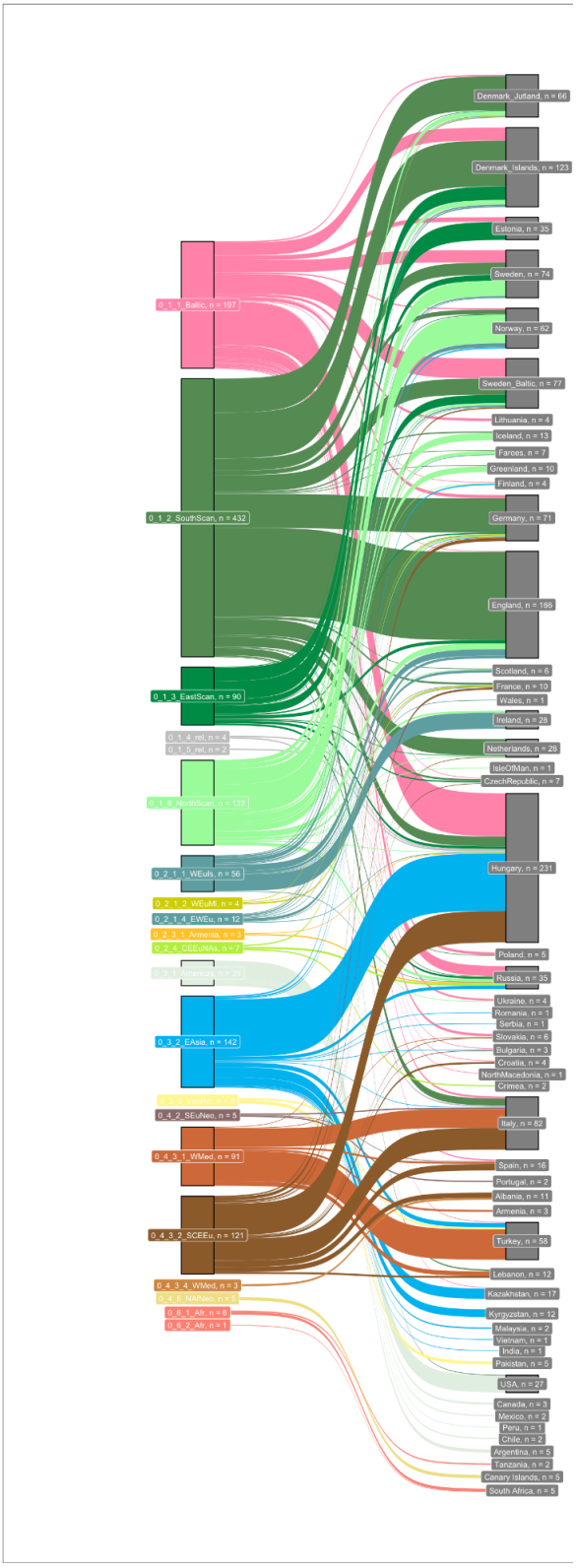


Figure S6.2.1.3. Sankey diagram showing the relationship between the deepest IBD clusters and countries around the world, in different time bins (15000 - 5000 BP, 5000 - 2800 BP, 2800 - 1575 BP, 1575 - 0 BP)

PDF UPLOADED SEPARATELY

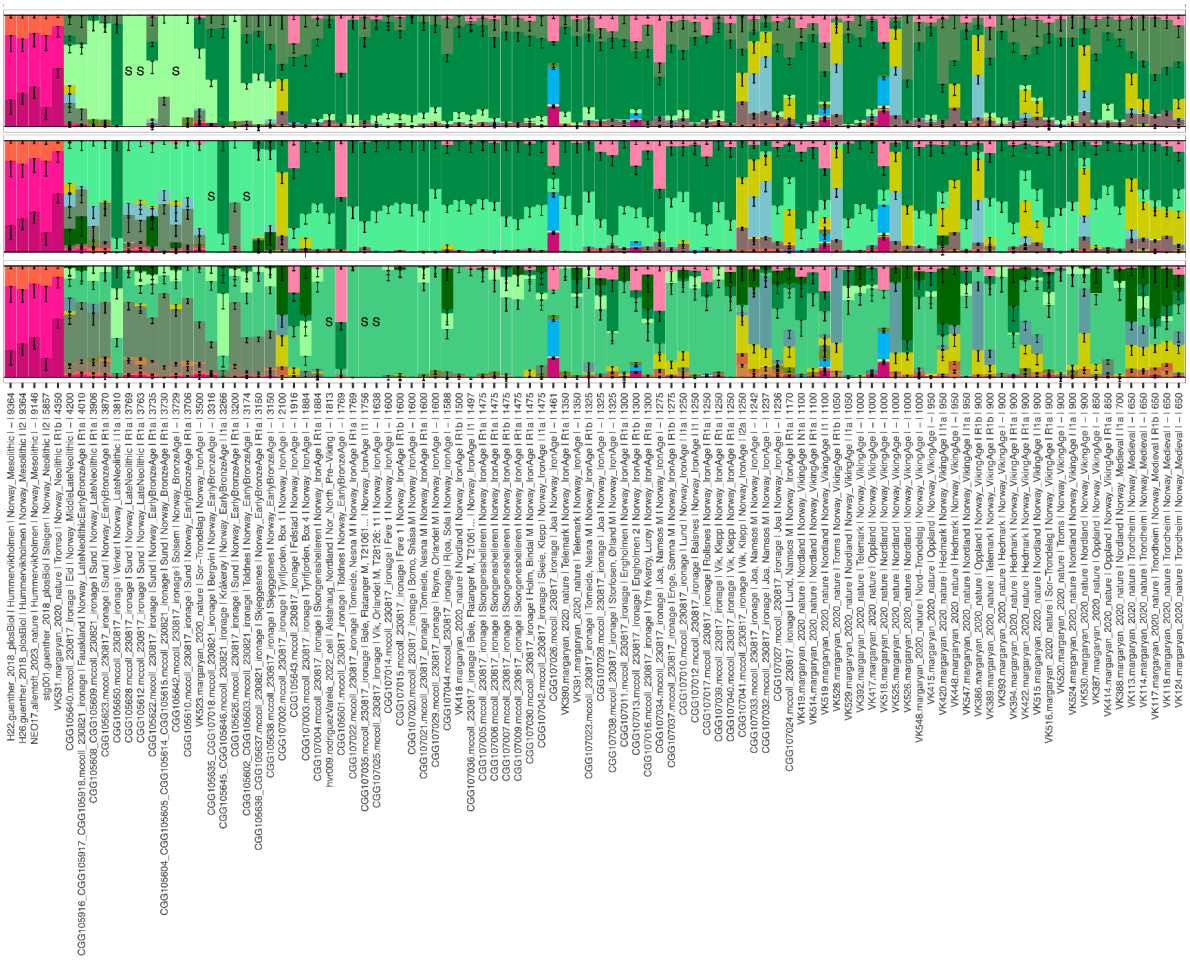
Figure S6.2.1.4. Plots showing deep clusters plotted across Western Eurasia

S6.2.2. 2800 + Datasets Clustering

Although the clustering represents genetic structure that corresponds to specific regions and times, basing results on the relationship between sub-clusters alone can be misleading. For example, the 0_1_6 Western Scandinavian cluster contains a series of subclusters, predominantly of Norwegian samples from the Bronze to the Viking Period. However, by using a set of sources including Bronze Age Western and Eastern Scandinavians source clusters, we find that the late Bronze Age Norwegian individuals are modelled as ~50% Eastern - 50% Western Scandinavian. By the Iron Age, most individuals are modelled with <10% Western Scandinavian and >90% Eastern Scandinavian ancestry, despite all falling within the 0_1_6 Western Scandinavian broader cluster. When using the late Bronze Age Western Scandinavians (modelled as 50% Western Scandinavian) as a source, we find the model 50% of the ancestry of the later Iron Age, which appears to create the link joining the clusters (Figure S6.2.2.1).

By excluding all individuals less dated to less than 2800BP (pre2.8k) and reclustering the dataset as in S6.2.1, we are able to identify the Bronze Age population structure without the impact of later admixture (Table S3, Figure S6.2.2.4). In the case of the Bronze Age Norwegians, we find in the pre2.8k reclustering they form a cluster together with the earliest

Swedish and Danish individuals (Figure 3), many of which previously feel in the Corded Ware (East) subcluster, and who are predominantly carriers of the R1a y-haplogroup.



B

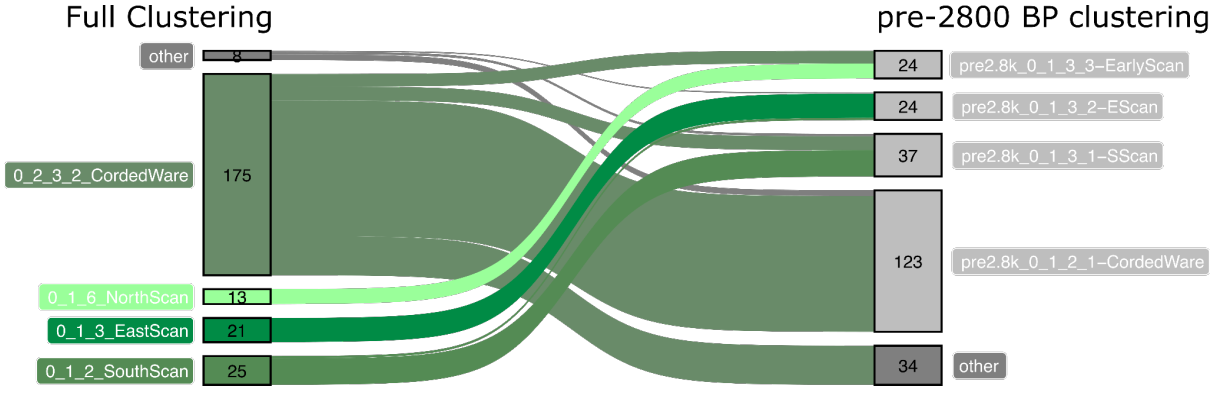


Figure S6.2.2.1. (A) Norwegian Subset of Mixture Modelling Set showing the link through time. Row 1 includes early Bronze Age Norwegian sources, revealing later admixed Bronze Age (50% W.Scan, 50% E.Scan) and Iron Age and Viking individuals (5% W.Scan, 95% E.Scan). Row 2 includes the admixed Bronze Age individuals being a good source for the ~50% of the

Iron Age and Viking Individuals ancestry, including those that were modelled with no early Bronze Age ancestry. 'S' indicates a source individual.
(B)

PDF UPLOADED SEPARATELY

Figure S6.2.2.2. Full plot for Figure S6.2.2.1A, showing all source and target individuals

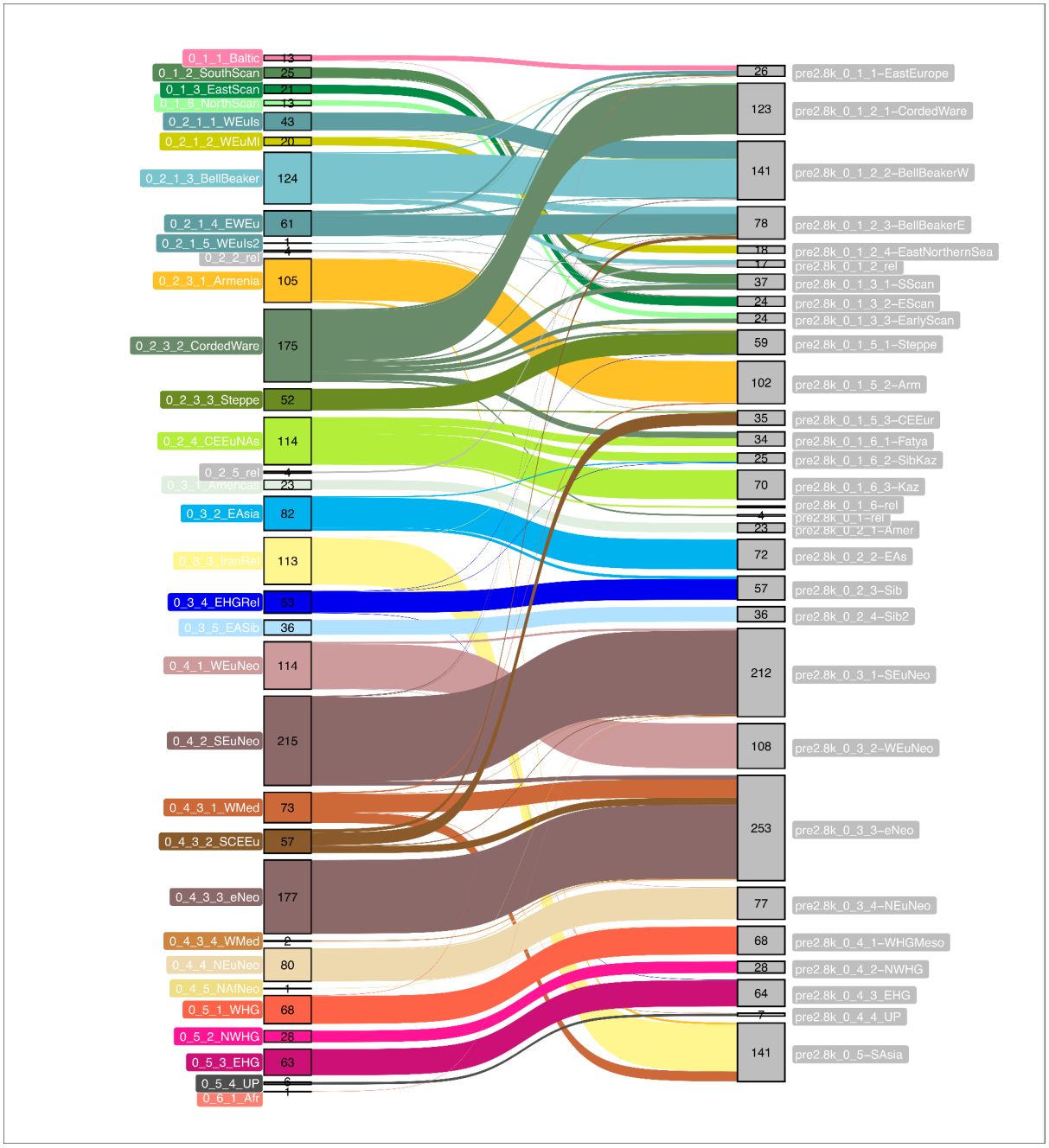
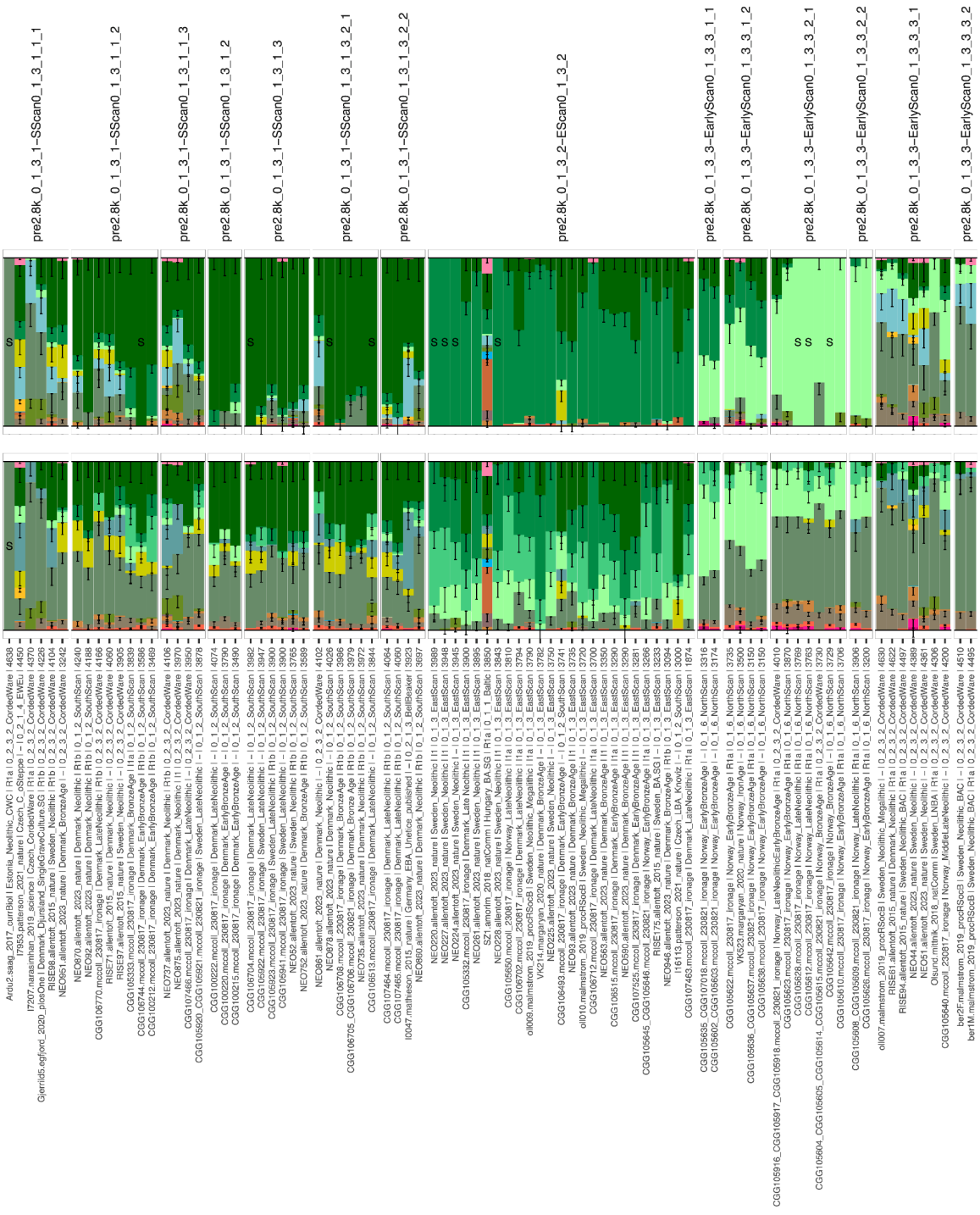


Figure S6.2.2.3. Sankey diagram showing relationship between full clustering (left) and pre2.8k clustering (right).



732

733

734 *Figure S6.2.2.4. Mixture Modelling Results for Set 6 (upper panel, includes Bronze Age*
735 *Scandinavian sources) and Set (lower panel, contains Iron Age Scandinavian Sources).*
736 *Samples are faceted on pre-2800 BP clusters, and the original cluster information is provided*
737 *in the sample label. Samples that were originally in the Corded Ware cluster are modelled with*
738 *high proportions of Corded Ware, Bell Beaker and ENS Ancestry, and only small proportions*
739 *of the local (Western Scandinavian, Southern Scandinavian) ancestries.*

740

S6.3. IBD Mixture Modelling.

S6.3.1. Main Mixture Modelling Sets

For IBD Mixture Modelling we calculated total shared IBD lengths as in S6.2, but with a minimal cut off of 1 cM per segment, and no lower limit on the total shared length. We undertook mixture modelling (Hofmanová *et al.*, 2016) on the total shared IBD lengths as described in (Allentoft, Sikora, Refoyo-Martínez, *et al.*, 2024), starting with the well-established diverse ancestries across Eurasia and adding more proximal sources (Table S4). These results were plotted as 'ADMIXTURE' style bar charts faceted by regions and country (Figure S6.3.1.1), by IBD Cluster (Figure S6.3.1.2) and on the Western Eurasian PCA (Figure S6.3.1.3).

The following 8 sets were used. For each set, increasingly proximal sources are added. The full list of source individuals can be found in Table S4 and Figure S6.3.1.5.

Set 1_base_ehg-swap

The initial set included representatives of the four major ancestries present in Bronze Age Europe: Early Anatolian Farmers, Caucusus Hunter Gatherers, Western Hunter Gatherers from Southern Europe and Ukrainian Eastern Hunter Gatherers. In addition, we included more distant source clusters, who are known to impact Western Eurasia from the Iron Age onwards: Russian Eastern Hunter Gatherers, East Asian, Iran Neolithic and South Africans.

Set 2_base_ehg-swap_w_LHG_wNHG_rmBBFarm

In addition to the clusters from Set 1, the second set included additional Mesolithic hunter-gatherers from along the Western Hunter Gatherer - Eastern Hunter Gatherer cline, to detect variation in hunter-gatherer ancestries in Bronze Age population. These included one cluster from Latvia and Lithuania, and a second from Norway).

Set 3_base_wYam_ehg-swap_w_LHG_wNHG_rmBBFarm

Set 3 is identical to Set 2, but includes Yamanya as a source, who were previously modelled as CHG and EHG.

Set 4_base_wYam_ehg-swap_w_LHG_wNHG_rmBBFarm_wGACSthF_wCWC

Set 4 is identical to Set 3, but includes three clusters with farming related ancestry, representing known admixture events in europe. The first cluster contains early European Farmers, here modelled as a mixture of early Anatolian Farmers and WHG ancestry. The second cluster contains Globular Amphora Culture (GAC) individuals, who carry some EHG ancestry similar to that of Yamnaya. A third cluster contains Bronze Age Anatolians, distinct from the early Anatolians in their additional CHG ancestry. Almost all Bronze Age and later Europeans carry GAC ancestry, whereas European Farmer ancestry is widespread in South and Western Europe, but not in the North.

Set 5_base_wYam_ehg-swap_w_LHG_wNHG_rmBBFarm_wBBCWGACSthF_wCWC

Set 5 is identical to set 4, but contains two additional clusters representing the major two archaeological cultures of the Bronze Age - Corded Ware ancestry and BellBeaker ancestry.

Set 6_base_wYam_ehg-swap_Glb3_rmBBFarm

Set 6 is similar to the set 5, except the hunter-gatherer sources from set 2 have been removed. In addition, clusters related to various West and Nothern European Bronze Age clusters have been added, a Bronze Age Southern Scandinavian cluster (0_1_2_SouthScan), and Bronze Age Eastern Scandinavian Cluster (0_1_3_EastScan), a Bronze Age Western Scandinavia cluster (0_1_6_NorthScan), and Bronze Age Baltic Cluster (0_1_1_Baltic), and a Bronze Age East North Sea cluster (0_2_1_2_WEuM).

Set 7_base_wYam_ehg-swap_ScIA_wSth

Set 7 is similar to Set 6, but now contains Iron Age Southern Scandinavians from Western Denmark (0_1_2_3_2800-), Iron Age Eastern Scandinavians from East Denmark (0_1_3_1_2_2_2800-), Iron Age Eastern Scandinavians from Southern Sweden (0_1_3_2_2800-), Iron Age Western Scandinavians from Norway (0_1_6_2_2800-) and Iron Age East North Sea individuals from the Netherlands (0_2_1_2_2_2800-)

Set 8_base_wYam_ehg-swap_ScIA_wSth_wGmN

Set 8 is similar to set 7, but includes a second more southern Iron Age Southern Scandinavian cluster from Northern Germany.

PDF UPLOADED SEPARATELY

Figure S6.3.1.1.Mixture Modelling Results for all ancient samples, ordered by country

PDF UPLOADED SEPARATELY

Figure S6.3.1.2.Mixture Modelling Results for all ancient samples, ordered by cluster

PDF UPLOADED SEPARATELY

Figure S6.3.1.3. Mixture Modelling Results displayed on the Western Eurasian PCA.

PDF UPLOADED SEPARATELY

Figure S6.3.1.4. Subset from Figure S6.3.1.2, plotting only samples older than 2800 BP

Figure S6.3.1.5. Mixture Modelling Results subset showing just Source individuals used for Sets 1-9

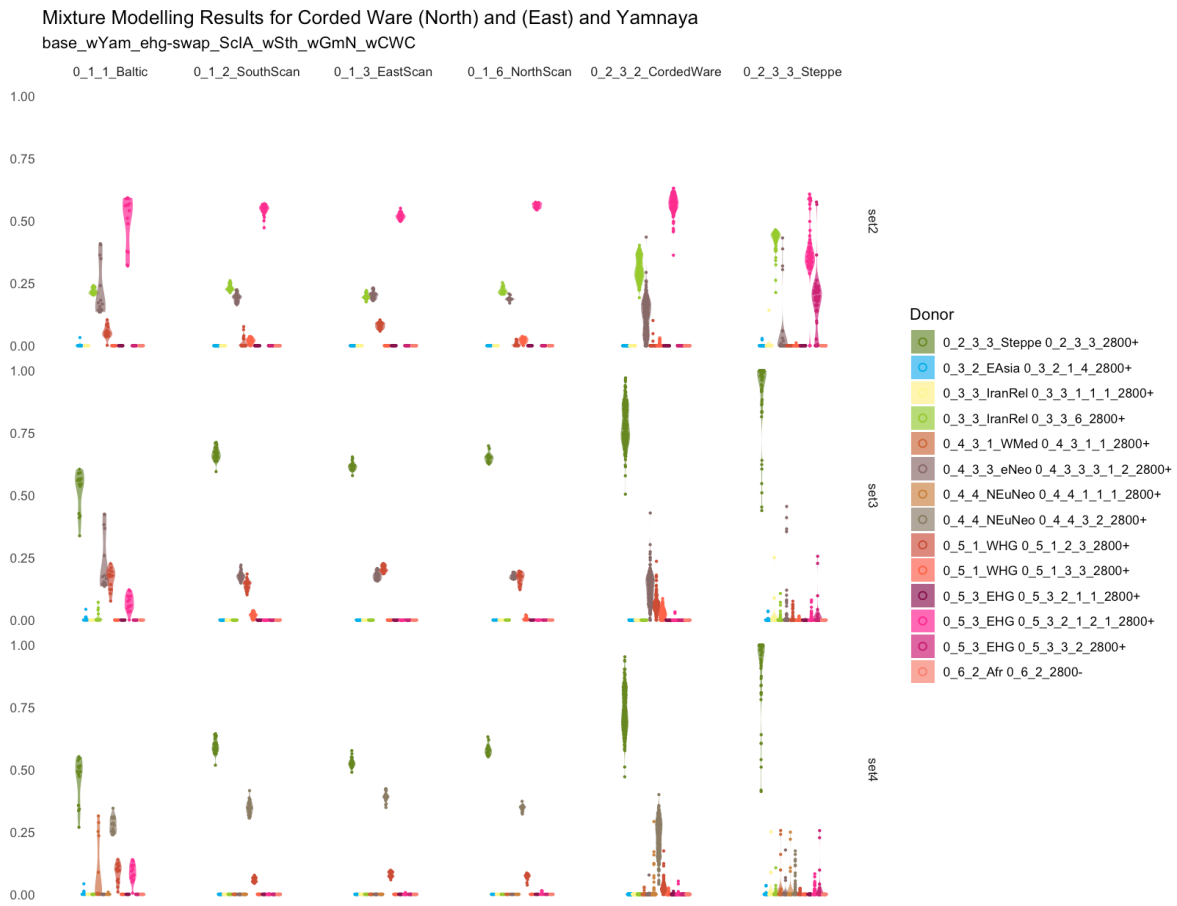


Figure S6.3.1.6. Violin Plots showing subsets of Figure S6.3.1.2 relevant for Corded Ware (East) and (North) clusters, for samples over 2800 BP. Sets 2 and 3 show the variation in ancestry of Western Hunter Gatherer (Southern Europe - 0_5_1_2_3_2800+) compared to Northern Western Hunter Gathers (Latvia, Lithuania - 0_5_1_3_3_2800+) for Southern, East and Western Scandinavians. Set 3 and 4 show the high Eastern Hunter Gatherer from Ukraine present in the Baltic Cluster. Set 3 shows the Yamnaya - Farmer ancestry cline present in the Corded Ware (East) cluster (0_2_3_3). Set 4 shows this ancestry modelled as GAC (0_4_4_3_2_2800+)

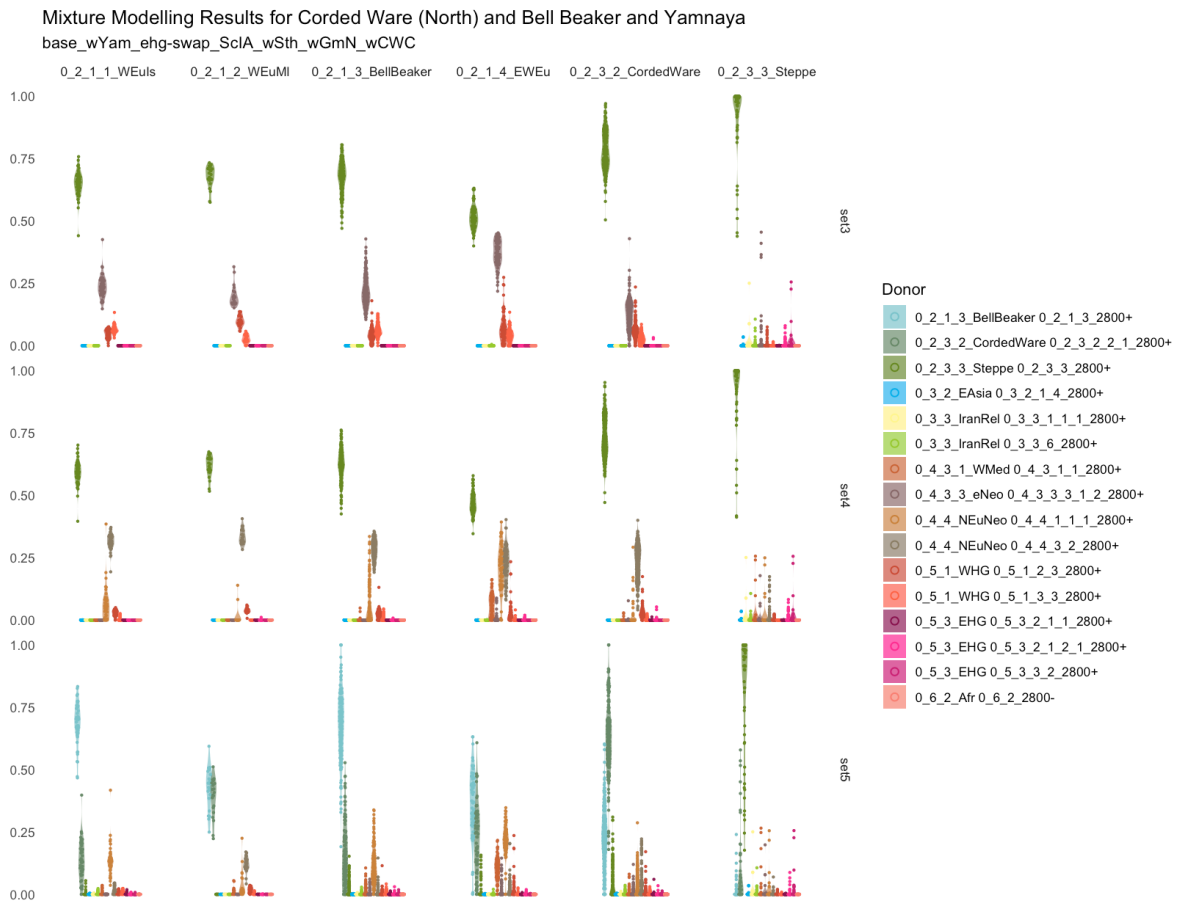


Figure S6.3.1.7. Violin Plots showing subsets of Figure S6.3.1.2 relevant for Corded Ware (East) and Bell Beaker clusters, for samples over 2800 BP. Set 4 shows the European Farmer (0_4_4_1_1_1_2800+) present in most Bell Beakers, in addition to the GAC (0_4_4_3_2_2800+) ancestry found in the Corded Ware (East) individuals. Set 5 shows the high proportion of Bell Beaker ancestry relative to Corded Ware ancestry in most individuals. The East North Sea (0_1_2_1_WEuMl) cluster is unique, in the high proportion of Northern Western Hunter Gatherer (Latvia, Lithuania - 0_5_1_3_3_2800+) ancestry, low European Farmer (0_4_4_1_1_1_2800+) ancestry and equal proportion of Bell Beaker and Corded Ware ancestry, reflecting both its geographical position between the two cultures and position in the Western Eurasian PCA space.

PDF UPLOADED SEPARATELY

Figure S6.3.1.8. Mixture Modelling Results overlaid on a map of Northern Europe for Set 7, including one Southern Scandinavian IA sources.

PDF UPLOADED SEPARATELY

Figure S6.3.1.9. Mixture Modelling Results overlaid on a map of Northern Europe for Set 8, including two Southern Scandinavian IA sources.

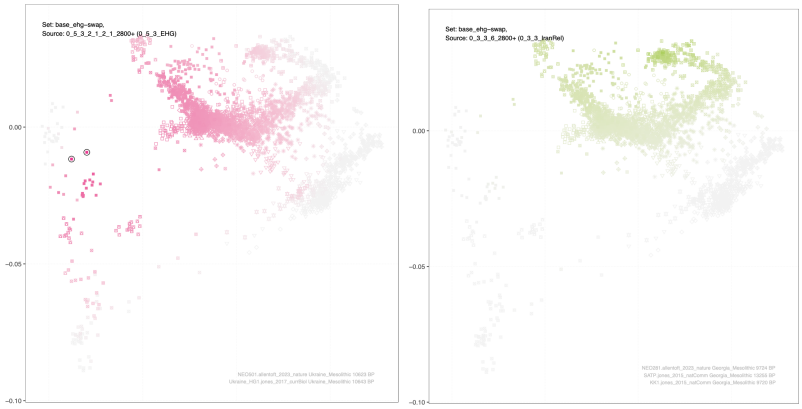


Figure S6.3.1.10. *PCA results coloured by Mixture modelling proportions when including more proximal European Steppe-related sources: Eastern Hunter Gatherer (dark green), Caucasus Hunter Gatherer (light green) source clusters (Set 1). Source individuals are circled. Subset of Figure S6.3.1.3*

S6.3.2. Auxiliary Mixture Modelling Sets

The 'main' mixture modelling sets were generated with more proximate source clusters added in each set. We also generated a set of auxiliary sets.

Set A1. Similar to main Set 3, except only a single EHG and WHG source cluster are included to demonstrate the WHG - WHG cline (Extended Data Figure 1) . All main sets have additional sources from the EHG - WHG cline. Results in Figure S6.2.2.1.

Set A2. Similar to main Set 6, but with Bronze Age early admixed Eastern and local (Western and Southern) Scandinanavians as sources for later Bronze Age individuals. Results in Figure S6.2.2.2.

PDF UPLOADED SEPARATELY

Figure S6.3.2.1. Mixture Modelling Results for Set A1.

PDF UPLOADED SEPARATELY

Figure S6.3.2.2. Mixture Modelling Results for Set A2.

S6.4. Analysis of uniparental markers

S6.4.1. Mitochondrial DNA analysis

Mitogenomes of the newly sequenced individuals were reconstructed using bcftools (Li, 2011) call and mpileup from reads mapped to rCRS from which we subsequently classified the haplogroups using haplogrep (Weissensteiner *et al.*, 2016). We then aligned the sequences with mafft (Katoh and Standley, 2013) and generated a phylogenetic tree with the maximum likelihood (ML) based tree inference tool raxML (Kozlov *et al.*, 2019) using under the GTR+I+G4 model with the options [--all –bs-trees 100]. The analysis was restricted to only the coding region ranging from 577-16,023 base pairs (bp) (rCRS coordinates).

The phylogenetic analysis and classification of haplogroups suggests that the introduction of eg. farmer related haplogroups such as H, V, T, and J were already widespread across Europe during the Iron Age. We, furthermore, identify U haplogroups, which have formerly been

commonly observed in Mesolithic Hunter-Gatherers, to be frequently observed among the sampled individuals, suggesting either continuation or a reintroduction of Hunter-Gatherer related haplogroups into Scandinavia. We see evidence of a broad level of continuity between the global IBD clusters for the European individuals from the period spanning from Bronze Age to Iron Age, and find farmer associated haplogroups such as haplogroup H to be predominantly observed among Western and Northern European populations suggesting that the migrations occurring in this period might not have caused complete replacement of mitochondrial lineages.

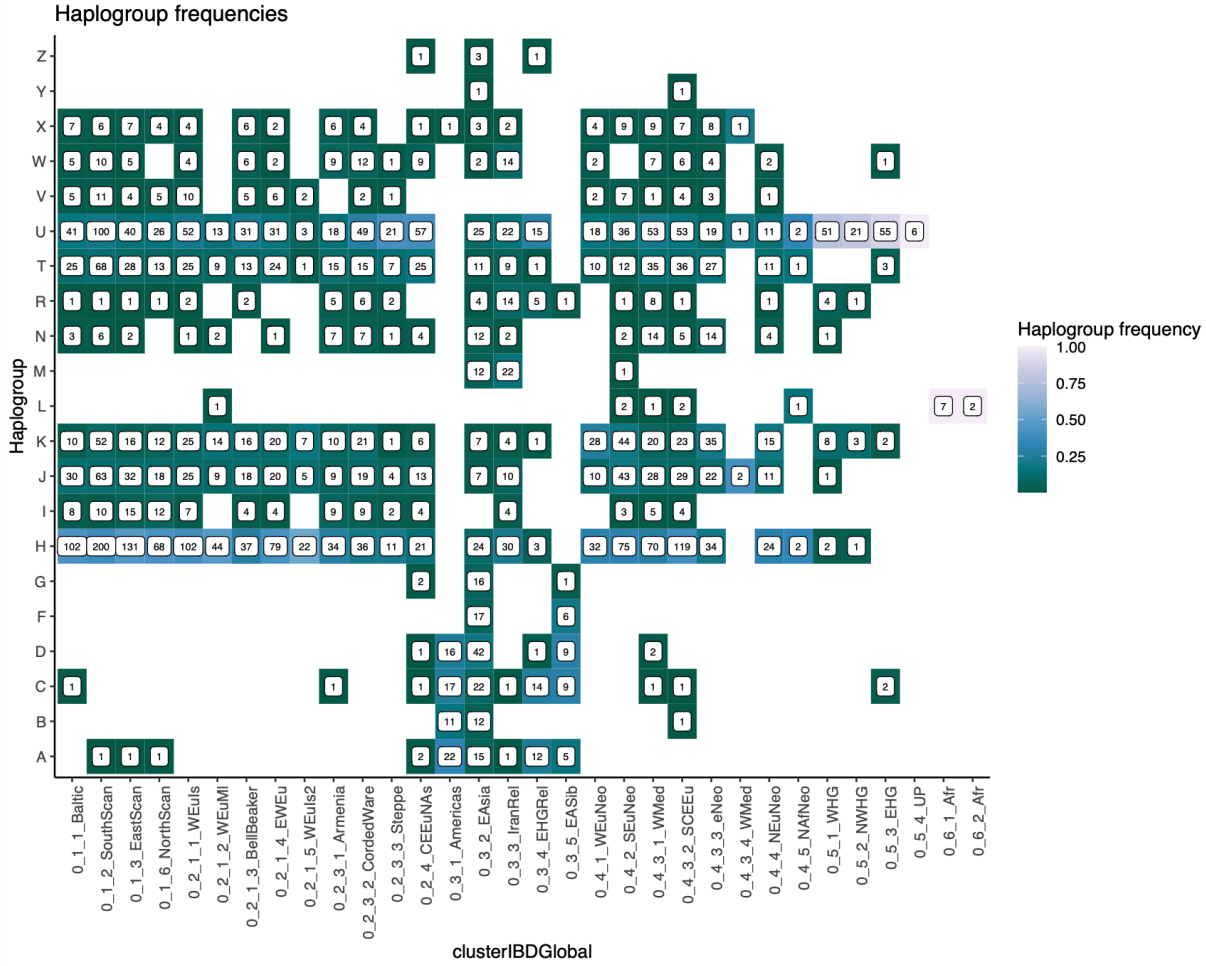
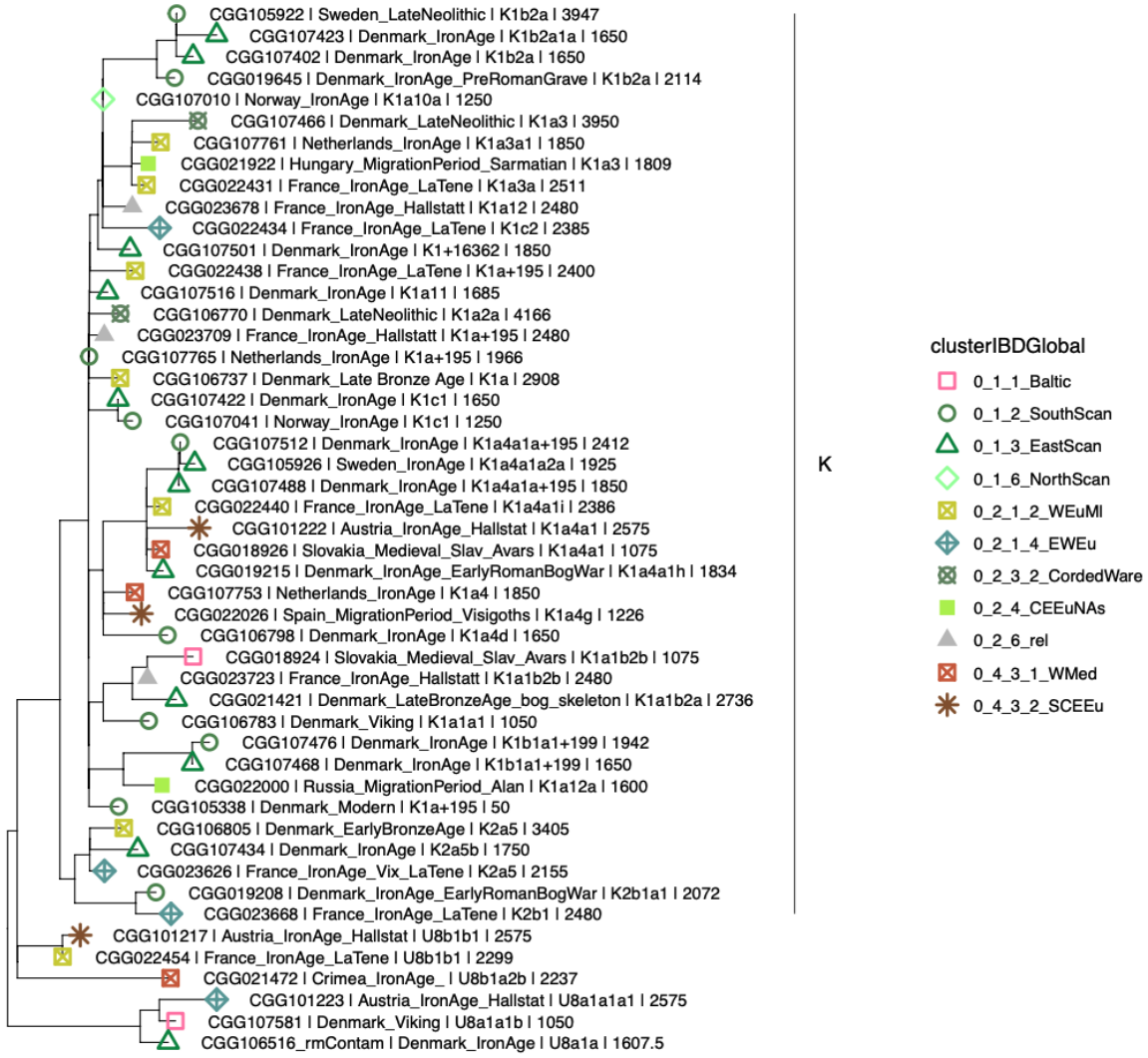


Fig S6.4.1.1 Haplogroup frequencies across Global IBD clusters

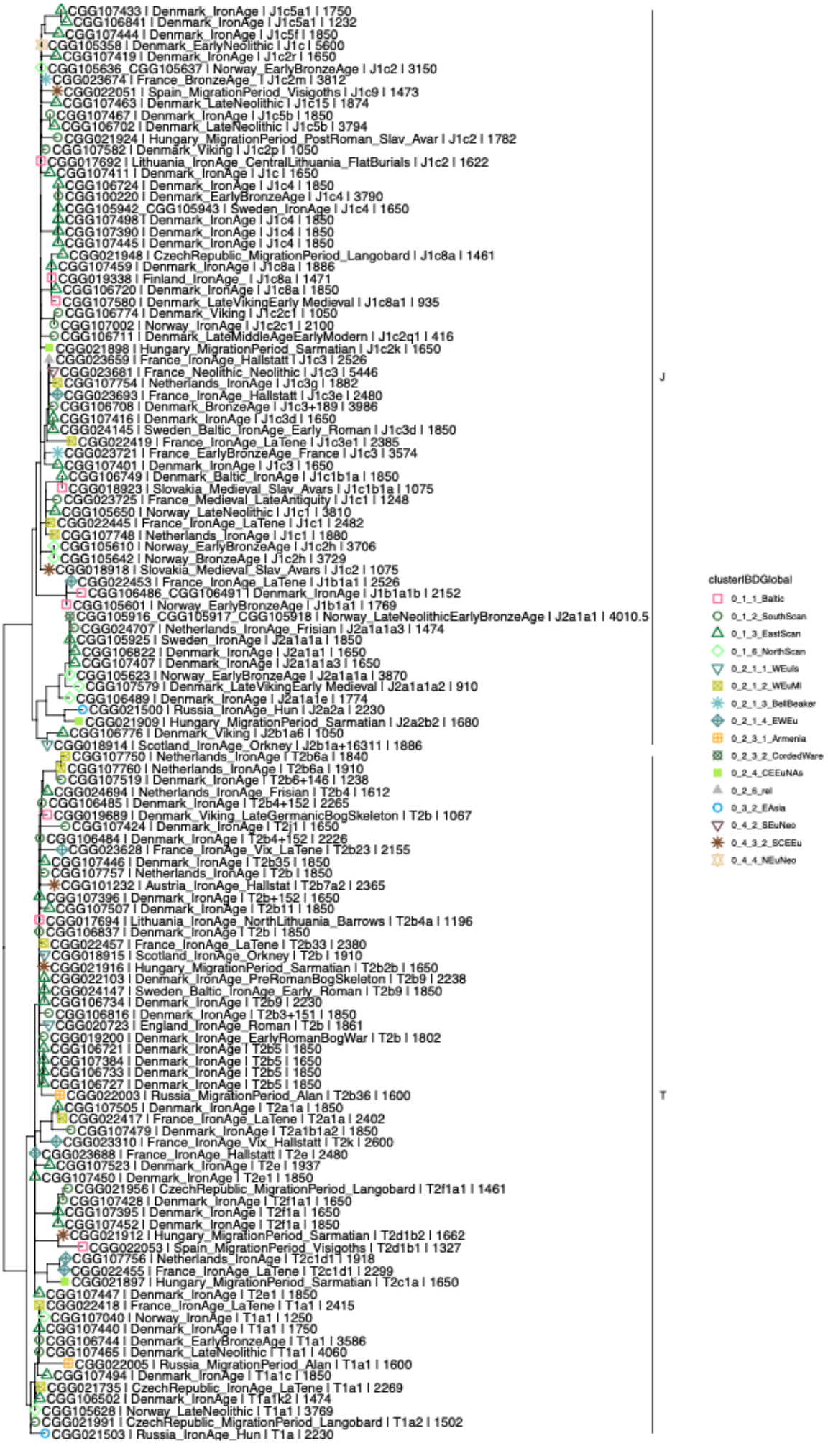
941

942 S6.4.1.2. Maximum likelihood of U haplogroup mitogenomes

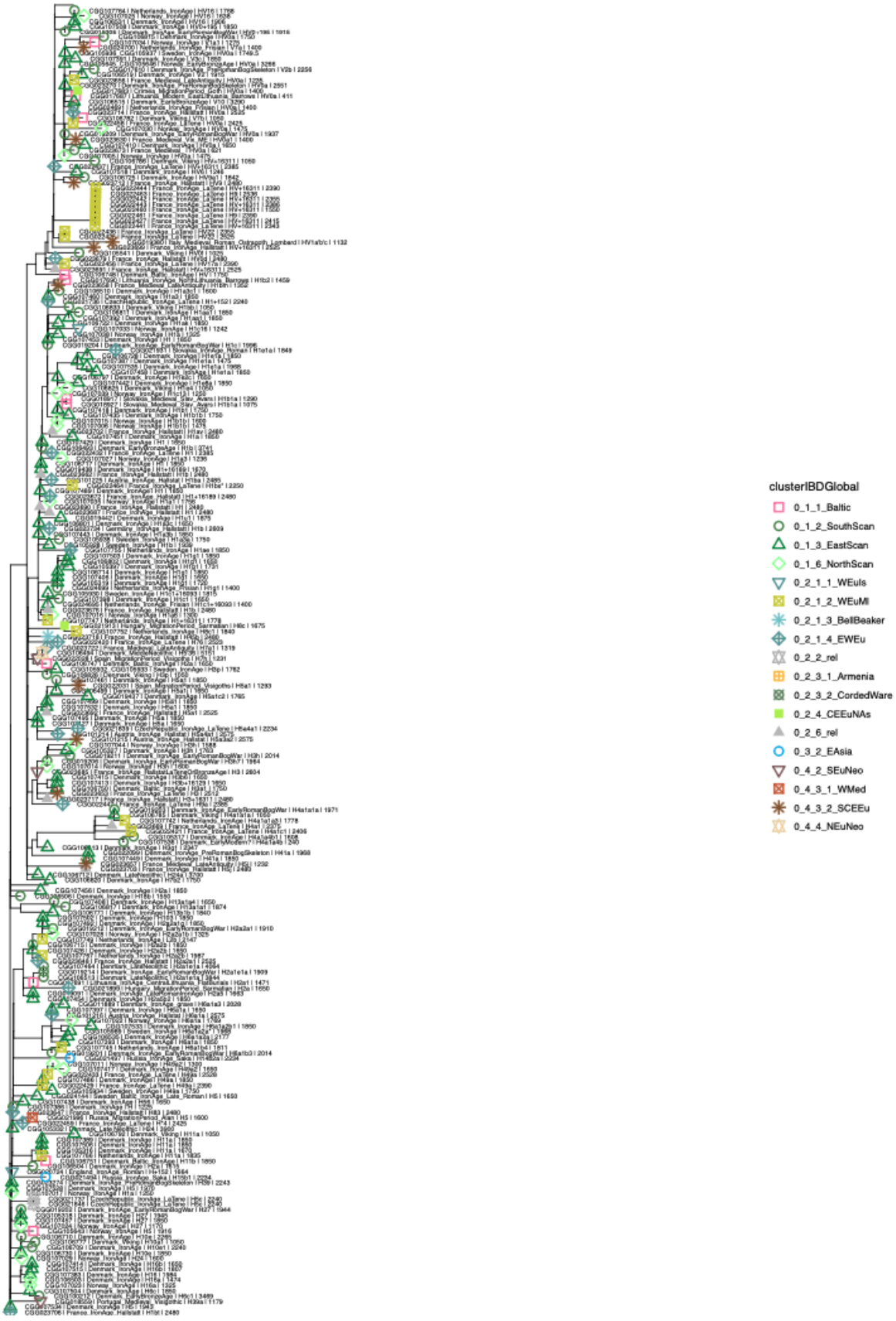
943



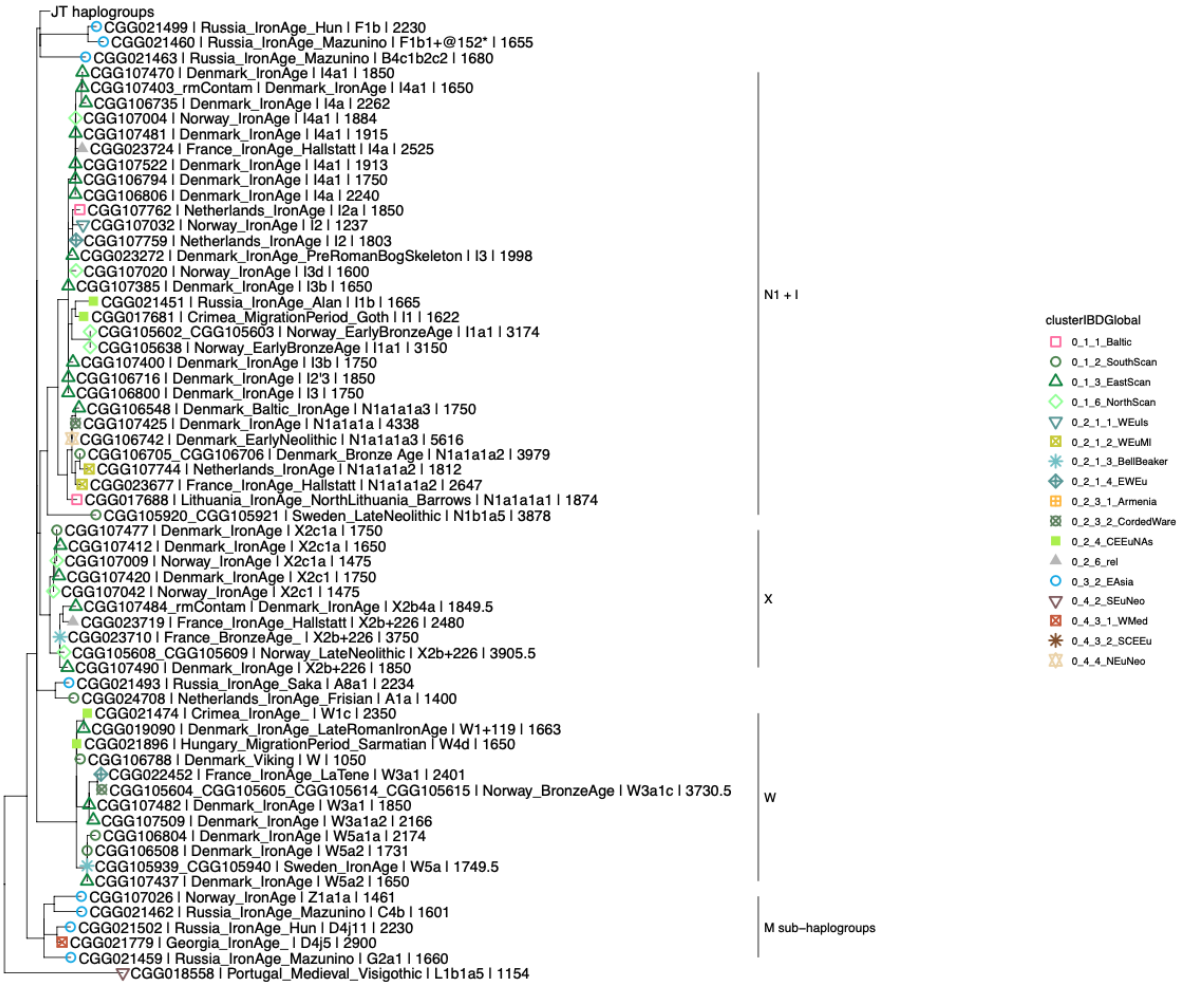
S6.4.1.3. Maximum likelihood of K haplogroup mitogenomes


CGG107433 | Denmark_IronAge | J1c5a1 | 1750
CGG106841 | Denmark_IronAge | J1c5a1 | 1232
CGG107444 | Denmark_IronAge | J1c5f | 1850
CGG105358 | Denmark_EarlyNeolithic | J1c | 5600
CGG107419 | Denmark_IronAge | J1c2r | 1650
CGG105636_CGG105637 | Norway_EarlyBronzeAge | J1c2 | 3150
CGG023674 | France_BronzeAge_ | J1c2m | 3812
CGG022051 | Spain_MigrationPeriod_Visigoths | J1c9 | 1473
CGG107463 | Denmark_LateNeolithic | J1c15 | 1874
CGG107467 | Denmark_IronAge | J1c5b | 1850
CGG106702 | Denmark_LateNeolithic | J1c5b | 3794
CGG021924 | Hungary_MigrationPeriod_PostRoman_Slav_Avar | J1c2 | 1782
CGG107582 | Denmark_Viking | J1c2p | 1050
CGG017692 | Lithuania_IronAge_CentralLithuania_FlatBurials | J1c2 | 1622
CGG107411 | Denmark_IronAge | J1c | 1650
CGG106724 | Denmark_IronAge | J1c4 | 1850
CGG100220 | Denmark_EarlyBronzeAge | J1c4 | 3790
CGG105942_CGG105943 | Sweden_IronAge | J1c4 | 1650
CGG107498 | Denmark_IronAge | J1c4 | 1850
CGG107390 | Denmark_IronAge | J1c4 | 1850
CGG107445 | Denmark_IronAge | J1c4 | 1850
CGG021948 | CzechRepublic_MigrationPeriod_Langobard | J1c8a | 1461
CGG107459 | Denmark_IronAge | J1c8a | 1886
CGG019338 | Finland_IronAge_ | J1c8a | 1471
CGG106720 | Denmark_IronAge | J1c8a | 1850
CGG107580 | Denmark_LateVikingEarly Medieval | J1c8a1 | 935
CGG106774 | Denmark_Viking | J1c2c1 | 1050
CGG107002 | Norway_IronAge | J1c2c1 | 2100
CGG106711 | Denmark_LateMiddleAgeEarlyModern | J1c2q1 | 416
CGG021898 | Hungary_MigrationPeriod_Sarmatian | J1c2k | 1650
CGG023659 | France_IronAge_Hallstatt | J1c3 | 2526
CGG023681 | France_Neolithic_Neolithic | J1c3 | 5446
CGG107754 | Netherlands_IronAge | J1c3g | 1882
CGG023693 | France_IronAge_Hallstatt | J1c3e | 2480
CGG106708 | Denmark_BronzeAge | J1c3+189 | 3986
CGG107416 | Denmark_IronAge | J1c3d | 1650
CGG024145 | Sweden_Baltic_IronAge_Early_Roman | J1c3d | 1850
CGG022419 | France_IronAge_LaTene | J1c3e1 | 2385
CGG023721 | France_EarlyBronzeAge_France | J1c3 | 3574
CGG107401 | Denmark_IronAge | J1c3 | 1650
CGG106749 | Denmark_Baltic_IronAge | J1c1b1a | 1850
CGG018923 | Slovakia_Medieval_Slav_Avars | J1c1b1a | 1075
CGG023725 | France_Medieval_LateAntiquity | J1c1 | 1248
CGG105650 | Norway_LateNeolithic | J1c1 | 3810
CGG022445 | France_IronAge_LaTene | J1c1 | 2482
CGG107748 | Netherlands_IronAge | J1c1 | 1880
CGG105610 | Norway_EarlyBronzeAge | J1c2h | 3706
CGG105642 | Norway_BronzeAge | J1c2h | 3729
CGG018918 | Slovakia_Medieval_Slav_Avars | J1c2 | 1075
CGG022453 | France_IronAge_LaTene | J1b1a1 | 2526
CGG106486_CGG106491 | Denmark_IronAge | J1b1a1b | 2152
CGG105601 | Norway_EarlyBronzeAge | J1b1a1 | 1769
CGG105916_CGG105917_CGG105918 | Norway_LateNeolithicEarlyBronzeAge | J2a1a1 | 4010.5
CGG024707 | Netherlands_IronAge_Frisian | J2a1a1a3 | 1474
CGG105925 | Sweden_IronAge | J2a1a1a | 1850
CGG106822 | Denmark_IronAge | J2a1a1 | 1650
CGG107407 | Denmark_IronAge | J2a1a1a3 | 1650
CGG105623 | Norway_EarlyBronzeAge | J2a1a1a | 3870
CGG107579 | Denmark_LateVikingEarly Medieval | J2a1a1a2 | 910
CGG106489 | Denmark_IronAge | J2a1a1e | 1774
CGG021500 | Russia_IronAge_Hun | J2a2a | 2230
CGG021909 | Hungary_MigrationPeriod_Sarmatian | J2a2b2 | 1680
CGG106776 | Denmark_Viking | J2b1a6 | 1050
CGG018914 | Scotland_IronAge_Orkney | J2b1a+16311 | 1886
J
CGG107750 | Netherlands_IronAge | T2b6a | 1840
CGG107760 | Netherlands_IronAge | T2b6a | 1910
CGG107519 | Denmark_IronAge | T2b6+146 | 1238
CGG024694 | Netherlands_IronAge_Frisian | T2b4 | 1612
CGG106485 | Denmark_IronAge | T2b4+152 | 2265
CGG019689 | Denmark_Viking_LateGermanicBogSkeleton | T2b | 1067
CGG107424 | Denmark_IronAge | T2j1 | 1650
CGG106484 | Denmark_IronAge | T2b4+152 | 2226
CGG023628 | France_IronAge_Vix_LaTene | T2b23 | 2155
CGG107446 | Denmark_IronAge | T2b35 | 1850
CGG107757 | Netherlands_IronAge | T2b | 1850
CGG101232 | Austria_IronAge_Hallstat | T2b7a2 | 2365
CGG107396 | Denmark_IronAge | T2b+152 | 1650
CGG107507 | Denmark_IronAge | T2b11 | 1850
CGG017694 | Lithuania_IronAge_NorthLithuania_Barrows | T2b4a | 1196
CGG106837 | Denmark_IronAge | T2b | 1850
CGG022457 | France_IronAge_LaTene | T2b33 | 2380
CGG018915 | Scotland_IronAge_Orkney | T2b | 1910
CGG021916 | Hungary_MigrationPeriod_Sarmatian | T2b2b | 1650
CGG022103 | Denmark_IronAge_PreRomanBogSkeleton | T2b9 | 2238
CGG024147 | Sweden_Baltic_IronAge_Early_Roman | T2b9 | 1850
CGG106734 | Denmark_IronAge | T2b9 | 2230
CGG106816 | Denmark_IronAge | T2b3+151 | 1850
CGG020723 | England_IronAge_Roman | T2b | 1861
CGG019200 | Denmark_IronAge_EarlyRomanBogWar | T2b | 1802
CGG106721 | Denmark_IronAge | T2b5 | 1850
CGG107384 | Denmark_IronAge | T2b5 | 1650
CGG106733 | Denmark_IronAge | T2b5 | 1850
CGG106727 | Denmark_IronAge | T2b5 | 1850
CGG022003 | Russia_MigrationPeriod_Alan | T2b36 | 1600
CGG107505 | Denmark_IronAge | T2a1a | 1850
CGG022417 | France_IronAge_LaTene | T2a1a | 2402
CGG107479 | Denmark_IronAge | T2a1b1a2 | 1850
CGG023310 | France_IronAge_Vix_Hallstatt | T2k | 2600
CGG023688 | France_IronAge_Hallstatt | T2e | 2480
CGG107523 | Denmark_IronAge | T2e | 1937
CGG107450 | Denmark_IronAge | T2e1 | 1850
CGG021956 | CzechRepublic_MigrationPeriod_Langobard | T2f1a1 | 1461
CGG107428 | Denmark_IronAge | T2f1a1 | 1650
CGG107395 | Denmark_IronAge | T2f1a | 1650
CGG107452 | Denmark_IronAge | T2f1a | 1850
CGG021912 | Hungary_MigrationPeriod_Sarmatian | T2d1b2 | 1662
CGG022053 | Spain_MigrationPeriod_Visigoths | T2d1b1 | 1327
CGG107756 | Netherlands_IronAge | T2c1d1 | 1918
CGG022455 | France_IronAge_LaTene | T2c1d1 | 2299
CGG021897 | Hungary_MigrationPeriod_Sarmatian | T2c1a | 1650
CGG107447 | Denmark_IronAge | T2e1 | 1850
CGG022418 | France_IronAge_LaTene | T1a1 | 2415
CGG107040 | Norway_IronAge | T1a1 | 1250
CGG107440 | Denmark_IronAge | T1a1 | 1750
CGG106744 | Denmark_EarlyBronzeAge | T1a1 | 3586
CGG107465 | Denmark_LateNeolithic | T1a1 | 4060
CGG022005 | Russia_MigrationPeriod_Alan | T1a1 | 1600
CGG107494 | Denmark_IronAge | T1a1c | 1850
CGG021735 | CzechRepublic_IronAge_LaTene | T1a1 | 2269
CGG106502 | Denmark_IronAge | T1a1k2 | 1474
CGG105628 | Norway_LateNeolithic | T1a1 | 3769
CGG021991 | CzechRepublic_MigrationPeriod_Langobard | T1a2 | 1502
CGG021503 | Russia_IronAge_Hun | T1a | 2230
T
clusterIBDGlobal
0_1_1_Baltic
0_1_2_SouthScan
0_1_3_EastScan
0_1_6_NorthScan
0_2_1_1_WEuIs
0_2_1_2_WEuMI
0_2_1_3_BellBeaker
0_2_1_4_EWEu
0_2_3_1_Armenia
0_2_3_2_CordedWare
0_2_4_CEEuNAs
0_2_6_rel
0_3_2_EAsia
0_4_2_SEuNeo
0_4_3_2_SCEEu
0_4_4_NEuNeo

S6.4.1.4. Maximum likelihood of J and T haplogroup mitogenomes



S6.4.1.5. Maximum likelihood of H and V haplogroup mitogenomes



S6.4.1.6. Maximum likelihood of N, X, W, and M haplogroup mitogenomes

S6.4.2. Y-chromosome analysis

The Y-chromosome is the largest uniparentally inherited haplotypic block of the human genome, and it has been correlated with many known events in human history (3,10,11). Due to its larger size, lower effective population size and because it traces primarily the movements of males, it has been shown to be more sensitive to diagnose past demographic fluctuations, such as migration and depopulation, than mitochondrial DNA (12). Initially, most of the developments in the Y-chromosome phylogeny were done by large public academic projects of worldwide diversity (11,13–16), but today most of the recent advances were largely first described in forensic or community-based citizen-scientist efforts using private genetic data (3), and often privately owned databases are far more complete than academic ones. While those are immensely biased towards Western Eurasian ancestry (particularly Northwest

Europe), this is also true of most genetic data overall (17,18). Ancient DNA poses particular problems to the study of Y-chromosome, such as pervasive missing data, relatively lower coverage, and confounding factors, as miscoding lesions ("damage") and potential contamination (19,20). Many approaches have been used to address those issues, such as giving different weights to transition and transversion variants (21), but in cases where dense sampling in the haplogroup exists and depth of coverage is not limiting a factor for ancient data, phylogenetic placement algorithms have proven to be the most effective (22–24).

Materials & Methods

We used bcftools (25) mpileup and call function to call genotypes specifying chromosome ploidy, in the 10 Mb accessible region of the Y (2) and excluded indels, triallelic positions and genotypes not called in more than 95% of the population of non-clonal reads. Because haplogroup-defining SNPs are defined from the root of the Y-chromosome phylogeny but the Y-chromosome reference is a mosaic of at least two donors of common European haplogroups (R1b and G2a – see also (26)), variants should be considered in terms of ancestral/derived instead of reference/alternate, particularly for individuals belonging to those two haplogroups. We therefore proceeded to call haplogroups by matching ancestral and derived calls to ISOGG 2019-2020 database, using an in-house script which generates haplogroup paths in root-to-tip fashion, sorted by number of supported variants seen, while also distinguishing C>T and G>A variants in the forward and reverse strand, which could arise from damage. To further incorporate results not listed in ISOGG and make use of the wealth of customer-based private datasets, we have also annotated VCFs using bcftools to all biallelic single nucleotide point mutation variants listed in yBrowse (https://ybrowse.org/gbrowse2/gff/) and matched those to the privately-owned Yfull tree (https://github.com/YFullTeam/YTree/blob/master/ytree/tree_10.07.0.zip). To visualise the results and confirm haplogroup placement of low-coverage samples, we used PathPhynder (23), which also takes into consideration the presence of ancestral variants to create a parsimonious phylogenetic path for the samples. Y-chromosome results from the new samples have been tabulated in Table S1. As we intended to match Y-chromosome haplogroups with

the autosome IBD clusters, we repeated the process above with samples from the literature, but restricting the subhaplogroup definition to informative clades discussed in the text.

Discussion

Almost the totality of the Scandinavians older than 4,500 BP belong to haplogroup I2-P215 (52 out of 57). Only starting at 4,500 BP we see the arrival of lineages more common today in the region, such as I1-M253, R1a1a1 (R1a-M417) and R1b1a1b1a (R1b-L11).

Europeans after 5,000 BP were clustered in four distinct groups, with majority of individuals from Scandinavia belonging to the cluster associated with northern Corded Ware ancestry (0_1 CordedWareNorth), but fifteen individuals dated between 4,600 to 3,700 BP instead were placed along other individuals of similar age in a more eastern group (0_2_3_2 CordedWareEast). The haplogroup frequencies in those is similar to the main cluster, with R1a1a1 (R1a-M417) and R1b1a1b1a (R1b-L11) representing together two thirds of the male lineages in both cases (66/99 for 0_2_3_2_CordedWareEast and 10/15 for the Scandinavian individuals in it).

In agreement with the finding that those four main clusters are genetically related, when investigating the Y-chromosome lineages of the older members of those groupings, we find them to be dominated by closely related, but clearly distinct, lineages. This can be seen by comparing individuals most associated with early Bell Beakers (0_2_1_3 BellBeaker) and Yamnaya (0_2_3_3 Steppe) clusters. While both are dominated by R1b1a1b (R1b-M269) lineages, the prevalent lineage among the Yamnaya – R1b1a1b1b (R1b-Z2103), found in 18 out of 29 males – is much more common in the Caucasus and Eastern Europe, and is a sister lineage (therefore not a direct source) of the present-day dominant Western European lineage R1b1a1b1a (R1b-L11) (Figure S1 and S2). Instead, almost the totality of males in the Bell Beaker cluster (70 out of 83) are downstream of this haplogroup, at R1b1a1b1a1a2 (R1b-P312), agreeing of modelling results of Bell Beakers being a better source for modern populations than Yamnaya.

When investigating the four main subgroups in the main cluster containing Scandinavian individuals (0_1 CordedWareNorth), we found no lineage occurring more than 5 times to be

exclusive to neither four, which reiterates the genetic similarity of those clusters. However, some lineages carry clear geographical correlations when using a time transect, as those clusters are shown with autosomes to have been admixed after 2,800 BP.

Haplogroup I1a-DF29 first appears in Scandinavia in the Bronze Age, and its vast number of local descendant lineages seem to indicate an in situ diversification and point of origin (Figure S3). Haplogroup I1a2a (I1a-Z59) is most the common variant of I1a-DF29 among both 0_1_2 SouthScan and 0_1_3 EastScan clusters, but when restricting to individuals older to 2,800 BP, all four are in the Eastern Scandinavian cluster (Figure S4). This likely indicates that its presence in the South Scandinavian cluster to be associated with the later mixture from an Eastern Scandinavian source.

The four more common subhaplogroups of R1a1a1 (R1a-M417) in Northern Europe also show frequency differences among those four clusters. Almost all R1a1a1b1a1 (R1a-M458) and R1a1a1b1a2 (R1a-Z280) individuals belong to the 0_1_1 Baltic cluster, which together with the presence of E1b1b1a1b1 (E1b-L618) individuals, represent a Central and Eastern European affinity of this cluster not seen in the remaining Corded Ware North subgroups.

Conversely, haplogroups R1a1a1a1 (R1a-L664) and R1a1a1b1a3a (R1b-Z284) are more frequent in three mainly Scandinavian clusters, with the majority of R1a1a1b1a3a (R1b-Z284) from the whole dataset deriving from 0_1_6 NorthScan. The oldest evidence of this haplogroup takes place in the Corded Ware East cluster (0_2_3_2 CordedWareEast) (at around 4,500 BP), but it is absent in Scandinavia outside of 0_1_6 NorthScan before 2,800 BP (Figure S5). Therefore, while we can ultimately trace this subhaplogroup to individuals related to eastern Neolithic Corded Ware ancestry, it likely spread later in Northern Europe, through mixture with individuals carrying North Scandinavian associated ancestry.

Downstream of R1b1a1b1a (R1b-L11), haplogroup R1b1a1b1a1a1 (R1b-U106) have been previously argued to be related to the expansion of the Germanic languages, due to its high frequency in places where those languages are spoken today (Figure S6). We found most of the individuals of the dataset positive for R1b-U106 to belong to two different downstream sublineages, which have starkly distinct distributions, particularly in the early Iron Age. R1b1a1b1a1a1c (R1b-Z19) is found almost exclusively in Northern Europe (with the only

exception being a Langobard from Hungary), and likely represents a local variant of R1b-U106 (Figure S7).

Instead, its sister lineage, R1b1a1b1a1a1b (R1b-S263), is absent in Scandinavia before the Iron Age (Figure S8), where it spreads, likely through an Eastern North Sea source, and becomes dominant in South Scandinavia during the Iron Age, before spreading through Northern Europe. This pattern strongly matches the one seen using autosomes, that detect gene flow back into Scandinavia related to the spread of Germanic languages. Another potential signal of this migration is the increase in frequency of R1b-U106 sister lineage, R1b1a1b1a1a2 (R1b-P312), that has a more continental distribution. and is almost absent in Scandinavia before 2,000 BP.

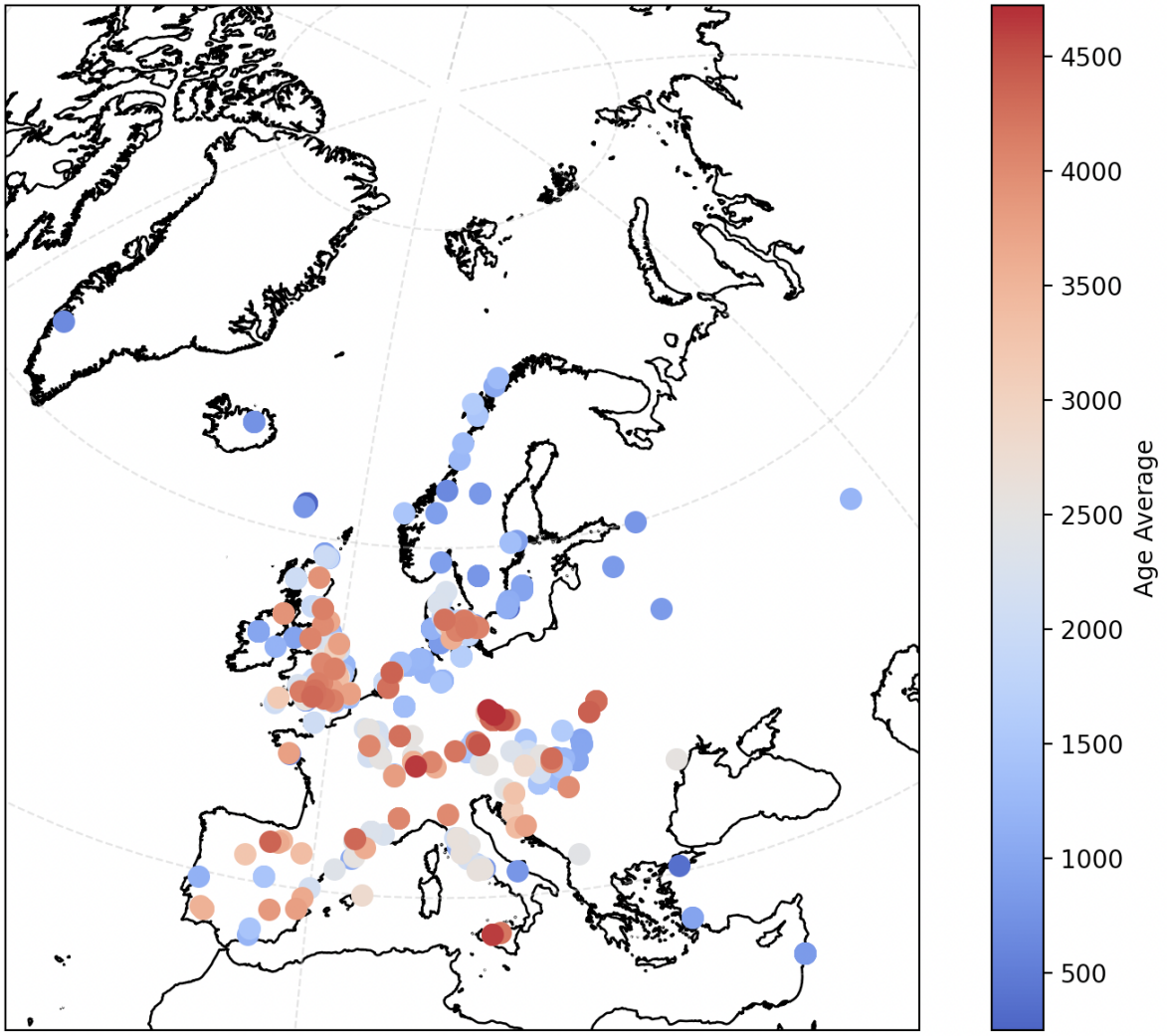


Figure S6.4.2.1: All individuals from the dataset belonging to haplogroup R1b1a1b1a (R1b-L11), coloured by age before the present.

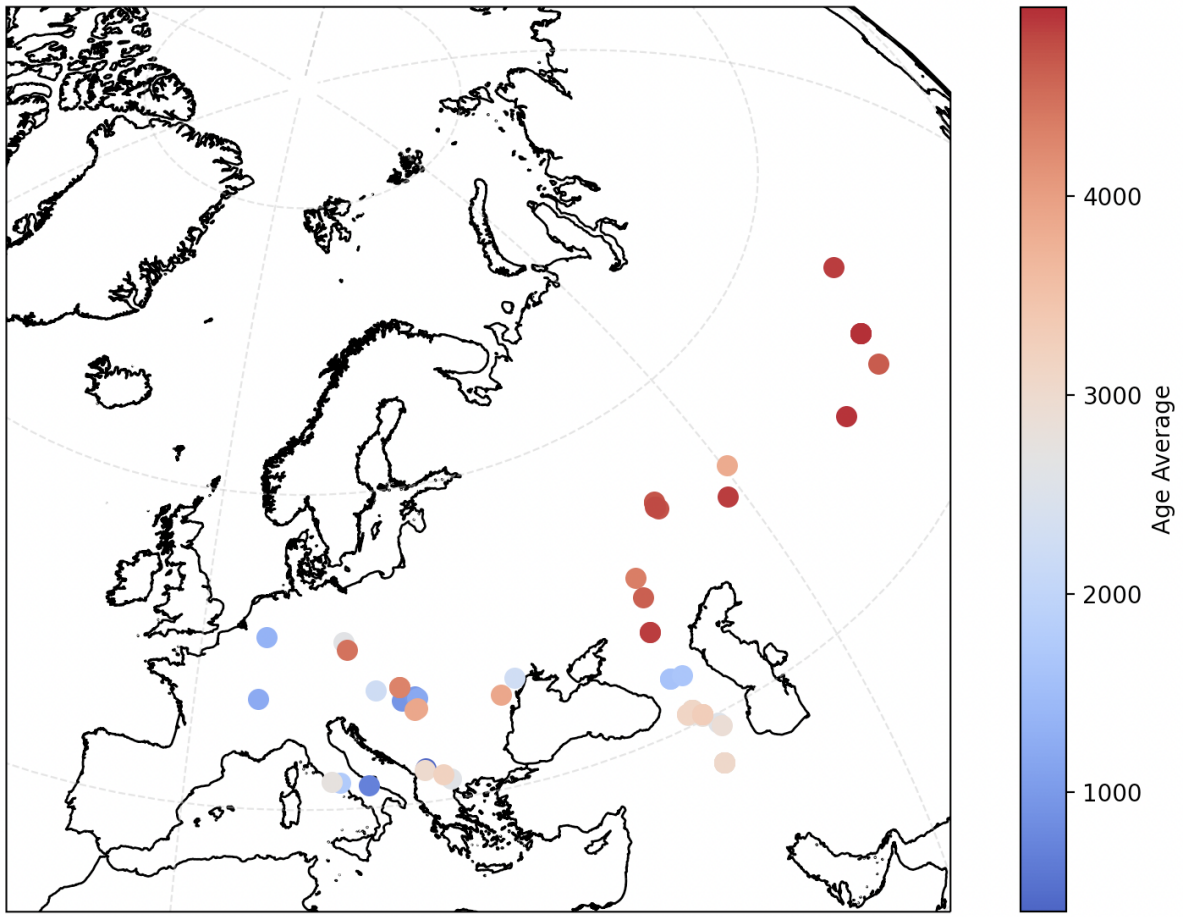


Figure S6.4.2.2: All individuals from the dataset belonging to haplogroup R1b1a1b1b (R1b-LZ2103), coloured by age before the present.

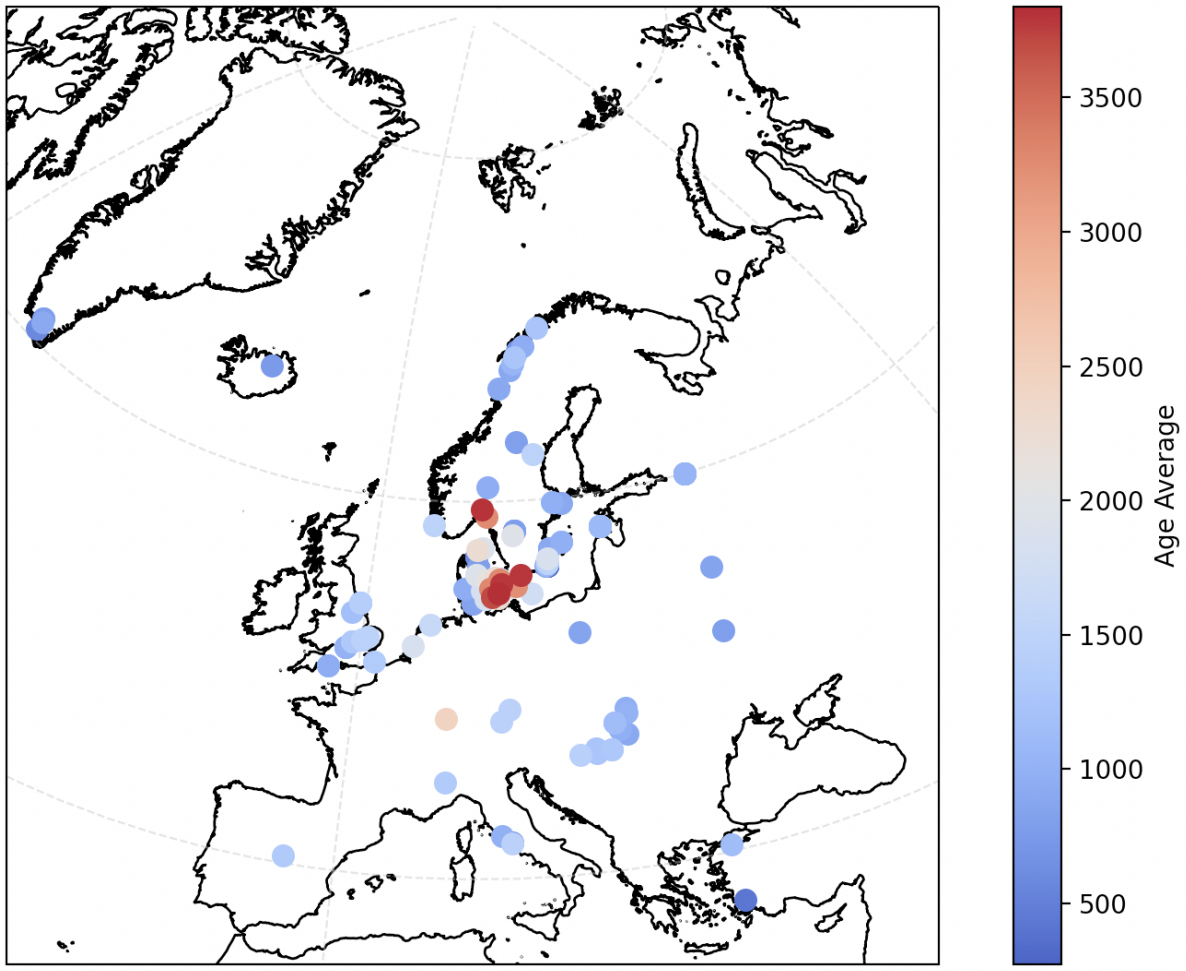


Figure S6.4.2.3: All individuals carrying haplogroup I1a-DF29 from the dataset, coloured by age before the present.

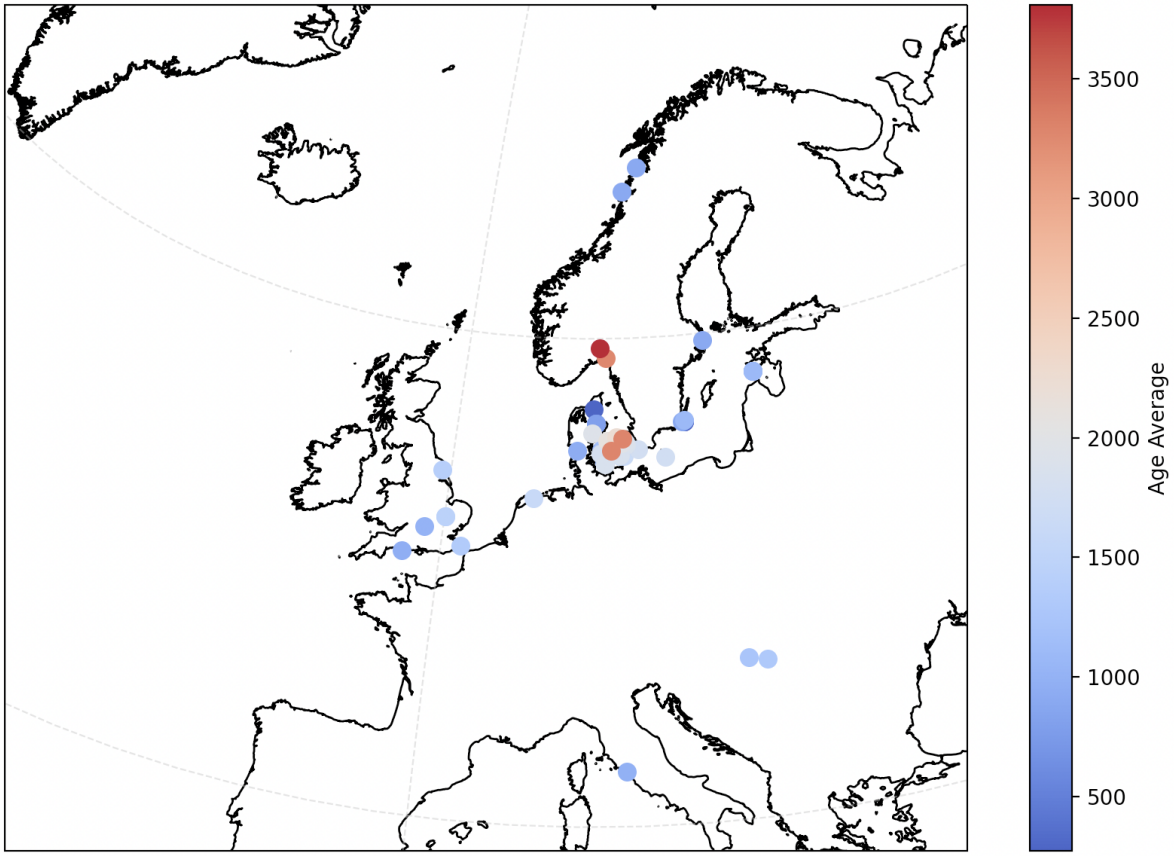


Figure S6.4.2.4: All individuals carrying haplogroup I1a1a (I1a-Z59) from the dataset, coloured by age before the present.

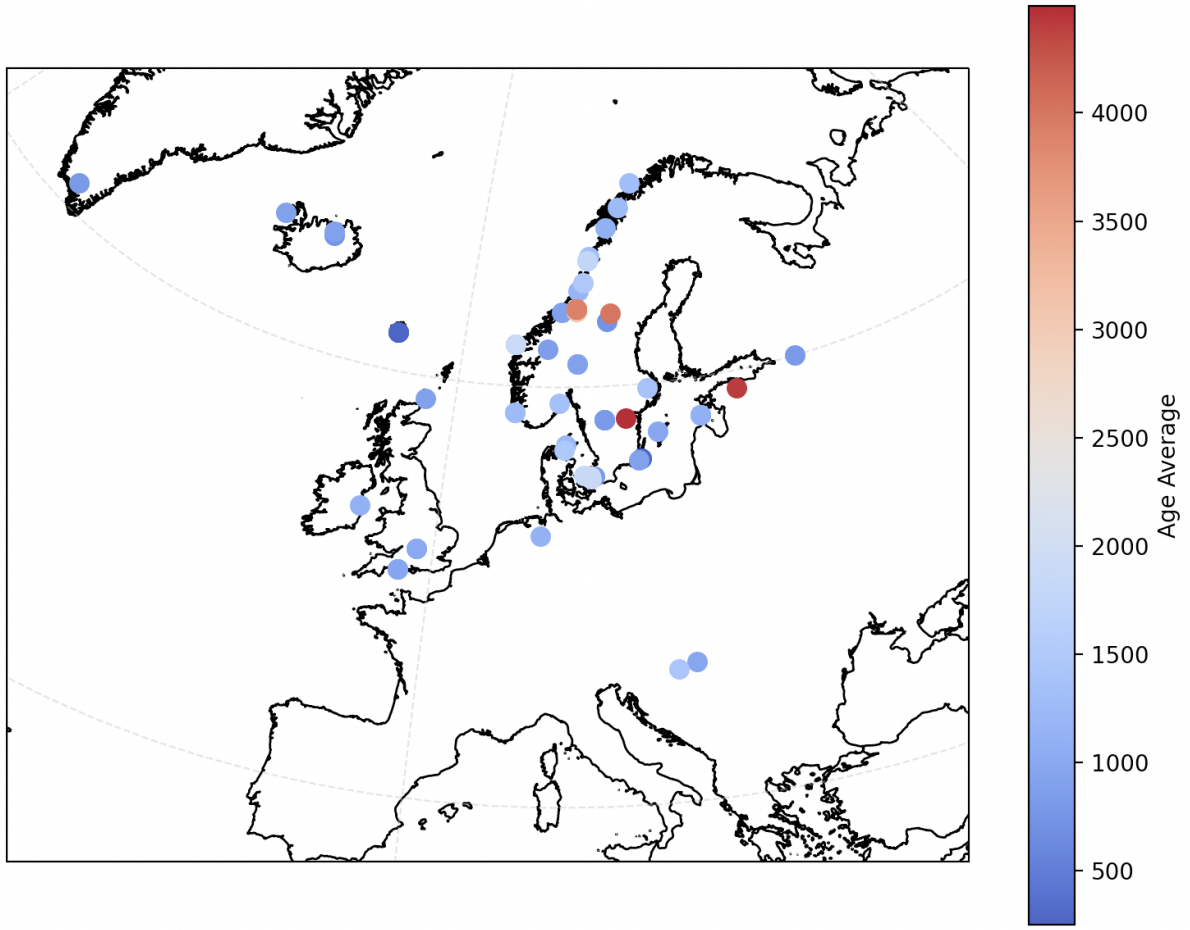


Figure S6.4.2.5: All individuals carrying haplogroup R1a1a1b1a3a (R1b-Z284) from the dataset, coloured by age before the present.

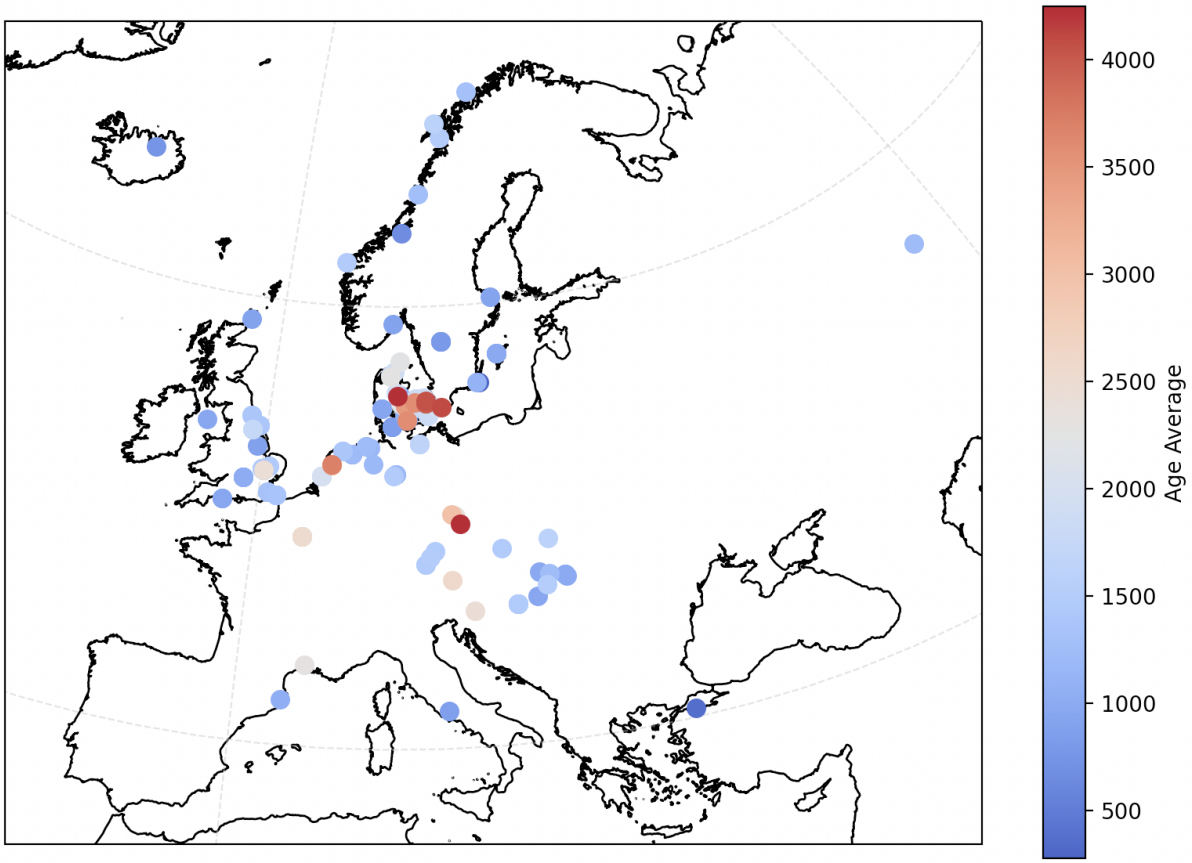


Figure S6.4.2.6: All individuals carrying haplogroup R1b1a1b1a1a1 (R1b-U106) from the dataset, coloured by age before present.

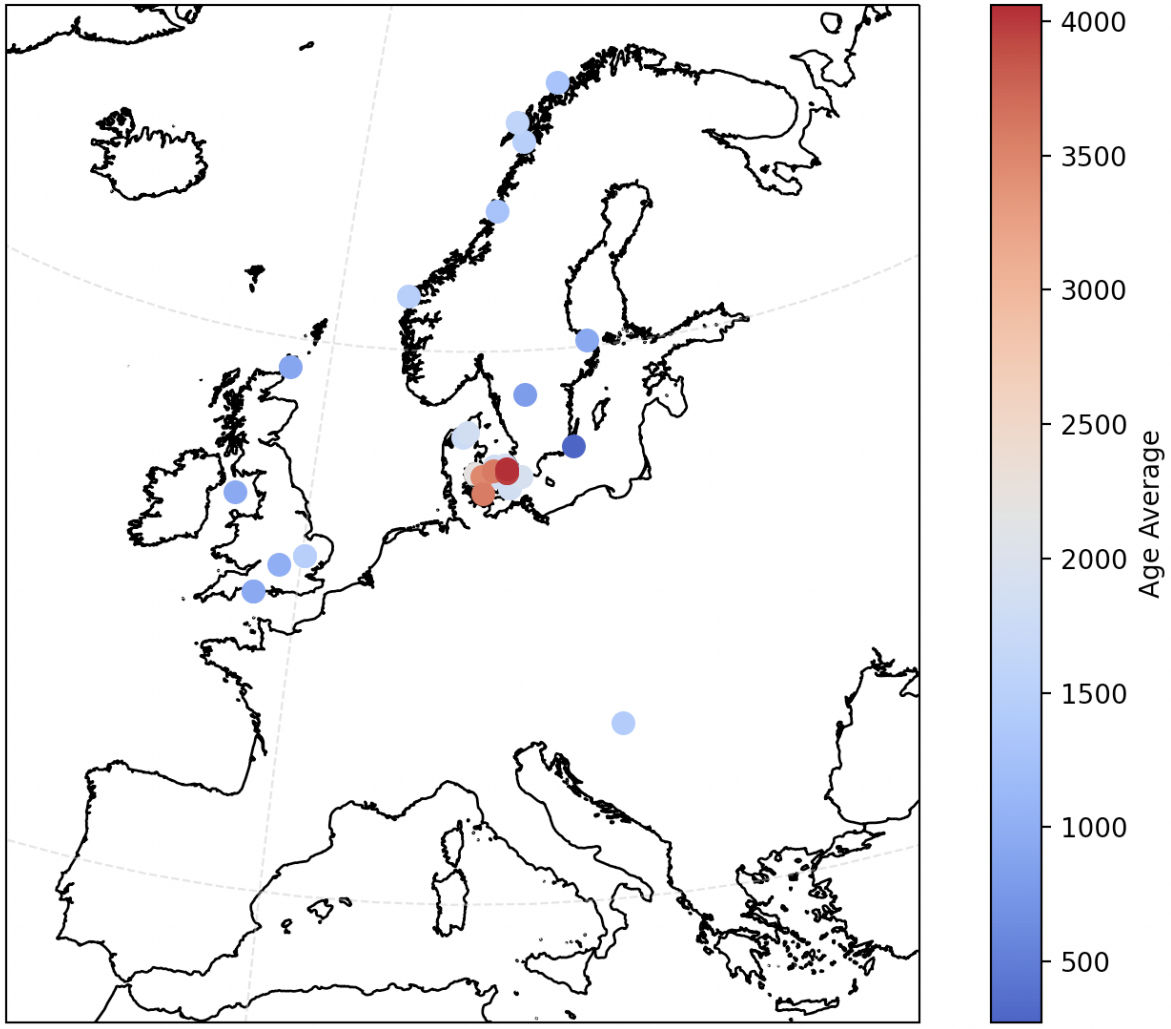


Figure S6.4.2.7: All individuals carrying haplogroup R1b1a1b1a1a1c (R1b-Z19) from the dataset, coloured by age before present.

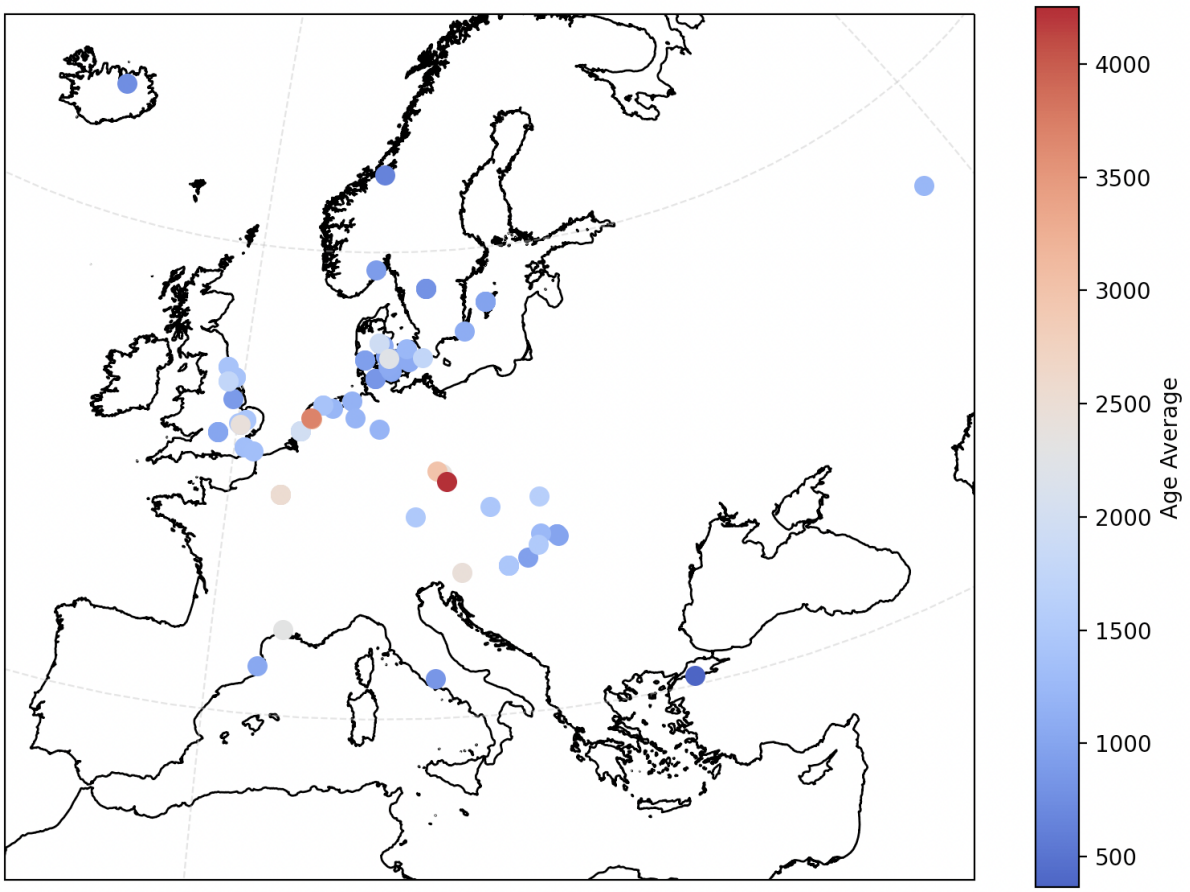


Figure S6.4.2.8: All individuals carrying haplogroup R1b1a1b1a1a1b (R1b-S263) from the dataset, coloured by age before present

References

1. Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature. 2003 Jun;423(6942):825–37.

2. Poznik GD, Henn BM, Yee MC, Sliwerska E, Euskirchen GM, Lin AA, et al. Sequencing Y Chromosomes Resolves Discrepancy in Time to Common Ancestor of Males versus Females. Science. 2013 Aug 2;341(6145):562–5.

3. Jobling MA, Tyler-Smith C. Human Y-chromosome variation in the genome-sequencing era. Nat Rev Genet. 2017 Aug;18(8):485–97.

4. Ebenesersdóttir SS, Sandoval-Velasco M, Gunnarsdóttir ED, Jagadeesan A, Guðmundsdóttir VB, Thordardóttir EL, et al. Ancient genomes from Iceland reveal the making of a human population. Science. 2018 Jun;360(6392):1028–32.

5. Margaryan A, Lawson DJ, Sikora M, Racimo F, Rasmussen S, Moltke I, et al. Population genomics of the Viking world. Nature. 2020 Sep;585(7825):390–6.
6. Villalba-Mouco V, Oliart C, Rihuete-Herrada C, Childebayeva A, Rohrlach AB, Fregeiro MI, et al. Genomic transformation and social organization during the Copper Age–Bronze Age transition in southern Iberia. Sci Adv. 2021 Nov 19;7(47):eabi7038.
7. Moilanen U, Kirkinen T, Saari NJ, Rohrlach AB, Krause J, Onkamo P, et al. A Woman with a Sword? – Weapon Grave at Suontaka Vesitorninmäki, Finland. Eur J Archaeol. 2022 Feb;25(1):42–60.
8. Roca-Rada X, Tereso S, Rohrlach AB, Brito A, Williams MP, Umbelino C, et al. A 1000-year-old case of Klinefelter's syndrome diagnosed by integrating morphology, osteology, and genetics. The Lancet. 2022 Aug;400(10353):691–2.
9. Anastasiadou K, Silva M, Booth T, Speidel L, Audsley T, Barrington C, et al. Detection of chromosomal aneuploidy in ancient genomes. Commun Biol. 2024 Jan 11;7(1):1–9.
10. Diamond J, Bellwood P. Farmers and Their Languages: The First Expansions. Science. 2003 Apr 25;300(5619):597–603.
11. Poznik GD, Xue Y, Mendez FL, Willems TF, Massaia A, Wilson Sayres MA, et al. Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. Nat Genet. 2016 Jun;48(6):593–9.
12. Karmin M, Saag L, Vicente M, Sayres MAW, Järve M, Talas UG, et al. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. Genome Res. 2015 Apr;25(4):459–66.
13. Consortium TYC. A Nomenclature System for the Tree of Human Y-Chromosomal Binary Haplogroups. Genome Res. 2002 Jan 2;12(2):339–48.
14. Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Res. 2008 Jan 5;18(5):830–8.
15. Hallast P, Batini C, Zadik D, Maisano Delser P, Wetton JH, Arroyo-Pardo E, et al. The Y-Chromosome Tree Bursts into Leaf: 13,000 High-Confidence SNPs Covering the Majority of Known Clades. Mol Biol Evol. 2015 Mar 1;32(3):661–73.
16. Bergström A, McCarthy SA, Hui R, Almarri MA, Ayub Q, Danecek P, et al. Insights into human genetic variation and population history from 929 diverse genomes. Science

[Internet]. 2020 Mar 20 [cited 2021 Dec 23]; Available from: https://www.science.org/doi/abs/10.1126/science.aay5012

17. Popejoy AB, Fullerton SM. Genomics is failing on diversity. Nature. 2016 Oct;538(7624):161–4.

18. Sirugo G, Williams SM, Tishkoff SA. The Missing Diversity in Human Genetic Studies. Cell. 2019 Mar 21;177(1):26–31.

19. Kivisild T. The study of human Y chromosome variation through ancient DNA. Hum Genet. 2017 May 1;136(5):529–46.

20. Cappellini E, Prohaska A, Racimo F, Welker F, Pedersen MW, Allentoft ME, et al. Ancient Biomolecules and Evolutionary Inference. Annu Rev Biochem. 2018;87(1):1029–60.

21. Moreno-Mayar JV, Vinner L, de Barros Damgaard P, de la Fuente C, Chan J, Spence JP, et al. Early human dispersals within the Americas. Science. 2018 Dec 7;362(6419):eaav2621.

22. Barbera P, Kozlov AM, Czech L, Morel B, Darriba D, Flouri T, et al. EPA-ng: Massively Parallel Evolutionary Placement of Genetic Sequences. Syst Biol. 2019 Mar 1;68(2):365–9.

23. Martiniano R, De Sanctis B, Hallast P, Durbin R. Placing Ancient DNA Sequences into Reference Phylogenies. Mol Biol Evol. 2022 Feb 1;39(2):msac017.

24. Allentoft ME, Sikora M, Refoyo-Martínez A, Irving-Pease EK, Fischer A, Barrie W, et al. Population Genomics of Stone Age Eurasia [Internet]. bioRxiv; 2022 [cited 2022 Sep 29]. p. 2022.05.04.490594. Available from: https://www.biorxiv.org/content/10.1101/2022.05.04.490594v4

25. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9.

26. Behar DM, van Oven M, Rosset S, Metspalu M, Loogväli EL, Silva NM, et al. A "Copernican" Reassessment of the Human Mitochondrial DNA Tree from its Root. Am J Hum Genet. 2012 Apr 6;90(4):675–84.

27. Scorrano G, Yediay FE, Pinotti T, Feizabadifarahani M, Kristiansen K. The genetic and cultural impact of the Steppe migration into Europe. Ann Hum Biol. 2021 Apr 3;48(3):223–33.

28. Lazaridis I, Alpaslan-Roodenberg S, Acar A, Açıkkol A, Agelarakis A, Aghikyan L, et al. The genetic history of the Southern Arc: A bridge between West Asia and Europe. Science. 2022 Aug 26;377(6609):eabm4247.

S6.5. DATES

Methods

We carried out DATES ((Narasimhan *et al.*, 2019) analyses to estimate the timing of relevant admixture events in this study. For each set of target/source populations listed in tables S6.2.2.1-5, we first subset plink files to relevant individuals and to the 1240k SNP panel. Next, we converted plink files to eigenstrat format using convertf (Patterson, Price and Reich, 2006; Price *et al.*, 2006) and ran DATES using default settings. Lastly, we calculated the absolute admixture date using the relative age of the admixture event from DATES (in generations ago), a mean generation time of 25 years, and the midpoint of the mean radiocarbon ages of all dated individuals in the target population (see figures S6.5.1 and S6.5.2).

Results

S6.5.2.1 Dating admixture of Eastern Scandinavian Bronze Age and Southern Scandinavian Bronze Age in Northern Jutlandic Iron Age individuals

DATES_set	DATES_cluster	sampleId	groupLabel	ageAverage
set1_0_1_2_3t	0_1_3_3	NEO220.allentoft_2023_nature	Sweden_Neolithic	3989
set1_0_1_2_3t	0_1_3_3	NEO224.allentoft_2023_nature	Sweden_Neolithic	3945
set1_0_1_2_3t	0_1_3_3	NEO227.allentoft_2023_nature	Sweden_Neolithic	3948
set1_0_1_2_3t	0_1_3_3	NEO228.allentoft_2023_nature	Sweden_Neolithic	3843
set1_0_1_2_3t	0_1_2_4	NEO878.allentoft_2023_nature	Denmark_Neolithic	4026
set1_0_1_2_3t	0_1_2_3t	CGG019201.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	2014
set1_0_1_2_3t	0_1_2_3t	CGG019203.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1971
set1_0_1_2_3t	0_1_2_3t	CGG019204.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1996
set1_0_1_2_3t	0_1_2_3t	CGG019205.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1915
set1_0_1_2_3t	0_1_2_3t	CGG019206.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1964
set1_0_1_2_3t	0_1_2_3t	CGG019209.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1937
set1_0_1_2_3t	0_1_2_3t	CGG019211.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	2014
set1_0_1_2_3t	0_1_2_3t	CGG019216.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1936
set1_0_1_2_3t	0_1_2_4	CGG106513.mccoll_230817_ironage	Denmark_LateNeolithic	3844
set1_0_1_2_3t	0_1_2_4	CGG106704.mccoll_230817_ironage	Denmark_LateNeolithic	3982
set1_0_1_2_3t	0_1_2_4	CGG106744.mccoll_230817_ironage	Denmark_EarlyBronzeAge	3586

DATES_set	DATES_cluster	sampleId	groupLabel	ageAverage
set2_0_1_2_3t-2114_1909BP	0_1_3_3_4_2800+	NEO220.allentoft_2023_nature	Sweden_Neolithic	3989
set2_0_1_2_3t-2114_1909BP	0_1_3_3_4_2800+	NEO224.allentoft_2023_nature	Sweden_Neolithic	3945
set2_0_1_2_3t-2114_1909BP	0_1_3_3_4_2800+	NEO227.allentoft_2023_nature	Sweden_Neolithic	3948
set2_0_1_2_3t-2114_1909BP	0_1_3_3_4_2800+	NEO228.allentoft_2023_nature	Sweden_Neolithic	3843
set2_0_1_2_3t-2114_1909BP	0_1_2_4_3_2_2_2800+	NEO878.allentoft_2023_nature	Denmark_Neolithic	4026
set2_0_1_2_3t-2114_1909BP	0_1_2_4_3_2_2_2800+	CGG106513.mccoll_230817_ironage	Denmark_LateNeolithic	3844
set2_0_1_2_3t-2114_1909BP	0_1_2_4_3_2_2_2800+	CGG106704.mccoll_230817_ironage	Denmark_LateNeolithic	3982
set2_0_1_2_3t-2114_1909BP	0_1_2_4_3_2_2_2800+	CGG106744.mccoll_230817_ironage	Denmark_EarlyBronzeAge	3586
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019214.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1909
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019212.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1910
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019205.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1915
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019210.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1918
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019216.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1936
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019209.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1937
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG107476.mccoll_230817_ironage	Denmark_IronAge	1942
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019202.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1944
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019206.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1964
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019203.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1971
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019204.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1996
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019201.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	2014
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019211.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	2014
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019208.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	2072
set2_0_1_2_3t-2114_1909BP	0_1_2_3t-2114_1909BP	CGG019645.mccoll_230707_ironage	Denmark_IronAge_PreRomanGrave	2114

DATES_set	DATES_cluster	sampleId	groupLabel	ageAverage
set3_0_1_2_3t-2050_1850BP	0_1_3_3_4_2800+	NEO220.allentoft_2023_nature	Sweden_Neolithic	3989
set3_0_1_2_3t-2050_1850BP	0_1_3_3_4_2800+	NEO224.allentoft_2023_nature	Sweden_Neolithic	3945
set3_0_1_2_3t-2050_1850BP	0_1_3_3_4_2800+	NEO227.allentoft_2023_nature	Sweden_Neolithic	3948
set3_0_1_2_3t-2050_1850BP	0_1_3_3_4_2800+	NEO228.allentoft_2023_nature	Sweden_Neolithic	3843
set3_0_1_2_3t-2050_1850BP	0_1_2_4_3_2_2_2800+	NEO878.allentoft_2023_nature	Denmark_Neolithic	4026
set3_0_1_2_3t-2050_1850BP	0_1_2_4_3_2_2_2800+	CGG106513.mccoll_230817_ironage	Denmark_LateNeolithic	3844
set3_0_1_2_3t-2050_1850BP	0_1_2_4_3_2_2_2800+	CGG106704.mccoll_230817_ironage	Denmark_LateNeolithic	3982
set3_0_1_2_3t-2050_1850BP	0_1_2_4_3_2_2_2800+	CGG106744.mccoll_230817_ironage	Denmark_EarlyBronzeAge	3586
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG106715.mccoll_230817_ironage	Denmark_IronAge	1850
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG106773.mccoll_230817_ironage	Denmark_IronAge	1850
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG106811.mccoll_230817_ironage	Denmark_IronAge	1850
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG106816.mccoll_230817_ironage	Denmark_IronAge	1850
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG106837.mccoll_230817_ironage	Denmark_IronAge	1850
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG107461.mccoll_230817_ironage	Denmark_IronAge	1850
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG107467.mccoll_230817_ironage	Denmark_IronAge	1850
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG107480.mccoll_230817_ironage	Denmark_IronAge	1850
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG016034.mccoll_230707_ironage	Denmark_IronAge_bog_skeleton	1874
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG106817.mccoll_230817_ironage	Denmark_IronAge	1874
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG019214.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1909
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG019212.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1910
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG019205.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1915
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG019210.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1918
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG019216.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1936
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG019209.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1937
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG107476.mccoll_230817_ironage	Denmark_IronAge	1942
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG019202.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1944
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG019206.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1964
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG019203.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1971
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG019204.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	1996
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG019201.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	2014
set3_0_1_2_3t-2050_1850BP	0_1_2_3t-2050_1850BP	CGG019211.mccoll_230707_ironage	Denmark_IronAge_EarlyRomanBogWar	2014

S6.5.2.2 Dating admixture of Eastern Scandinavian Bronze Age and Southern Scandinavian Bronze Age in Danish Isles Iron Age individuals

DATES_set	DATES_cluster	sampleId	groupLabel	ageAverage
set4_0_1_3_2t-1900_1800BP	0_1_3_3_4_2800+	NEO220.allentoft_2023_nature	Sweden_Neolithic	3989
set4_0_1_3_2t-1900_1800BP	0_1_3_3_4_2800+	NEO224.allentoft_2023_nature	Sweden_Neolithic	3945
set4_0_1_3_2t-1900_1800BP	0_1_3_3_4_2800+	NEO227.allentoft_2023_nature	Sweden_Neolithic	3948
set4_0_1_3_2t-1900_1800BP	0_1_3_3_4_2800+	NEO228.allentoft_2023_nature	Sweden_Neolithic	3843
set4_0_1_3_2t-1900_1800BP	0_1_2_4_3_2_2_2800+	NEO878.allentoft_2023_nature	Denmark_Neolithic	4026
set4_0_1_3_2t-1900_1800BP	0_1_2_4_3_2_2_2800+	CGG106513.mccoll_230817_ironage	Denmark_LateNeolithic	3844
set4_0_1_3_2t-1900_1800BP	0_1_2_4_3_2_2_2800+	CGG106704.mccoll_230817_ironage	Denmark_LateNeolithic	3982
set4_0_1_3_2t-1900_1800BP	0_1_2_4_3_2_2_2800+	CGG106744.mccoll_230817_ironage	Denmark_EarlyBronzeAge	3586
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107515.mccoll_230817_ironage	Denmark_IronAge	1807
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107484_rmContam.mccoll_230823_ironage	Denmark_IronAge	1849.5
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106713.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106714.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106716.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106717.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106718.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106719.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106720.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106721.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106722.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106723.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106724.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106726.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106727.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106728.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106730.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG106733.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107389.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107390.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107391.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107393.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107441.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107442.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107443.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107444.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107445.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107446.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107447.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107449.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107450.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107451.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107452.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107453.mccoll_230817_ironage	Denmark_IronAge	1850

set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107456.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107457.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107458.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107460.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107470.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107486.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107488.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107489.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107490.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107494.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107495.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107499.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107501.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107502.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107503.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107504.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107506.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107507.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107508.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107532.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG107533.mccoll_230817_ironage	Denmark_IronAge	1850
set4_0_1_3_2t-1900_1800BP	0_1_3_2t-1900_1800BP	CGG019442.mccoll_230817_ironage	Denmark_IronAge	1875

S6.5.2.3 Dating admixture of Eastern Scandinavian Bronze Age and Southern Scandinavian Bronze Age in admixed Danish Bronze Age individuals

DATES_set	DATES_cluster	sampleId	groupLabel	ageAverage
set5_0_1_3_3_2t-3794_3735BP	0_1_3_3_4_2800+	NEO220.allentoft_2023_nature	Sweden_Neolithic	3989
set5_0_1_3_3_2t-3794_3735BP	0_1_3_3_4_2800+	NEO224.allentoft_2023_nature	Sweden_Neolithic	3945
set5_0_1_3_3_2t-3794_3735BP	0_1_3_3_4_2800+	NEO227.allentoft_2023_nature	Sweden_Neolithic	3948
set5_0_1_3_3_2t-3794_3735BP	0_1_3_3_4_2800+	NEO228.allentoft_2023_nature	Sweden_Neolithic	3843
set5_0_1_3_3_2t-3794_3735BP	0_1_2_4_3_2_2_2800+	NEO878.allentoft_2023_nature	Denmark_Neolithic	4026
set5_0_1_3_3_2t-3794_3735BP	0_1_2_4_3_2_2_2800+	CGG106513.mccoll_230817_ironage	Denmark_LateNeolithic	3844
set5_0_1_3_3_2t-3794_3735BP	0_1_2_4_3_2_2_2800+	CGG106704.mccoll_230817_ironage	Denmark_LateNeolithic	3982
set5_0_1_3_3_2t-3794_3735BP	0_1_2_4_3_2_2_2800+	CGG106744.mccoll_230817_ironage	Denmark_EarlyBronzeAge	3586
set5_0_1_3_3_2t-3794_3735BP	0_1_3_3_2t-3794_3735BP	NEO93.allentoft_2023_nature	Denmark_BronzeAge	3735
set5_0_1_3_3_2t-3794_3735BP	0_1_3_3_2t-3794_3735BP	CGG106702.mccoll_230817_ironage	Denmark_LateNeolithic	3794

Datin

DATES_set	DATES_cluster	sampleId	groupLabel	ageAverage
set6_0_1_3_3_2_2t-3350-3281BP	0_1_3_3_4_2800+	NEO220.allentoft_2023_nature	Sweden_Neolithic	3989
set6_0_1_3_3_2_2t-3350-3281BP	0_1_3_3_4_2800+	NEO224.allentoft_2023_nature	Sweden_Neolithic	3945
set6_0_1_3_3_2_2t-3350-3281BP	0_1_3_3_4_2800+	NEO227.allentoft_2023_nature	Sweden_Neolithic	3948
set6_0_1_3_3_2_2t-3350-3281BP	0_1_3_3_4_2800+	NEO228.allentoft_2023_nature	Sweden_Neolithic	3843
set6_0_1_3_3_2_2t-3350-3281BP	0_1_2_4_3_2_2_2800+	NEO878.allentoft_2023_nature	Denmark_Neolithic	4026
set6_0_1_3_3_2_2t-3350-3281BP	0_1_2_4_3_2_2_2800+	CGG106513.mccoll_230817_ironage	Denmark_LateNeolithic	3844
set6_0_1_3_3_2_2t-3350-3281BP	0_1_2_4_3_2_2_2800+	CGG106704.mccoll_230817_ironage	Denmark_LateNeolithic	3982
set6_0_1_3_3_2_2t-3350-3281BP	0_1_2_4_3_2_2_2800+	CGG106744.mccoll_230817_ironage	Denmark_EarlyBronzeAge	3586
set6_0_1_3_3_2_2t-3350-3281BP	0_1_3_3_2_2t-3350-3281BP	NEO563.allentoft_2023_nature	Denmark_BronzeAge	3350
set6_0_1_3_3_2_2t-3350-3281BP	0_1_3_3_2_2t-3350-3281BP	NEO590.allentoft_2023_nature	Denmark_BronzeAge	3290
set6_0_1_3_3_2_2t-3350-3281BP	0_1_3_3_2_2t-3350-3281BP	CGG106515.mccoll_230817_ironage	Denmark_EarlyBronzeAge	3290
set6_0_1_3_3_2_2t-3350-3281BP	0_1_3_3_2_2t-3350-3281BP	CGG107525.mccoll_230817_ironage	Denmark_EarlyBronzeAge	3281

S6.52.4 Dating admixture of Eastern Scandinavian Bronze Age and Western Scandinavian Bronze Age in admixed Norwegian Bronze Age individuals

DATES_set	DATES_cluster	sampleId	groupLabel	ageAverage
set7_0_1_6_4_1t-3316_3150BP	0_1_3_3_4_2800+	NEO220.allentoft_2023_nature	Sweden_Neolithic	3989
set7_0_1_6_4_1t-3316_3150BP	0_1_3_3_4_2800+	NEO224.allentoft_2023_nature	Sweden_Neolithic	3945
set7_0_1_6_4_1t-3316_3150BP	0_1_3_3_4_2800+	NEO227.allentoft_2023_nature	Sweden_Neolithic	3948
set7_0_1_6_4_1t-3316_3150BP	0_1_3_3_4_2800+	NEO228.allentoft_2023_nature	Sweden_Neolithic	3843
set7_0_1_6_4_1t-3316_3150BP	0_1_6_4_2_2_2800+	CGG105610.mccoll_230817_ironage	Norway_EarlyBronzeAge	3706
set7_0_1_6_4_1t-3316_3150BP	0_1_6_4_2_2_2800+	CGG105612.mccoll_230817_ironage	Norway_LateNeolithic	3763
set7_0_1_6_4_1t-3316_3150BP	0_1_6_4_2_2_2800+	CGG105623.mccoll_230817_ironage	Norway_EarlyBronzeAge	3870
set7_0_1_6_4_1t-3316_3150BP	0_1_6_4_2_2_2800+	CGG105628.mccoll_230817_ironage	Norway_LateNeolithic	3769
set7_0_1_6_4_1t-3316_3150BP	0_1_6_4_2_2_2800+	CGG105642.mccoll_230817_ironage	Norway_BronzeAge	3729
set7_0_1_6_4_1t-3316_3150BP	0_1_6_4_1t-3316_3150BP	CGG105602_CGG105603.mccoll_230821_ironage	Norway_EarlyBronzeAge	3174
set7_0_1_6_4_1t-3316_3150BP	0_1_6_4_1t-3316_3150BP	CGG105635_CGG107018.mccoll_230821_ironage	Norway_EarlyBronzeAge	3316
set7_0_1_6_4_1t-3316_3150BP	0_1_6_4_1t-3316_3150BP	CGG105636_CGG105637.mccoll_230821_ironage	Norway_EarlyBronzeAge	3150
set7_0_1_6_4_1t-3316_3150BP	0_1_6_4_1t-3316_3150BP	CGG105638.mccoll_230817_ironage	Norway_EarlyBronzeAge	3150

S6.5.2.5 Dating admixture of Yamnaya and Latvian-Lithuanian Hunter Gatherers in admixed Eastern Scandinavian Bronze Age individuals

DATES_set	DATES_cluster	sampleId	groupLabel	ageAverage
set8_0_1_3_3_4t-3989_3843BP	0_2_3_3_2800+	RISE546.allentoft_2015_nature	Russia_BronzeAge_Yamnaya	4850
set8_0_1_3_3_4t-3989_3843BP	0_2_3_3_2800+	RISE550.allentoft_2015_nature	Russia_BronzeAge_Yamnaya	4935
set8_0_1_3_3_4t-3989_3843BP	0_2_3_3_2800+	RISE548.allentoft_2015_nature	Russia_BronzeAge_Yamnaya	4850
set8_0_1_3_3_4t-3989_3843BP	0_2_3_3_2800+	RISE547.allentoft_2015_nature	Russia_BronzeAge_Yamnaya	4711
set8_0_1_3_3_4t-3989_3843BP	0_5_1_2_3_2800+	Latvia_HG1.jones_2017_currBiol	Latvia_Mesolithic	8308
set8_0_1_3_3_4t-3989_3843BP	0_5_1_2_3_2800+	Latvia_HG2.jones_2017_currBiol	Latvia_Mesolithic	7689
set8_0_1_3_3_4t-3989_3843BP	0_5_1_2_3_2800+	Spiginas4.mittnik_2018_natComm	Lithuania_Mesolithic	8276
set8_0_1_3_3_4t-3989_3843BP	0_1_3_3_4t-3989_3843BP	NEO228.allentoft_2023_nature	Sweden_Neolithic	3843
set8_0_1_3_3_4t-3989_3843BP	0_1_3_3_4t-3989_3843BP	NEO227.allentoft_2023_nature	Sweden_Neolithic	3948
set8_0_1_3_3_4t-3989_3843BP	0_1_3_3_4t-3989_3843BP	NEO220.allentoft_2023_nature	Sweden_Neolithic	3989
set8_0_1_3_3_4t-3989_3843BP	0_1_3_3_4t-3989_3843BP	NEO224.allentoft_2023_nature	Sweden_Neolithic	3945

DATES_set	DATES_cluster	sampleId	groupLabel	ageAverage
set9_0_1_3_3t-3989_3750BP	0_2_3_3_2800+	RISE546.allentoft_2015_nature	Russia_BronzeAge_Yamnaya	4850
set9_0_1_3_3t-3989_3750BP	0_2_3_3_2800+	RISE550.allentoft_2015_nature	Russia_BronzeAge_Yamnaya	4935
set9_0_1_3_3t-3989_3750BP	0_2_3_3_2800+	RISE548.allentoft_2015_nature	Russia_BronzeAge_Yamnaya	4850
set9_0_1_3_3t-3989_3750BP	0_2_3_3_2800+	RISE547.allentoft_2015_nature	Russia_BronzeAge_Yamnaya	4711
set9_0_1_3_3t-3989_3750BP	0_5_1_2_3_2800+	Latvia_HG1.jones_2017_currBiol	Latvia_Mesolithic	8308
set9_0_1_3_3t-3989_3750BP	0_5_1_2_3_2800+	Latvia_HG2.jones_2017_currBiol	Latvia_Mesolithic	7689
set9_0_1_3_3t-3989_3750BP	0_5_1_2_3_2800+	Spiginas4.mittnik_2018_natComm	Lithuania_Mesolithic	8276
set9_0_1_3_3t-3989_3750BP	0_1_3_3t-3989_3750BP	NEO225.allentoft_2023_nature	Sweden_Neolithic	3750
set9_0_1_3_3t-3989_3750BP	0_1_3_3t-3989_3750BP	ol1009.malmstrom_2019_procRSocB	Sweden_Neolithic_Megalithic	3790
set9_0_1_3_3t-3989_3750BP	0_1_3_3t-3989_3750BP	CGG105650.mccoll_230817_ironage	Norway_LateNeolithic	3810
set9_0_1_3_3t-3989_3750BP	0_1_3_3t-3989_3750BP	NEO228.allentoft_2023_nature	Sweden_Neolithic	3843
set9_0_1_3_3t-3989_3750BP	0_1_3_3t-3989_3750BP	NEO261.allentoft_2023_nature	Sweden_Neolithic	3895
set9_0_1_3_3t-3989_3750BP	0_1_3_3t-3989_3750BP	NEO224.allentoft_2023_nature	Sweden_Neolithic	3945
set9_0_1_3_3t-3989_3750BP	0_1_3_3t-3989_3750BP	NEO227.allentoft_2023_nature	Sweden_Neolithic	3948
set9_0_1_3_3t-3989_3750BP	0_1_3_3t-3989_3750BP	NEO220.allentoft_2023_nature	Sweden_Neolithic	3989

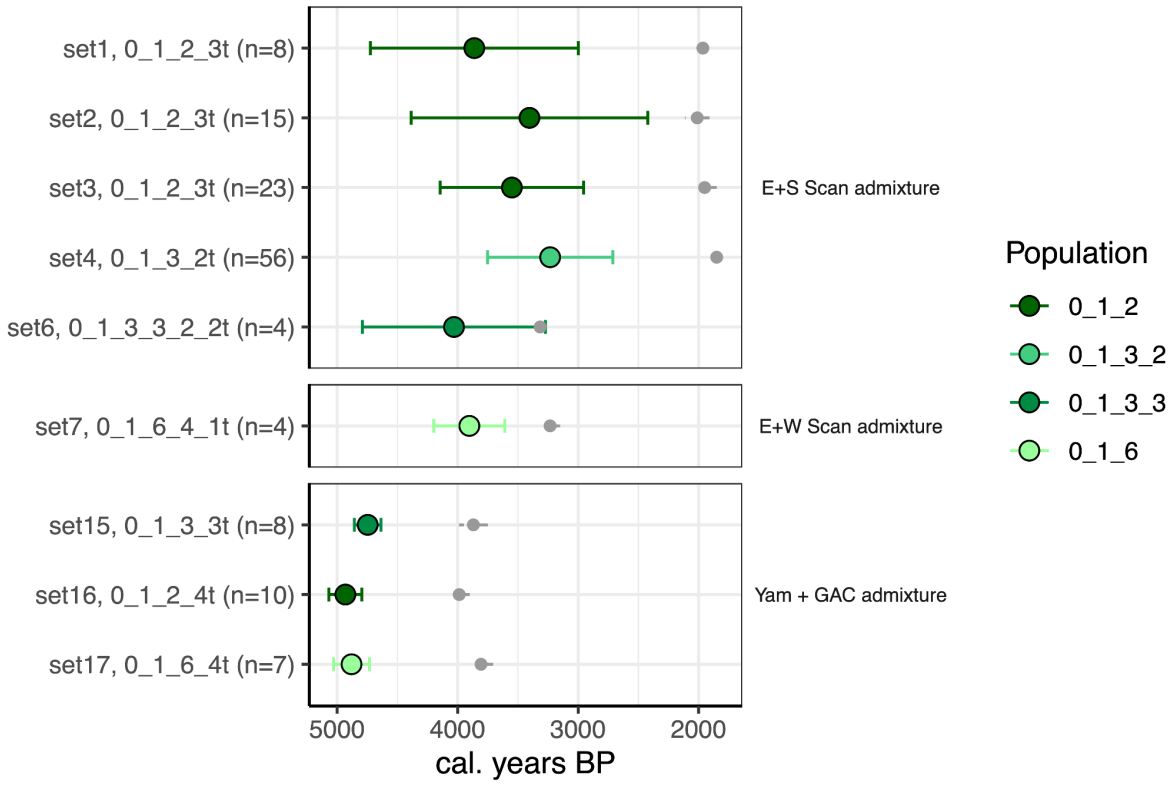


Figure S6.5.1. Suggested date of admixture between Eastern and Southern Bronze Age ancestries (top panel), Eastern and Western Scandinavian Bronze Age ancestries (middle panel), and Yamnaya and Globular Amphora Farmers (lower panel), for a Bronze and Iron Age target populations (0_1_2 = Southern Scandinavians, 0_1_3_2 Danish Isles Eastern Scandinavians, 0_1_3_3 = Bronze Age Eastern Scandinvains, 0_1_6 = Western Scandinavains)

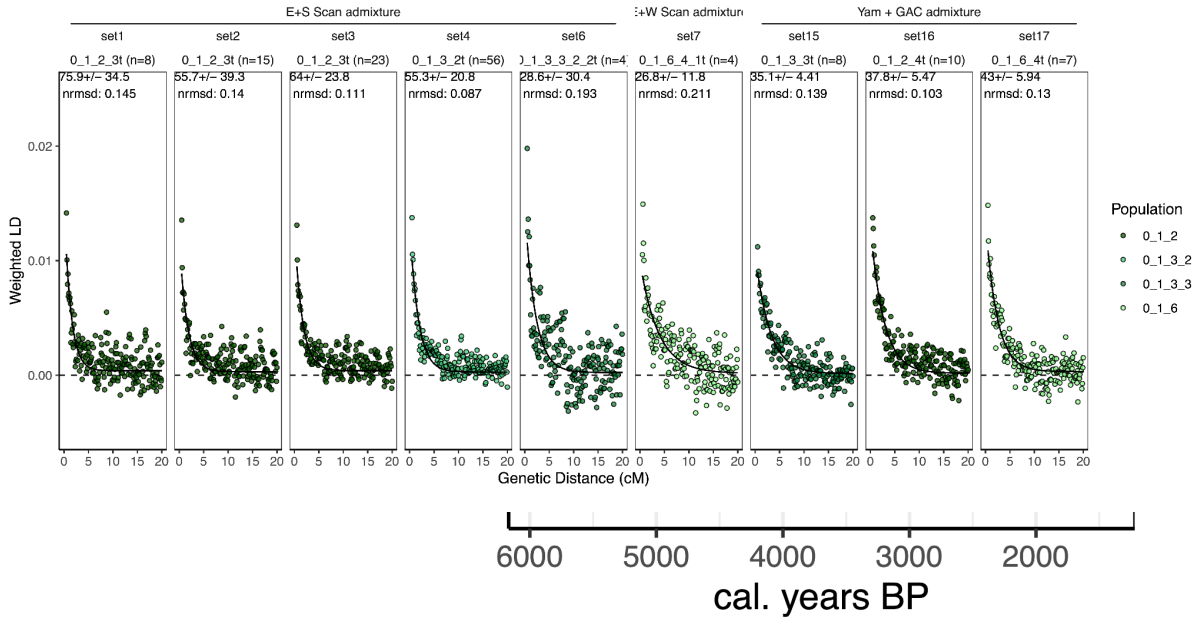


Figure S6.5.2.

S6.6. Kinship

In order to identify pairs of close relatives in our dataset, we ran ngsRelate (v2) on the imputed dataset (see section S5) subset to individuals of European ancestry. We used the pairwise relatedness estimate (rab) from ngsRelate to categorise pairs of individuals into degrees of relatedness (0, 1, 2 or unrelated). We define cutoffs for each degree as the midpoint between the theoretical value of the following and preceeding degrees of relatedness as follows: 0 degree [1, 0.75), 1st degree [0.75, 0.375), 2nd degree [0.375, 0.1875), unrelated [0.1875,0]. Furthermore, for 1st degree relatives we categorise each pair of individuals as either parent-offspring or siblings based on the R0 estimate from ngsRelate as follows: PO [0,2-2), sib [2-2,1].

In total, we find 82 pairs of close relatives among our samples sequenced for this study, represented by 110 individuals. Among the pairs of close relatives we identify 26 parent-offspring relationships, 13 pairs of full siblings, and 43 pairs of 2nd degree relationships (grandparent-grandchild, avuncular relationship or half-siblings). We identify several clusters

where multiple individuals are related. Where possible, we resolved these clusters into pedigrees (Figures 6.8.1-6.8.5). Generally, owing to a relatively low number of samples from each site, we do not find any large extended pedigrees. We do, however, identify multiple examples of relatives coming from different sites. Most striking is the finding of two pairs of siblings and one pair of parent-offspring, all of which have one individual buried at the site Les Moidons and the other at Parançot 2.4 km away (Figure S6.6.4).

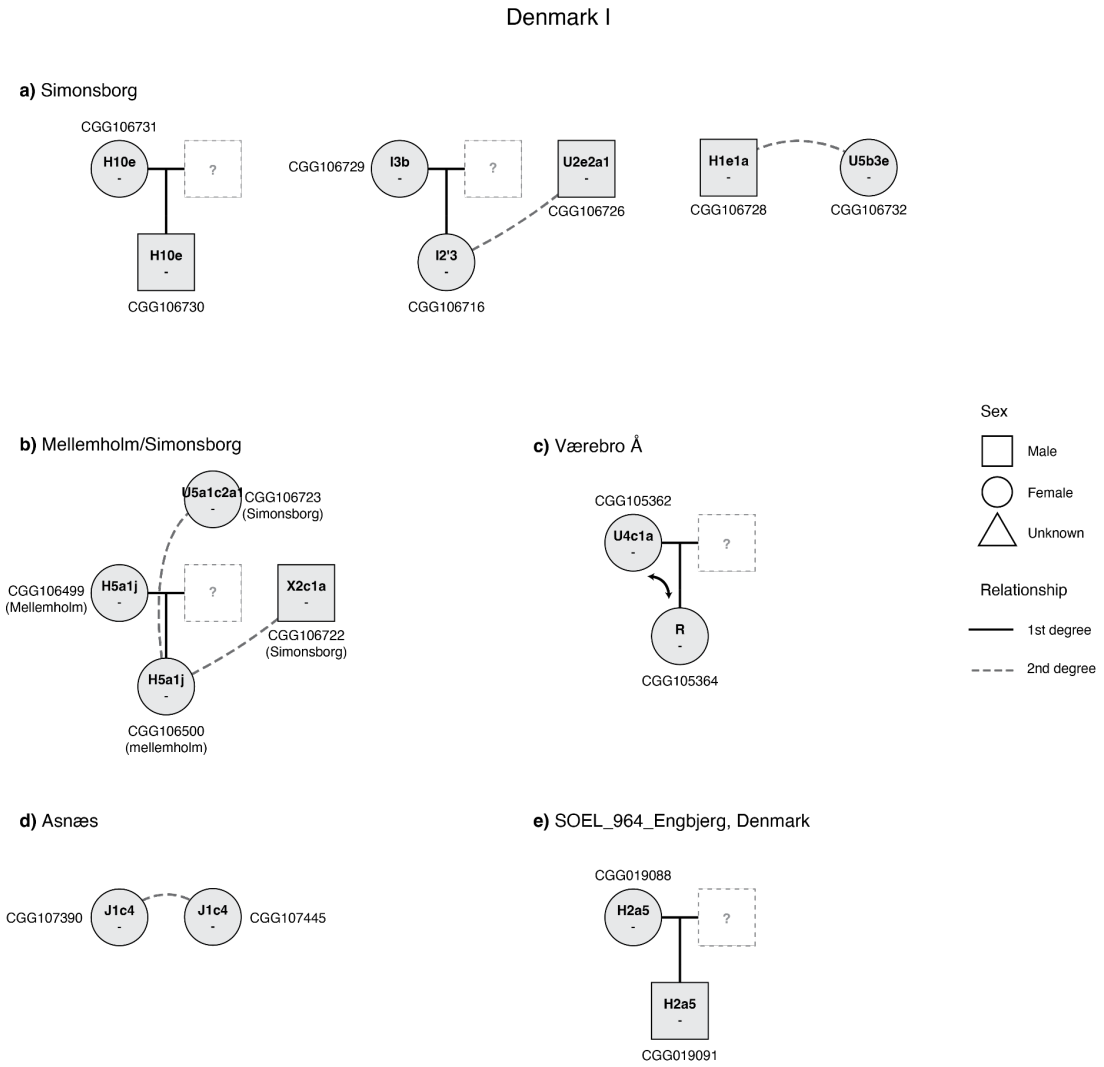


Figure S6.6.1. Pedigrees from Denmark (part 1). Pedigrees were constructed manually, based on results from ngsRelate. Solid black line indicates first degree relationships, whereas stippled line indicates unknown second-degree relationships. Shape for each individual indicates sex. Mitochondrial haplogroup is denoted inside the shape for each individual.

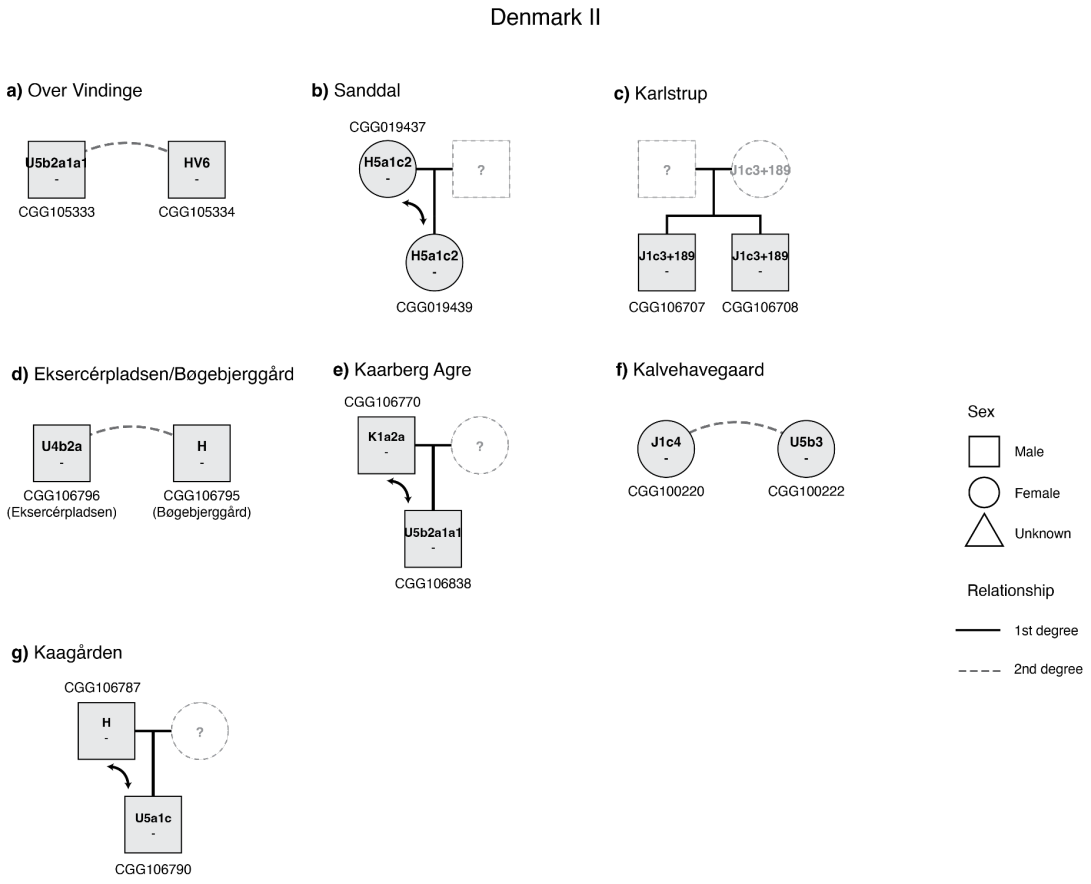


Figure S6.6.2. Pedigrees from Denmark (part 2). Pedigrees were constructed manually, based on results from ngsRelate. Solid black line indicates first degree relationships, whereas stippled line indicates unknown second-degree relationships. Shape for each individual indicates sex. Mitochondrial haplogroup is denoted inside the shape for each individual.

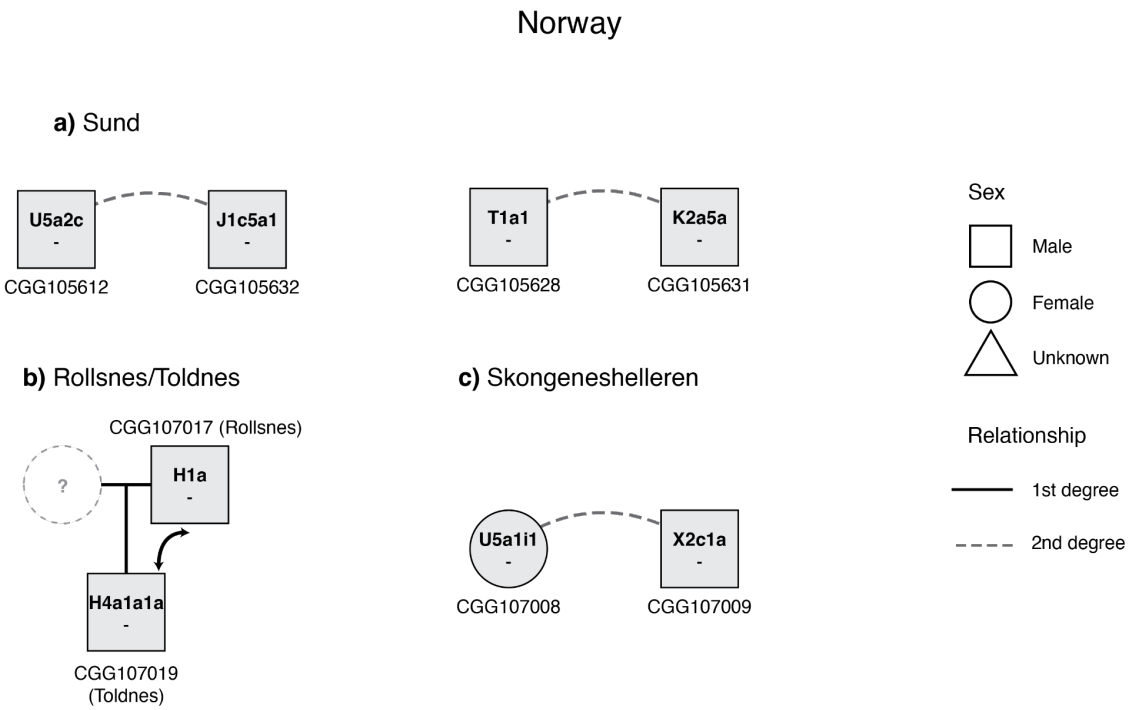


Figure S6.6.3. Pedigrees from Norway. Pedigrees were constructed manually, based on results from ngsRelate. Solid black line indicates first degree relationships, whereas stippled line indicates unknown second-degree relationships. The shape for each individual indicates sex. Mitochondrial haplogroup is denoted inside the shape for each individual. The incompatibility of the context dates between the Rollsnes and Toldnes individuals suggest C14 dating these individuals is required before the context can be properly understood.

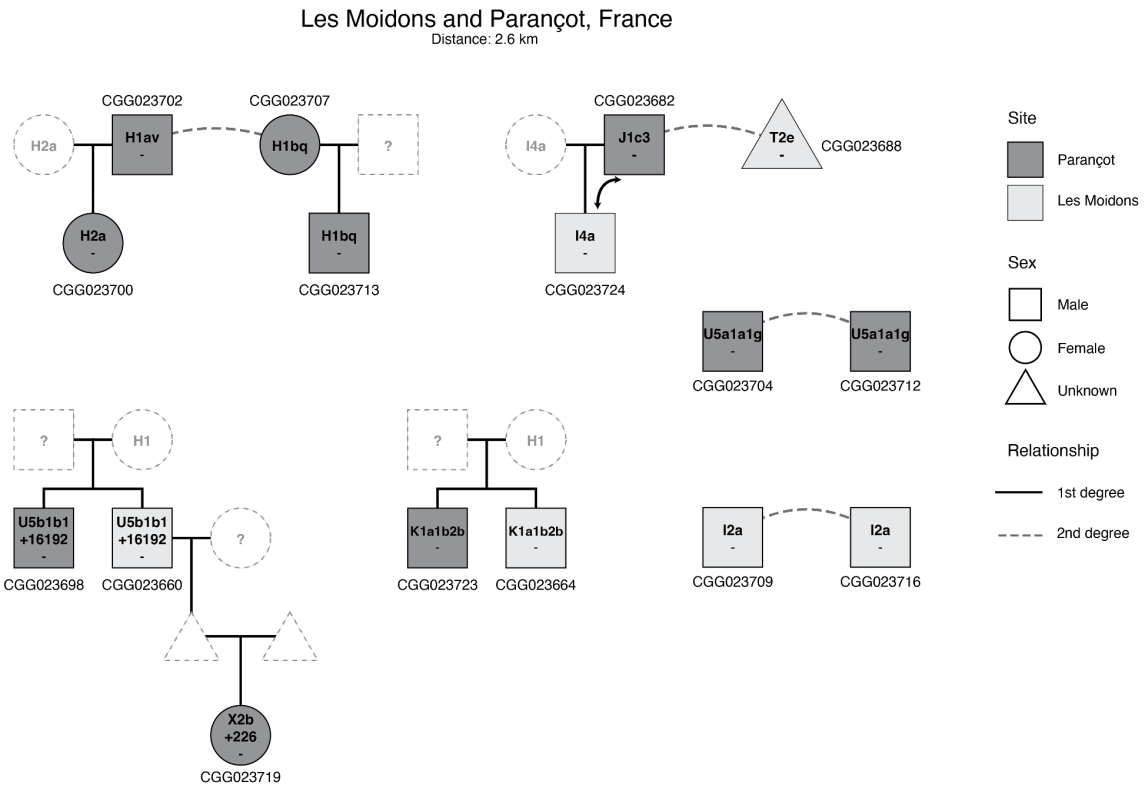


Figure S6.6.4. Pedigrees from Les Moidons and Parançot, France. Pedigrees were constructed manually, based on results from ngsRelate. Solid black line indicates first degree relationships, whereas stippled line indicates unknown second-degree relationships. Color and shape for each individual indicate site and sex respectively. Mitochondrial haplogroup is denoted inside the shape for each individual.

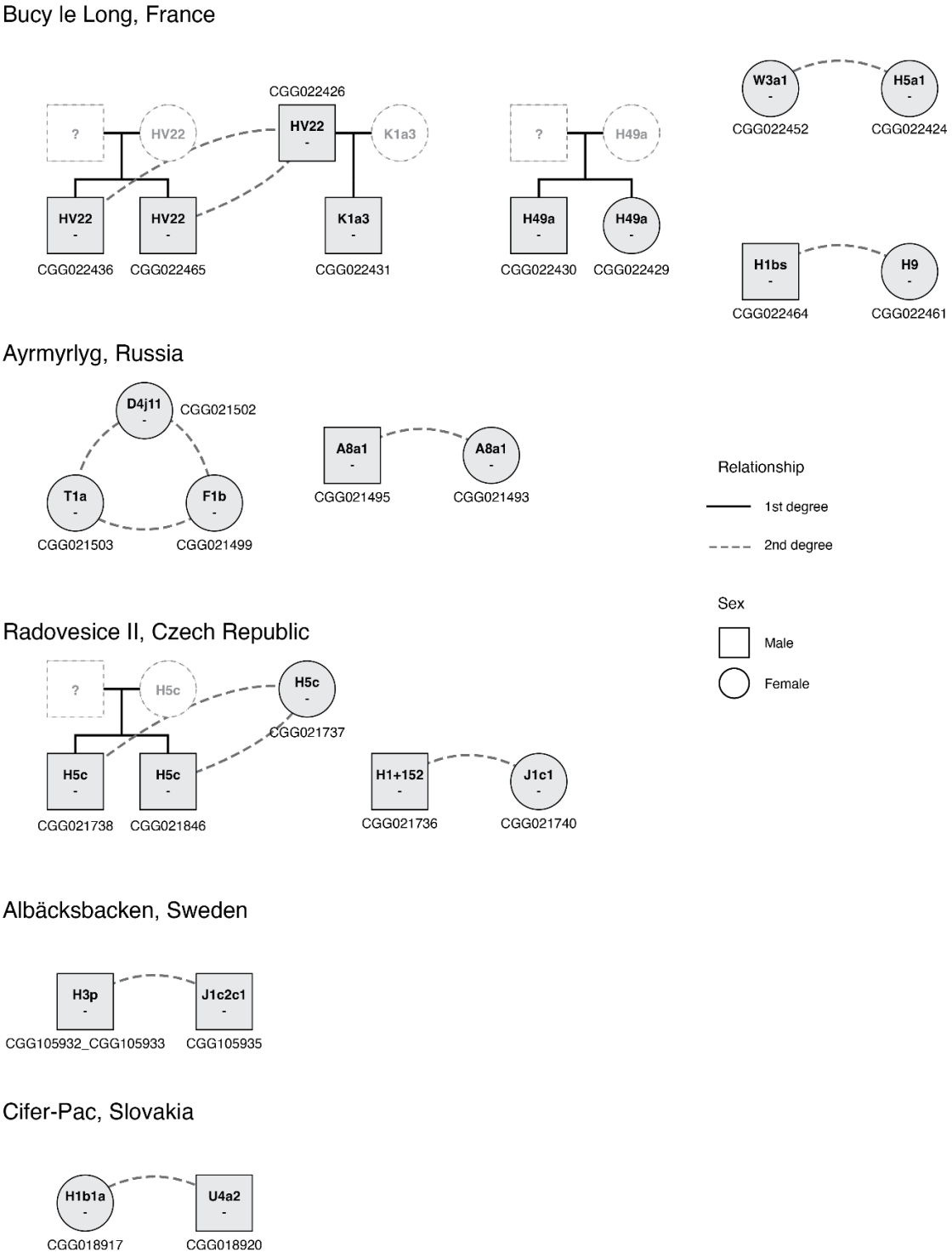


Figure S6.6.5. Smaller pedigrees from various sites. Pedigrees were constructed manually, based on results from ngsRelate. Solid black line indicates first degree relationships, whereas stippled line indicates unknown second-degree relationships. Shape for each individual indicate sex. Mitochondrial haplogroup is denoted inside the shape for each individual.

S6.7. Extended discussion

S6.7.1 Genetic Outliers

The combination of dense sampling and high resolution IBD modelling from the Iron Age sources also allows for the identification of outliers and their origins within Scandinavia. From Northern Jutland, 5 individuals cluster together with Iron Age Norwegians (CGG106489, Sondrup Østergaard, Ulstrup sogn, CGG106503, Hyldebjerg, Vaarst CGG106553, Hellevad, CGG106810, Mellemholm, CGG107417, Gammel Hasseris Grusgrav), representing the contact between the two regions at a time in which crossing the sea to the North was relatively easy compared to travelling south on land (Christiansen, 2017). Throughout the Bronze and Iron Age, individuals with ancestral origins across the Baltic were found in Norway and Denmark (CGG105601, CGG106751, CGG105643, CGG107034, CGG106747, CGG106748, CGG106486_CGG106491). Individuals from the Salme Viking burial on the Island of Saaremaa cluster together with and are modelled with similar profiles to those of Iron Age of Central Sweden rather than the later Vikings, supporting origins in the Mälaren valley of Sweden. At Kalmargården, two beheaded individuals appear to be non-locals, one (CGG107579, 2654/79, East) clustering with and modelled with ancestry of Norway, and the second (CGG107580, 2654/79, West) of the Baltic region. In pre-Viking Norway, as early as ~1240 BP we find individuals with Celtic British Isles ancestry, representing early migration between the two regions.

S6.7.2 The Netherlands

The Bell Beaker sub-cluster 0_2_1_2 WEuIsMl located primarily from the Eastern North Sea (ENS) region (present day the Netherlands) is unique in its high NWHG ancestry, low European Farmer, and inability to be modelled primarily as Bell Beaker ancestry, like most others from Bell Beaker sub-clusters (Figure S6.3.6).

When using early representatives of this cluster as a source, we see a large degree of genetic continuity from 3700 - 1700 BP. From Valkenburg (ZH) however, there are a number of individuals that do not fit the profile. The Roman cemetery Valkenburg Marktveld, located south of the auxiliary fort, was used between 50 – 300 CE for the entire military community that consisted of men, women and children, who lived in the vicinity of the auxiliary fort. Over 650 individuals were recovered, 145 of which are inhumations (41 adults, 104 children and

infants); an extraordinary number as cremation dominates the Roman burial record in the Netherlands. The individuals included in this study are possibly associated with different departments of the Roman army, so the presence of non-local individuals is not unexpected (DeCoster et al., in prep.). For the individuals that do not fit the local profile, most are Celtic (similar to contemporaneous people from British Isles or France and one is similar to people deriving their ancestry from the Bronze Age Eastern Mediterranean. A single individual is modelled with North East European ancestry.

A transition by at least 1612 BP is apparent, Frisian individuals are modelled primarily as Southern Scandinavian ancestries, but possessing small amounts of the local ENS ancestry.

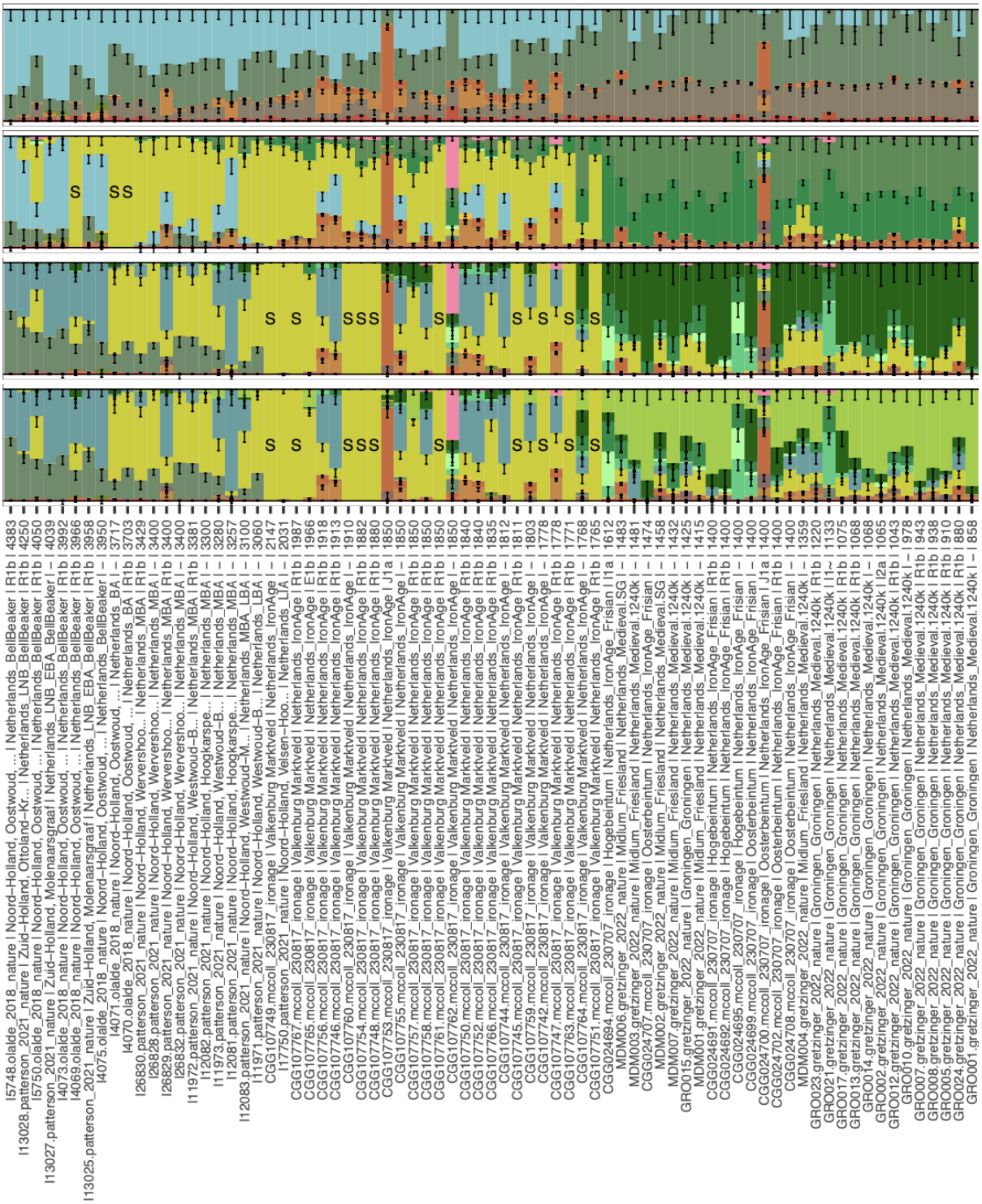


Figure S6.7.2.1. Subset of Mixture Modellings Results from Figure S6.3.1, for the Netherlands

S6.7.4 Norway

As discussed in Supplementary Note S6.2.2 and shown in Figure S6.2.2.1A, from the Bronze Age the expansions of Eastern Scandinavians had a large impact on the population structure of Norway. Despite only a few % of local Bronze Age Norwegian ancestry being present in Iron Age in Norway, we see the majority of Norwegian individuals from the Bronze Age, Iron Age and Viking Period clustering within the deep 0_1_6 Corded Ware (North) sub-cluster. This appears to result from some admixed late Bronze Age individuals within this cluster who are modelled with ~50% local Bronze Age ancestry, and who themselves model 50% of the Iron Age Ancestry (Figure S6.2.2.1A). The remaining 50% is Eastern Scandinavian BA ancestry, suggesting multiple waves or constant migration from the east to the west by the Iron Age. From the Iron Age onwards, there seems to be genetic continuity through the Viking period. Unlike further south in Scandinavia, the Vikings in Norway appear to descend from the local Iron Age individuals.

In addition to the interactions between Jutland and Norway (Supplementary Note S6.7.1), we also find evidence of interactions with Celtic Britain and Ireland, during the Late Iron Age and Migration Period. Many of these individuals are not admixed with the local Norwegians, suggesting they are visitors, first generation migrations,or remain genetically isolated from the local populations.

S6.7.5 Langobards and Goths

To compare with the recently published (Stolarek *et al.*, 2023) genomes, we generated a new panel of imputed genomes and re-clustered with the additional genomes as described in Supplementary Notes S5 and S6. We re-ran mixture modelling using the same source and target clusters as described in S6, but with the new samples in additional target clusters. Our results can be seen in Figure S6.7.6.1 and S6.7.6.2, and further reinforce our previous results.

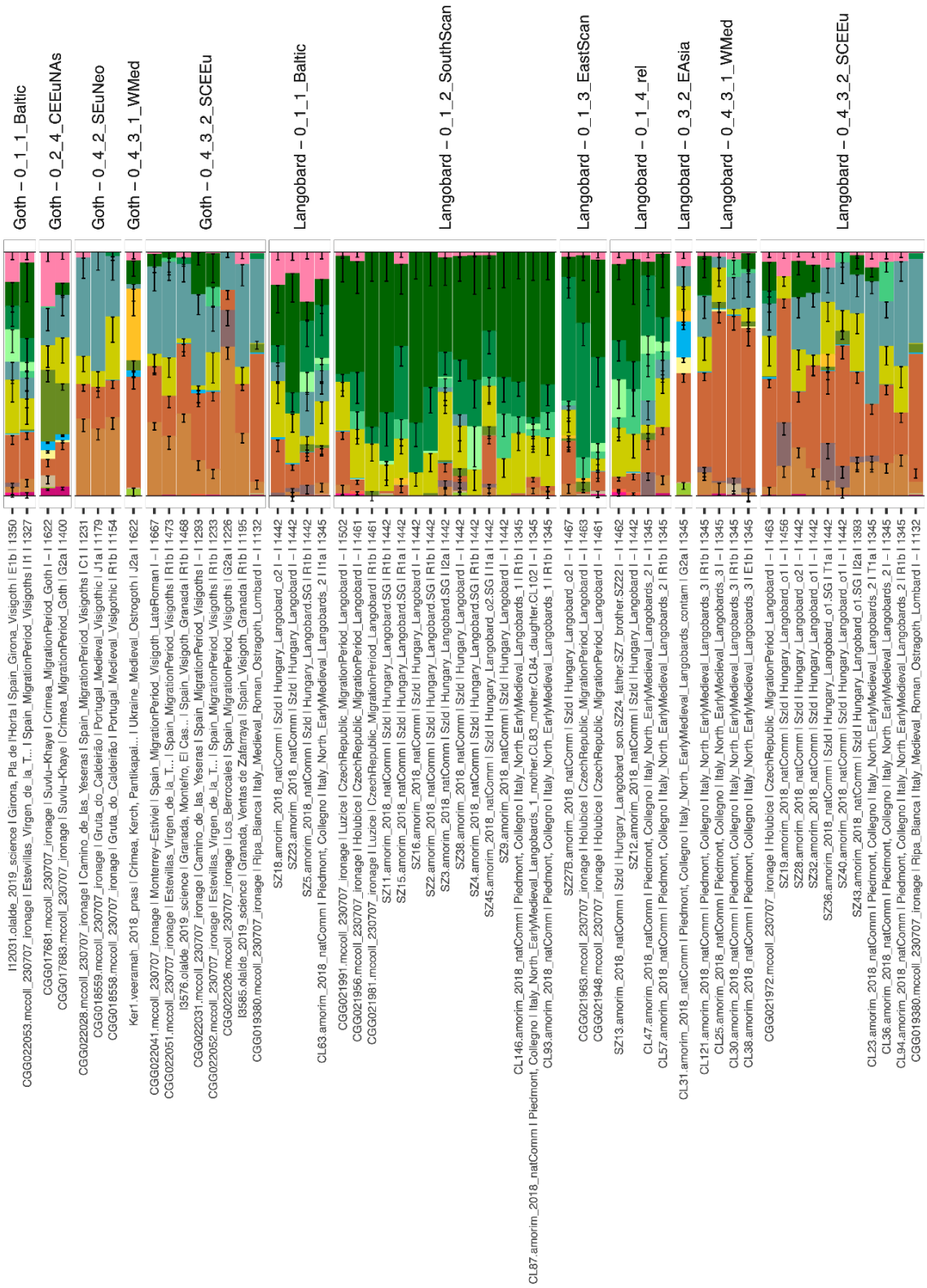


Figure S6.7.6.1 Mixture Modelling Results (Set 7) for Langobards and Goth, faceted by broad IBD clusters.

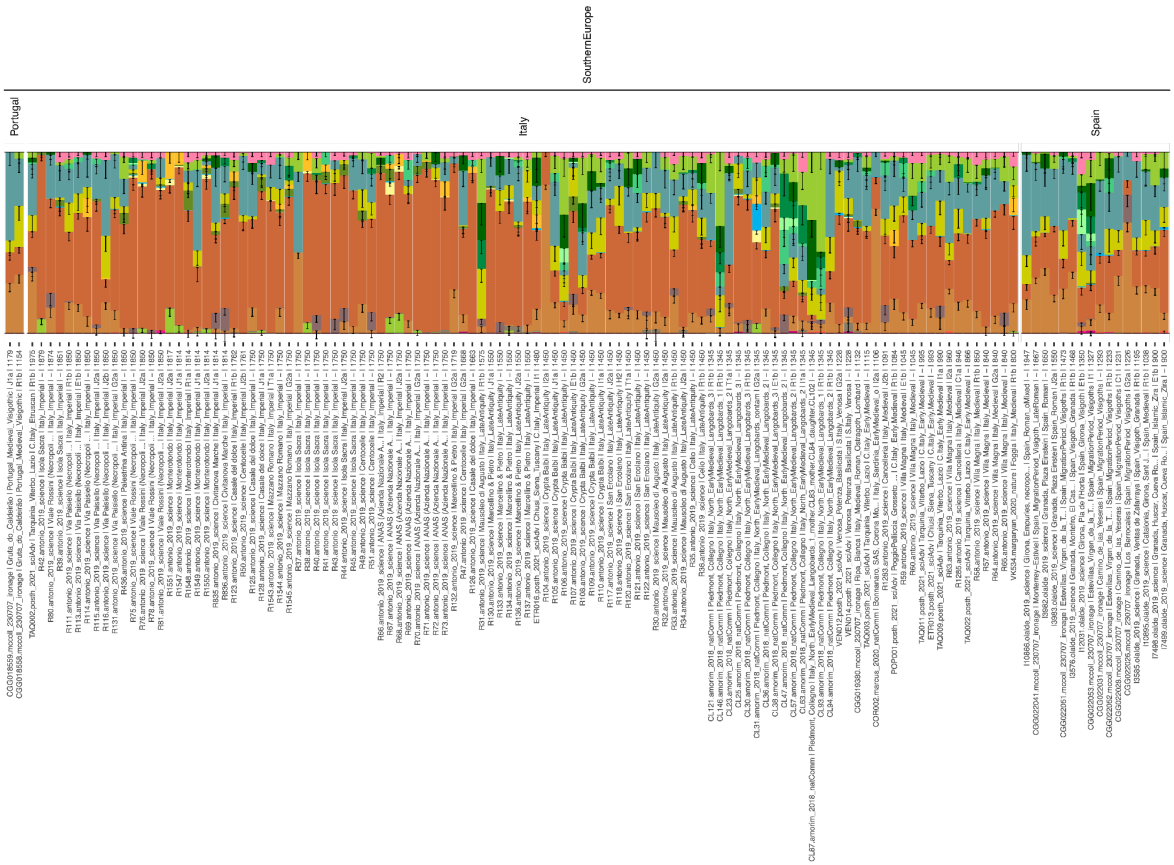


Figure S6.7.6.2. Mixture Modelling Results (Set 8) for Italy and Iberia. Excluding the Langobards, only three individuals cluster as Southern Scandinavians (R31, R106 and R108 from) (Antonio *et al.*, 2019)

S6.7.7. Britain and Ireland

After the arrival of Southern Scandinavian ancestry to the British Isles around 1500 BP, the majority of samples are modelled with North Germanic (Mecklenburg) ancestry and fall within the Southern Scandinavian subclusters 0_1_2_1 (n= 80), 0_1_2_2 (n = 18), 0_1_2_4 (n=1) and 0_1_2_5 (n=25).

Figure S6.7.7. Full set of mixture modelling results for set X for Britain and its Isles.

There are a number of individuals present that otherwise cluster with individuals from Northern Europe. Throughout the Saxon and Viking periods, n = 11 individuals fall within the Southern Scandinavian - Jutland sub cluster (0_1_2_3). The remaining individuals are primarily from the Viking period: the Baltic cluster (0_1_1, n = 1), the Eastern Scandinavian cluster (0_1_3, n = 5), the Western Scandinavian sub-cluster (n = 14)

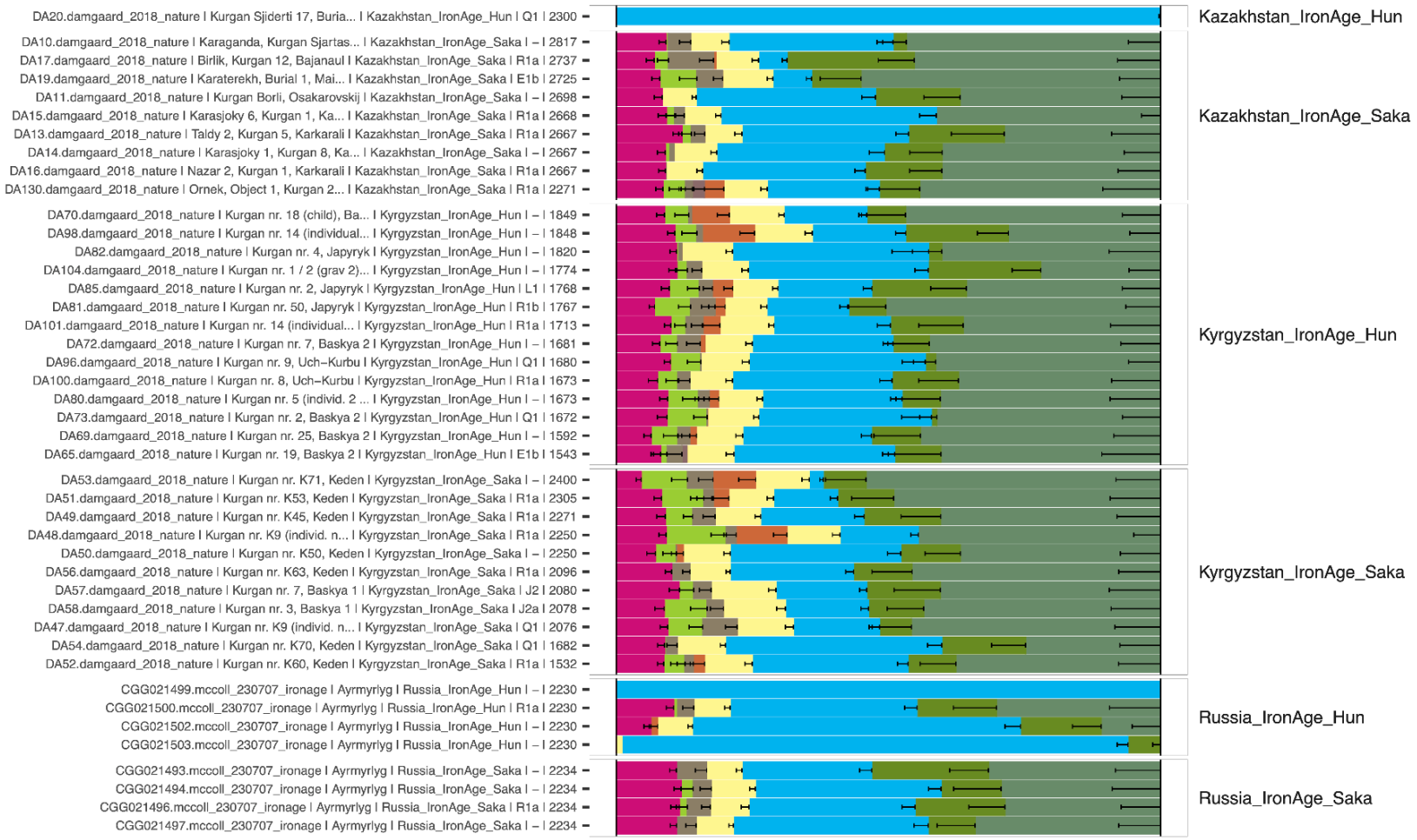
These outlier individuals, their age (BP) and specific subcluster are listed below.

Baltic-cluster				
VK145.margaryan_2020_nature	England	Britain_VikingAge	1010	0_1_1_2_2800-
Eastern Scandinavian				
BUK037.gretzinger_2022_nature	England	England_Medieval.1240k	1350	0_1_3_1_2_1_1_2800-
VK176.margaryan_2020_nature	England	Britain_VikingAge	1010	0_1_3_1_2800-
VK175.margaryan_2020_nature	England	Britain_VikingAge	1010	0_1_3_1_1_2_2_2800-
VK174.margaryan_2020_nature	England	Britain_VikingAge	1010	0_1_3_1_2800-
VK168.margaryan_2020_nature	England	Britain_VikingAge	1010	0_1_3_1_2800-
Southern Scandinavian - Jutland cluster				
HAD002.gretzinger_2022_nature	England	England_Medieval.SG	1472	0_1_2_3_2800-
BUK039.gretzinger_2022_nature	England	England_Medieval.1240k	1470	0_1_2_3_2_1_2_2800-
I20654.gretzinger_2022_nature	England	England_Medieval.1240k	1450	0_1_2_3_4_2_2800-
I20674.gretzinger_2022_nature	England	England_Medieval.1240k	1400	0_1_2_3_4_2_2800-
BUK057.gretzinger_2022_nature	England	England_Medieval.1240k	1350	0_1_2_3_2800-
BUK040.gretzinger_2022_nature	England	England_Medieval.1240k	1350	0_1_2_3_1_2_2_2800-
BUK030.gretzinger_2022_nature	England	England_Medieval.1240k	1350	0_1_2_3_2_1_1_2800-
BUK029.gretzinger_2022_nature	England	England_Medieval.1240k	1350	0_1_2_3_2_1_2_2800-
BUK001.gretzinger_2022_nature	England	England_Medieval.1240k	1350	0_1_2_3_2_1_2_2800-
VK165.margaryan_2020_nature	England	Britain_VikingAge	1010	0_1_2_3_2800-
I3044.gretzinger_2022_nature	England	England_Medieval.1240k	895	0_1_2_3_2800-
Western Scandinavian cluster				
VK546.margaryan_2020_nature	Ireland	Ireland_Viking.SG	1100	0_1_6_2_2_1_1_1_2800-
VK172.margaryan_2020_nature	England	Britain_VikingAge	1010	0_1_6_2_2800-
VK144.margaryan_2020_nature	England	Britain_VikingAge	1010	0_1_6_6_2800-
VK544.margaryan_2020_nature	Ireland	Ireland_Medieval	1000	0_1_6_2_2_3_2_2800-
VK543.margaryan_2020_nature	Ireland	Ireland_Medieval	1000	0_1_6_1_1_2800-
VK449.margaryan_2020_nature	England	Britain_VikingAge	953	0_1_6_1_3_2800-
VK263.margaryan_2020_nature	England	Britain_VikingAge	953	0_1_6_1_1_2_2_2800-

VK262.margaryan_2020_nature	England	Britain_VikingAge	953	0_1_6_1_1_2_1_2800-
VK260.margaryan_2020_nature	England	Britain_VikingAge	953	0_1_6_1_1_2800-
VK259.margaryan_2020_nature	England	Britain_VikingAge	953	0_1_6_1_3_2800-
VK258.margaryan_2020_nature	England	Britain_VikingAge	953	0_1_6_2_1_1_2_2800-
VK257.margaryan_2020_nature	England	Britain_VikingAge	953	0_1_6_2_5_2_2800-
VK170.margaryan_2020_nature	IsleOfMan	Britain_VikingAge	950	0_1_6_2_2800-
VK204.margaryan_2020_nature	Scotland	Britain_VikingAge	900	0_1_6_7_1_2800-

S6.7.8. Ayrmyrlyg

The site of Ayrmyrlyg in Russia around 2200 BP is home to individuals from both the Saka and Hun cultures. Under IBD Mixture Modelling Set 5, the Saka individuals are modelled with similar ancestry proportions to the Sakas of Kazakhstan and Kyrgyzstan, with ~50% Steppe ancestry, smaller proportions of Iranian, Neolithic Farming and Caucausus Hunter Gatherer and Eastern Hunter-gatherer ancestry. Of the Huns, two are modelled here entirely as East Asian ancestry (light blue), similar to the Hun from Kazakhstan. The third Hun is modelled similar to the Saka (CGG021500), and the last has proportions expected from an individual deriving 50% of their ancestry from an East Asian source, and 50% from a Saka source.


DA20.damgaard_2018_nature | Kurgan Sjiderti 17, Buria... | Kazakhstan_IronAge_Hun | Q1 | 2300
DA10.damgaard_2018_nature | Karaganda, Kurgan Sjartas... | Kazakhstan_IronAge_Saka | – | 2817
DA17.damgaard_2018_nature | Birlik, Kurgan 12, Bajanaul | Kazakhstan_IronAge_Saka | R1a | 2737
DA19.damgaard_2018_nature | Karaterekh, Burial 1, Mai... | Kazakhstan_IronAge_Saka | E1b | 2725
DA11.damgaard_2018_nature | Kurgan Borli, Osakarovskij | Kazakhstan_IronAge_Saka | – | 2698
DA15.damgaard_2018_nature | Karasjoky 6, Kurgan 1, Ka... | Kazakhstan_IronAge_Saka | R1a | 2668
DA13.damgaard_2018_nature | Taldy 2, Kurgan 5, Karkarali | Kazakhstan_IronAge_Saka | R1a | 2667
DA14.damgaard_2018_nature | Karasjoky 1, Kurgan 8, Ka... | Kazakhstan_IronAge_Saka | – | 2667
DA16.damgaard_2018_nature | Nazar 2, Kurgan 1, Karkarali | Kazakhstan_IronAge_Saka | R1a | 2667
DA130.damgaard_2018_nature | Ornek, Object 1, Kurgan 2... | Kazakhstan_IronAge_Saka | R1a | 2271
DA70.damgaard_2018_nature | Kurgan nr. 18 (child), Ba... | Kyrgyzstan_IronAge_Hun | – | 1849
DA98.damgaard_2018_nature | Kurgan nr. 14 (individual... | Kyrgyzstan_IronAge_Hun | – | 1848
DA82.damgaard_2018_nature | Kurgan nr. 4, Japyryk | Kyrgyzstan_IronAge_Hun | – | 1820
DA104.damgaard_2018_nature | Kurgan nr. 1 / 2 (grav 2)... | Kyrgyzstan_IronAge_Hun | – | 1774
DA85.damgaard_2018_nature | Kurgan nr. 2, Japyryk | Kyrgyzstan_IronAge_Hun | L1 | 1768
DA81.damgaard_2018_nature | Kurgan nr. 50, Japyryk | Kyrgyzstan_IronAge_Hun | R1b | 1767
DA101.damgaard_2018_nature | Kurgan nr. 14 (individual... | Kyrgyzstan_IronAge_Hun | R1a | 1713
DA72.damgaard_2018_nature | Kurgan nr. 7, Baskya 2 | Kyrgyzstan_IronAge_Hun | – | 1681
DA96.damgaard_2018_nature | Kurgan nr. 9, Uch−Kurbu | Kyrgyzstan_IronAge_Hun | Q1 | 1680
DA100.damgaard_2018_nature | Kurgan nr. 8, Uch−Kurbu | Kyrgyzstan_IronAge_Hun | R1a | 1673
DA80.damgaard_2018_nature | Kurgan nr. 5 (individ. 2 ... | Kyrgyzstan_IronAge_Hun | – | 1673
DA73.damgaard_2018_nature | Kurgan nr. 2, Baskya 2 | Kyrgyzstan_IronAge_Hun | Q1 | 1672
DA69.damgaard_2018_nature | Kurgan nr. 25, Baskya 2 | Kyrgyzstan_IronAge_Hun | – | 1592
DA65.damgaard_2018_nature | Kurgan nr. 19, Baskya 2 | Kyrgyzstan_IronAge_Hun | E1b | 1543
DA53.damgaard_2018_nature | Kurgan nr. K71, Keden | Kyrgyzstan_IronAge_Saka | – | 2400
DA51.damgaard_2018_nature | Kurgan nr. K53, Keden | Kyrgyzstan_IronAge_Saka | R1a | 2305
DA49.damgaard_2018_nature | Kurgan nr. K45, Keden | Kyrgyzstan_IronAge_Saka | R1a | 2271
DA48.damgaard_2018_nature | Kurgan nr. K9 (individ. n... | Kyrgyzstan_IronAge_Saka | R1a | 2250
DA50.damgaard_2018_nature | Kurgan nr. K50, Keden | Kyrgyzstan_IronAge_Saka | – | 2250
DA56.damgaard_2018_nature | Kurgan nr. K63, Keden | Kyrgyzstan_IronAge_Saka | R1a | 2096
DA57.damgaard_2018_nature | Kurgan nr. 7, Baskya 1 | Kyrgyzstan_IronAge_Saka | J2 | 2080
DA58.damgaard_2018_nature | Kurgan nr. 3, Baskya 1 | Kyrgyzstan_IronAge_Saka | J2a | 2078
DA47.damgaard_2018_nature | Kurgan nr. K9 (individ. n... | Kyrgyzstan_IronAge_Saka | Q1 | 2076
DA54.damgaard_2018_nature | Kurgan nr. K70, Keden | Kyrgyzstan_IronAge_Saka | Q1 | 1682
DA52.damgaard_2018_nature | Kurgan nr. K60, Keden | Kyrgyzstan_IronAge_Saka | R1a | 1532
CGG021499.mccoll_230707_ironage | Ayrmyrlyg | Russia_IronAge_Hun | – | 2230
CGG021500.mccoll_230707_ironage | Ayrmyrlyg | Russia_IronAge_Hun | R1a | 2230
CGG021502.mccoll_230707_ironage | Ayrmyrlyg | Russia_IronAge_Hun | – | 2230
CGG021503.mccoll_230707_ironage | Ayrmyrlyg | Russia_IronAge_Hun | – | 2230
CGG021493.mccoll_230707_ironage | Ayrmyrlyg | Russia_IronAge_Saka | – | 2234
CGG021494.mccoll_230707_ironage | Ayrmyrlyg | Russia_IronAge_Saka | – | 2234
CGG021496.mccoll_230707_ironage | Ayrmyrlyg | Russia_IronAge_Saka | R1a | 2234
CGG021497.mccoll_230707_ironage | Ayrmyrlyg | Russia_IronAge_Saka | – | 2234
Kazakhstan_IronAge_Hun
Kazakhstan_IronAge_Saka
Kyrgyzstan_IronAge_Hun
Kyrgyzstan_IronAge_Saka
Russia_IronAge_Hun
Russia_IronAge_Saka

Allentoft, M.E. *et al.* (2015) 'Population genomics of Bronze Age Eurasia', *Nature*, 522(7555), pp. 167–172. Available at: https://doi.org/10.1038/nature14507.
Allentoft, M.E. *et al.* (2022) 'Population Genomics of Stone Age Eurasia'. Evolutionary Biology. Available at: https://doi.org/10.1101/2022.05.04.490594.
Allentoft, M.E., Sikora, M., Fischer, A., *et al.* (2024) '100 ancient genomes show repeated population turnovers in Neolithic Denmark', *Nature*, 625(7994), pp. 329–337. Available at: https://doi.org/10.1038/s41586-023-06862-3.
Allentoft, M.E., Sikora, M., Refoyo-Martínez, A., *et al.* (2024) 'Population genomics of post-glacial western Eurasia', *Nature*, 625(7994), pp. 301–311. Available at: https://doi.org/10.1038/s41586-023-06865-0.
Amorim, C.E.G. *et al.* (2018) 'Understanding 6th-century barbarian social organization and migration through paleogenomics', *Nature Communications*, 9(1), p. 3547. Available at: https://doi.org/10.1038/s41467-018-06024-4.
Antonio, M.L. *et al.* (2019) 'Ancient Rome: A genetic crossroads of Europe and the Mediterranean', *Science*, 366(6466), pp. 708–714. Available at: https://doi.org/10.1126/science.aay6826.
Auton, A. *et al.* (2015) 'A global reference for human genetic variation', *Nature*, 526(7571), pp. 68–74. Available at: https://doi.org/10.1038/nature15393.
Barrie, W. *et al.* (2024) 'Elevated genetic risk for multiple sclerosis emerged in steppe pastoralist populations', *Nature*, 625(7994), pp. 321–328. Available at: https://doi.org/10.1038/s41586-023-06618-z.
de Barros Damgaard, P. *et al.* (2018) 'The first horse herders and the impact of early Bronze Age steppe expansions into Asia', *Science*, 360(6396), p. eaar7711. Available at: https://doi.org/10.1126/science.aar7711.
Brace, S. *et al.* (2019) 'Ancient genomes indicate population replacement in Early Neolithic Britain', *Nature Ecology & Evolution*, 3(5), pp. 765–771. Available at: https://doi.org/10.1038/s41559-019-0871-9.
Broushaki, F. *et al.* (2016) 'Early Neolithic genomes from the eastern Fertile Crescent', *Science*, 353(6298), pp. 499–503. Available at: https://doi.org/10.1126/science.aaf7943.
Browning, B.L. and Browning, S.R. (2013) 'Detecting Identity by Descent and Estimating Genotype Error Rates in Sequence Data', *The American Journal of Human Genetics*, 93(5), pp. 840–851. Available at: https://doi.org/10.1016/j.ajhg.2013.09.014.
Brunel, S. *et al.* (2020) 'Ancient genomes from present-day France unveil 7,000 years of its demographic history', *Proceedings of the National Academy of Sciences*, 117(23), pp. 12791–12798. Available at: https://doi.org/10.1073/pnas.1918034117.
Cassidy, L.M. *et al.* (2016) 'Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome', *Proceedings of the National Academy of Sciences*, 113(2), pp. 368–373. Available at: https://doi.org/10.1073/pnas.1518445113.
Cassidy, L.M. *et al.* (2020) 'A dynastic elite in monumental Neolithic society', *Nature*, 582(7812), pp. 384–388. Available at: https://doi.org/10.1038/s41586-020-2378-6.
Christiansen, T.T. (2017) *The Productive Limfjord Region in perspective: A Study of Metal Detecting Sites and Socioeconomic Development in Denmark, AD 400-1150*. Aarhus University. Available at: https://nordjyskemuseer.dk/den-fundrige-limfjordsregion-i-perspektiv/ (Accessed: 11 December 2023).
Clemente, F. *et al.* (2021) 'The genomic history of the Aegean palatial civilizations', *Cell*, 184(10), pp. 2565-2586.e21. Available at: https://doi.org/10.1016/j.cell.2021.03.039.
Coutinho, A. *et al.* (2020) 'The Neolithic Pitted Ware culture foragers were culturally but not

genetically influenced by the Battle Axe culture herders', *American Journal of Physical Anthropology*, 172(4), pp. 638–649. Available at: https://doi.org/10.1002/ajpa.24079.
Damgaard, P. de B. *et al.* (2018) '137 ancient human genomes from across the Eurasian steppes', *Nature*, 557(7705), pp. 369–374. Available at: https://doi.org/10.1038/s41586-018-0094-2.
Ebenesersdóttir, S.S. *et al.* (2018) 'Ancient genomes from Iceland reveal the making of a human population', *Science*, 360(6392), pp. 1028–1032. Available at: https://doi.org/10.1126/science.aar2625.
Egfjord, A.F.-H. *et al.* (2021) 'Genomic Steppe ancestry in skeletons from the Neolithic Single Grave Culture in Denmark', *PLOS ONE*, 16(1), p. e0244872. Available at: https://doi.org/10.1371/journal.pone.0244872.
Fernandes, D.M. *et al.* (2018) 'A genomic Neolithic time transect of hunter-farmer admixture in central Poland', *Scientific Reports*, 8(1), p. 14879. Available at: https://doi.org/10.1038/s41598-018-33067-w.
Fernandes, D.M. *et al.* (2020) 'The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean', *Nature Ecology & Evolution*, 4(3), pp. 334–345. Available at: https://doi.org/10.1038/s41559-020-1102-0.
Fischer, C.-E. *et al.* (2022) 'Origin and mobility of Iron Age Gaulish groups in present-day France revealed through archaeogenomics', *iScience*, 25(4). Available at: https://doi.org/10.1016/j.isci.2022.104094.
Fregel, R. *et al.* (2018) 'Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe', *Proceedings of the National Academy of Sciences*, 115(26), pp. 6774–6779. Available at: https://doi.org/10.1073/pnas.1800851115.
Fu, Q. *et al.* (2014) 'Genome sequence of a 45,000-year-old modern human from western Siberia', *Nature*, 514(7523), pp. 445–449. Available at: https://doi.org/10.1038/nature13810.
Fu, Q. *et al.* (2016) 'The genetic history of Ice Age Europe', *Nature*. Available at: https://doi.org/10.1038/nature17993.
Furtwängler, A. *et al.* (2020) 'Comparison of target enrichment strategies for ancient pathogen DNA', *BioTechniques*, 69(6), pp. 455–459. Available at: https://doi.org/10.2144/btn-2020-0100.
Gallego-Llorente, M. *et al.* (2016) 'The genetics of an early Neolithic pastoralist from the Zagros, Iran', *Scientific Reports*, 6(1), p. 31326. Available at: https://doi.org/10.1038/srep31326.
Gamba, C. *et al.* (2014) 'Genome flux and stasis in a five millennium transect of European prehistory', *Nature Communications*, 5(1), p. 5257. Available at: https://doi.org/10.1038/ncomms6257.
González-Fortes, G. *et al.* (2017) 'Paleogenomic Evidence for Multi-generational Mixing between Neolithic Farmers and Mesolithic Hunter-Gatherers in the Lower Danube Basin', *Current Biology*, 27(12), pp. 1801-1810.e10. Available at: https://doi.org/10.1016/j.cub.2017.05.023.
González-Fortes, G. *et al.* (2019) 'A western route of prehistoric human migration from Africa into the Iberian Peninsula', *Proceedings of the Royal Society B: Biological Sciences*, 286(1895), p. 20182288. Available at: https://doi.org/10.1098/rspb.2018.2288.
Greenbaum, G. *et al.* (2019) 'Network-based hierarchical population structure analysis for large genomic data sets', *Genome Research*, 29(12), pp. 2020–2033. Available at: https://doi.org/10.1101/gr.250092.119.
Gretzinger, J. *et al.* (2022) 'The Anglo-Saxon migration and the formation of the early English gene pool', *Nature*, 610(7930), pp. 112–119. Available at: https://doi.org/10.1038/s41586-022-05247-2.
Günther, T. *et al.* (2015) 'Ancient genomes link early farmers from Atapuerca in Spain to

modern-day Basques', *Proceedings of the National Academy of Sciences*, 112(38), pp. 11917–11922. Available at: https://doi.org/10.1073/pnas.1509851112.
Günther, T. *et al.* (2018) 'Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation', *PLOS Biology*, 16(1), p. e2003703. Available at: https://doi.org/10.1371/journal.pbio.2003703.
Haber, M. *et al.* (2017) 'Continuity and Admixture in the Last Five Millennia of Levantine History from Ancient Canaanite and Present-Day Lebanese Genome Sequences', *The American Journal of Human Genetics*, 101(2), pp. 274–282. Available at: https://doi.org/10.1016/j.ajhg.2017.06.013.
Haber, M. *et al.* (2019) 'A Transient Pulse of Genetic Admixture from the Crusaders in the Near East Identified from Ancient Genome Sequences', *The American Journal of Human Genetics*, 104(5), pp. 977–984. Available at: https://doi.org/10.1016/j.ajhg.2019.03.015.
Hansen, H.B. *et al.* (2017) 'Comparing Ancient DNA Preservation in Petrous Bone and Tooth Cementum', *PLoS One*. Available at: https://doi.org/10.1371/journal.pone.0170940.
Hofmanová, Z. *et al.* (2016) 'Early farmers from across Europe directly descended from Neolithic Aegeans', *Proceedings of the National Academy of Sciences*, 113(25), pp. 6886–6891. Available at: https://doi.org/10.1073/pnas.1523951113.
Järve, M. *et al.* (2019) 'Shifts in the Genetic Landscape of the Western Eurasian Steppe Associated with the Beginning and End of the Scythian Dominance', *Current Biology*, 29(14), pp. 2430-2441.e10. Available at: https://doi.org/10.1016/j.cub.2019.06.019.
Jensen, T.Z.T. *et al.* (2019) 'A 5700 year-old human genome and oral microbiome from chewed birch pitch', *Nature Communications*, 10(1), p. 5520. Available at: https://doi.org/10.1038/s41467-019-13549-9.
Jones, E.R. *et al.* (2015) 'Upper Palaeolithic genomes reveal deep roots of modern Eurasians', *Nature Communications*, 6(1), p. 8912. Available at: https://doi.org/10.1038/ncomms9912.
Jones, E.R. *et al.* (2017) 'The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers', *Current Biology*, 27(4), pp. 576–582. Available at: https://doi.org/10.1016/j.cub.2016.12.060.
Kanzawa-Kiriyama, H. *et al.* (2019) 'Late Jomon male and female genome sequences from the Funadomari site in Hokkaido, Japan', *Anthropological Science*, 127(2), pp. 83–108. Available at: https://doi.org/10.1537/ase.190415.
Kashuba, N. *et al.* (2019) 'Ancient DNA from mastics solidifies connection between material culture and genetics of mesolithic hunter–gatherers in Scandinavia', *Communications Biology*, 2(1), pp. 1–10. Available at: https://doi.org/10.1038/s42003-019-0399-1.
Katoh, K. and Standley, D.M. (2013) 'MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability', *Molecular Biology and Evolution*, 30(4), pp. 772–780. Available at: https://doi.org/10.1093/molbev/mst010.
Kılınç, G.M. *et al.* (2016) 'The Demographic Development of the First Farmers in Anatolia', *Current Biology*, 26(19), pp. 2659–2666. Available at: https://doi.org/10.1016/j.cub.2016.07.057.
Kılınç, G.M. *et al.* (2021) 'Human population dynamics and Yersinia pestis in ancient northeast Asia', *Science Advances*, 7(2), p. eabc4587. Available at: https://doi.org/10.1126/sciadv.abc4587.
Kozlov, A.M. *et al.* (2019) 'RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference', *Bioinformatics*, 35(21), pp. 4453–4455. Available at: https://doi.org/10.1093/bioinformatics/btz305.
Krzewińska, M., Kılınç, G.M., *et al.* (2018) 'Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads', *Science Advances*, 4(10), p. eaat4457. Available at: https://doi.org/10.1126/sciadv.aat4457.

Krzewińska, M., Kjellström, A., *et al.* (2018) 'Genomic and Strontium Isotope Variation Reveal Immigration Patterns in a Viking Age Town', *Current Biology*, 28(17), pp. 2730-2738.e10. Available at: https://doi.org/10.1016/j.cub.2018.06.053.
Lazaridis, I. *et al.* (2014) 'Ancient human genomes suggest three ancestral populations for present-day Europeans', *Nature*, 513(7518), pp. 409–413. Available at: https://doi.org/10.1038/nature13673.
Lazaridis, I. *et al.* (2022) 'The genetic history of the Southern Arc: A bridge between West Asia and Europe', *Science*, 377(6609), p. eabm4247. Available at: https://doi.org/10.1126/science.abm4247.
Li, H. (2011) 'A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data', *Bioinformatics*, 27(21), pp. 2987–2993. Available at: https://doi.org/10.1093/bioinformatics/btr509.
Linderholm, A. *et al.* (2020) 'Corded Ware cultural complexity uncovered using genomic and isotopic analysis from south-eastern Poland', *Scientific Reports*, 10(1), p. 6885. Available at: https://doi.org/10.1038/s41598-020-63138-w.
Lindo, J. *et al.* (2018) 'The genetic prehistory of the Andean highlands 7000 years BP though European contact', *Science Advances*, 4(11), p. eaau4921. Available at: https://doi.org/10.1126/sciadv.aau4921.
Lipson, M. and Reich, D. (2017) 'A working model of the deep relationships of diverse modern human genetic lineages outside of Africa', *Molecular Biology and Evolution*, p. msw293. Available at: https://doi.org/10.1093/molbev/msw293.
Llorente, M.G. *et al.* (2015) 'Ancient Ethiopian genome reveals extensive Eurasian admixture in Eastern Africa', *Science*, 350(6262), pp. 820–822. Available at: https://doi.org/10.1126/science.aad2879.
Mallick, S. *et al.* (2023) 'The Allen Ancient DNA Resource (AADR): A curated compendium of ancient human genomes'. bioRxiv, p. 2023.04.06.535797. Available at: https://doi.org/10.1101/2023.04.06.535797.
Malmström, H. *et al.* (2019) 'The genomic ancestry of the Scandinavian Battle Axe Culture people and their relation to the broader Corded Ware horizon', *Proceedings of the Royal Society B: Biological Sciences*, 286(1912), p. 20191528. Available at: https://doi.org/10.1098/rspb.2019.1528.
Marcus, J.H. *et al.* (2020) 'Genetic history from the Middle Neolithic to present on the Mediterranean island of Sardinia', *Nature Communications*, 11(1), p. 939. Available at: https://doi.org/10.1038/s41467-020-14523-6.
Margaryan, A. *et al.* (2020) 'Population genomics of the Viking world', *Nature*, 585(7825), pp. 390–396. Available at: https://doi.org/10.1038/s41586-020-2688-8.
Maróti, Z. *et al.* (2022) 'The genetic origin of Huns, Avars, and conquering Hungarians', *Current Biology*, 32(13), pp. 2858-2870.e7. Available at: https://doi.org/10.1016/j.cub.2022.04.093.
Martiniano, R. *et al.* (2016) 'Genomic signals of migration and continuity in Britain before the Anglo-Saxons', *Nature Communications*, 7(1), p. 10326. Available at: https://doi.org/10.1038/ncomms10326.
Martiniano, R. *et al.* (2017) 'The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods', *PLOS Genetics*, 13(7), p. e1006852. Available at: https://doi.org/10.1371/journal.pgen.1006852.
Mathieson, I. *et al.* (2015) 'Genome-wide patterns of selection in 230 ancient Eurasians', *Nature*, 528(7583), pp. 499–503. Available at: https://doi.org/10.1038/nature16152.
McColl, H. *et al.* (2018) 'The prehistoric peopling of Southeast Asia', *Science*, 361(6397),

pp. 88–92. Available at: https://doi.org/10.1126/science.aat3628.
Mittnik, A. *et al.* (2018) 'The genetic prehistory of the Baltic Sea region', *Nature Communications*, 9(1), p. 442. Available at: https://doi.org/10.1038/s41467-018-02825-9.
Moots, H.M. *et al.* (2023) 'A genetic history of continuity and mobility in the Iron Age central Mediterranean', *Nature Ecology & Evolution*, pp. 1–10. Available at: https://doi.org/10.1038/s41559-023-02143-4.
Moreno-Mayar, J.V., Vinner, L., *et al.* (2018) 'Early human dispersals within the Americas', *Science*, 362(6419), p. eaav2621. Available at: https://doi.org/10.1126/science.aav2621.
Moreno-Mayar, J.V., Potter, B.A., *et al.* (2018) 'Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans', *Nature*, 553(7687), pp. 203–207. Available at: https://doi.org/10.1038/nature25173.
Mota, B.S. da *et al.* (2022) 'Imputation of ancient genomes'. bioRxiv, p. 2022.07.19.500636. Available at: https://doi.org/10.1101/2022.07.19.500636.
Narasimhan, V.M. *et al.* (2019) 'The formation of human populations in South and Central Asia', *Science*, 365(6457), p. eaat7487. Available at: https://doi.org/10.1126/science.aat7487.
Ning, C. *et al.* (2019) 'Ancient Genomes Reveal Yamnaya-Related Ancestry and a Potential Source of Indo-European Speakers in Iron Age Tianshan', *Current Biology*, 29(15), pp. 2526-2532.e4. Available at: https://doi.org/10.1016/j.cub.2019.06.044.
Ning, C. *et al.* (2020) 'Ancient genomes from northern China suggest links between subsistence changes and human migration', *Nature Communications*, 11(1), p. 2700. Available at: https://doi.org/10.1038/s41467-020-16557-2.
Olalde, I. *et al.* (2014) 'Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European', *Nature*, 507(7491), pp. 225–228. Available at: https://doi.org/10.1038/nature12960.
Olalde, I. *et al.* (2015) 'A Common Genetic Origin for Early Farmers from Mediterranean Cardial and Central European LBK Cultures', *Molecular Biology and Evolution*, 32(12), pp. 3132–3142. Available at: https://doi.org/10.1093/molbev/msv181.
Olalde, I. *et al.* (2018) 'The Beaker phenomenon and the genomic transformation of northwest Europe', *Nature*, 555(7695), pp. 190–196. Available at: https://doi.org/10.1038/nature25738.
Olalde, I. *et al.* (2019) 'The genomic history of the Iberian Peninsula over the past 8000 years', *Science*, 363(6432), pp. 1230–1234. Available at: https://doi.org/10.1126/science.aav4040.
Papac, L. *et al.* (2021) 'Dynamic changes in genomic and social structures in third millennium BCE central Europe', *Science Advances*, 7(35), p. eabi6941. Available at: https://doi.org/10.1126/sciadv.abi6941.
Patterson, N. *et al.* (2022) 'Large-scale migration into Britain during the Middle to Late Bronze Age', *Nature*, 601(7894), pp. 588–594. Available at: https://doi.org/10.1038/s41586-021-04287-4.
Patterson, N., Price, A.L. and Reich, D. (2006) 'Population structure and eigenanalysis', *PLoS Genet*. Available at: https://doi.org/10.1371/journal.pgen.0020190.
Posth, C. *et al.* (2021) 'The origin and legacy of the Etruscans through a 2000-year archeogenomic time transect', *Science Advances*, 7(39), p. eabi7673. Available at: https://doi.org/10.1126/sciadv.abi7673.
Price, A.L. *et al.* (2006) 'Principal components analysis corrects for stratification in genome-wide association studies', *Nature Genetics*, 38(8), pp. 904–909. Available at: https://doi.org/10.1038/ng1847.
Raghavan, M. *et al.* (2014) 'The genetic prehistory of the New World Arctic', *Science*, 345(6200), p. 1255832. Available at: https://doi.org/10.1126/science.1255832.
Raghavan, M. *et al.* (2015) 'Genomic evidence for the Pleistocene and recent population

history of Native Americans', *Science*, 349(6250), p. aab3884. Available at: https://doi.org/10.1126/science.aab3884.
Rasmussen, M. *et al.* (2010) 'Ancient human genome sequence of an extinct Palaeo-Eskimo', *Nature*, 463(7282), pp. 757–762. Available at: https://doi.org/10.1038/nature08835.
Rasmussen, M. *et al.* (2014) 'The genome of a Late Pleistocene human from a Clovis burial site in western Montana', *Nature*, 506(7487), pp. 225–229. Available at: https://doi.org/10.1038/nature13025.
Rasmussen, M. *et al.* (2015) 'The ancestry and affiliations of Kennewick Man', *Nature*. Available at: https://doi.org/10.1038/nature14625.
Rodríguez-Varela, R. *et al.* (2017) 'Genomic Analyses of Pre-European Conquest Human Remains from the Canary Islands Reveal Close Affinity to Modern North Africans', *Current biology: CB*, 27(21), pp. 3396-3402.e5. Available at: https://doi.org/10.1016/j.cub.2017.09.059.
Rodríguez-Varela, R. *et al.* (2023) 'The genetic history of Scandinavia from the Roman Iron Age to the present', *Cell*, 186(1), pp. 32-46.e19. Available at: https://doi.org/10.1016/j.cell.2022.11.024.
Saag, Lehti *et al.* (2017) 'Extensive Farming in Estonia Started through a Sex-Biased Migration from the Steppe', *Current Biology*, 27(14), pp. 2185-2193.e6. Available at: https://doi.org/10.1016/j.cub.2017.06.022.
Saag, Lehti *et al.* (2019) 'The Arrival of Siberian Ancestry Connecting the Eastern Baltic to Uralic Speakers further East', *Current Biology*, 29(10), pp. 1701-1711.e16. Available at: https://doi.org/10.1016/j.cub.2019.04.026.
Saag, Lehti *et al.* (2021) 'Genetic ancestry changes in Stone to Bronze Age transition in the East European plain', *Science Advances*, 7(4), p. eabd6535. Available at: https://doi.org/10.1126/sciadv.abd6535.
Sánchez-Quinto, F. *et al.* (2019) 'Megalithic tombs in western and northern Neolithic Europe were linked to a kindred society', *Proceedings of the National Academy of Sciences*, 116(19), pp. 9469–9474. Available at: https://doi.org/10.1073/pnas.1818037116.
Scheib, C.L. *et al.* (2018) 'Ancient human parallel lineages within North America contributed to a coastal expansion', *Science*, 360(6392), pp. 1024–1027. Available at: https://doi.org/10.1126/science.aar6851.
Schiffels, S. *et al.* (2016) 'Iron Age and Anglo-Saxon genomes from East England reveal British migration history', *Nature Communications*, 7(1), p. 10408. Available at: https://doi.org/10.1038/ncomms10408.
Schlebusch, C.M. *et al.* (2017) 'Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago', *Science*, 358(6363), pp. 652–655. Available at: https://doi.org/10.1126/science.aao6266.
Schroeder, H. *et al.* (2019) 'Unraveling ancestry, kinship, and violence in a Late Neolithic mass grave', *Proceedings of the National Academy of Sciences*, 116(22), pp. 10705–10710. Available at: https://doi.org/10.1073/pnas.1820210116.
Seguin-Orlando, A. *et al.* (2014) 'Genomic structure in Europeans dating back at least 36,200 years', *Science*, 346(6213), pp. 1113–1118. Available at: https://doi.org/10.1126/science.aaa0114.
Seguin-Orlando, A. *et al.* (2021) 'Heterogeneous Hunter-Gatherer and Steppe-Related Ancestries in Late Neolithic and Bell Beaker Genomes from Present-Day France', *Current Biology*, 31(5), pp. 1072-1083.e10. Available at: https://doi.org/10.1016/j.cub.2020.12.015.
Sikora, M. *et al.* (2017) 'Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers', *Science*, 358(6363), pp. 659–662. Available at: https://doi.org/10.1126/science.aao1807.
Sikora, M. *et al.* (2019) 'The population history of northeastern Siberia since the

Pleistocene', *Nature*, 570(7760), pp. 182–188. Available at: https://doi.org/10.1038/s41586-019-1279-z.
Skoglund, P. *et al.* (2014) 'Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal', *Proceedings of the National Academy of Sciences*, 111(6), pp. 2229–2234. Available at: https://doi.org/10.1073/pnas.1318934111.
Skoglund, P. *et al.* (2017) 'Reconstructing Prehistoric African Population Structure', *Cell*, 171(1), pp. 59-71.e21. Available at: https://doi.org/10.1016/j.cell.2017.08.049.
Stolarek, I. *et al.* (2023) 'Genetic history of East-Central Europe in the first millennium CE', *Genome Biology*, 24(1), p. 173. Available at: https://doi.org/10.1186/s13059-023-03013-9.
Taliun, D. *et al.* (2021) 'Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program', *Nature*, 590(7845), pp. 290–299. Available at: https://doi.org/10.1038/s41586-021-03205-y.
Unterländer, M. *et al.* (2017) 'Ancestry and demography and descendants of Iron Age nomads of the Eurasian Steppe', *Nature Communications*, 8(1), p. 14615. Available at: https://doi.org/10.1038/ncomms14615.
Valdiosera, C. *et al.* (2018) 'Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia', *Proceedings of the National Academy of Sciences*, 115(13), pp. 3428–3433. Available at: https://doi.org/10.1073/pnas.1717762115.
Veeramah, K.R. *et al.* (2018) 'Population genomic analysis of elongated skulls reveals extensive female-biased immigration in Early Medieval Bavaria', *Proceedings of the National Academy of Sciences*, 115(13), pp. 3494–3499. Available at: https://doi.org/10.1073/pnas.1719880115.
Weissensteiner, H. *et al.* (2016) 'HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing', *Nucleic Acids Res*. Available at: https://doi.org/10.1093/nar/gkw233.

S7. Pre-viking migration into Scandinavia

S7.1. Summary

Ralph Fyfe, Marie-José Gaillard

Large-scale demographic trends are reflected in the history of land use, and informed by data from dendrochonology and palynology. This evidence suggests profound changes across wide parts of Scandinavia between 1600-1200 BP.

Data from dendrochonology shows an abrupt cooling episode after c. 535 CE, a trend associated with the global Late Antique Little Ice Age (LALIA). Across Scandinavia, the impact has been shown to vary regionally (Gundersen, 2022), however studies of tree-rings from Denmark show the most significant decrease of growth in 539, compared to 535 (Ellegård Larsen, 2023), and transitions in land use over this time. By reconstructing the vegetation

through pollen records, we are able to gain insight into human activity during this period. The expectation for a gradual population decline followed by gradual immigration would be a decrease and discontinuity in land use resulting in the recovery of vegetation over an extended period of time, with heterogeneity in the forms of land use occurring over time. In contrast, the expectation for a rapid replacement would be a transition from heterogeneity in the region that has developed over time, to homogeneity of land use of the incoming population.

O'Dwyer et al. (2021) used pollen data from Skåne, Halland, Blekinge and Småland to quantify local vegetation cover for a number of time windows from 7500 BP to present using the Landscape Reconstruction Algorithm (REVEALS and LOVE models, Sugita 2007a and b). The published maps of land cover indicate a clear increase in broadleaf deciduous woodland and a decrease in open land in the time window 1400-1500 BP (450-550 AD) compared to the previous and following time windows, 1900-2000 BP (50 BC-50 AD) and 1000-1100 BP (850-950 AD), respectively. The changes are most pronounced in southernmost Scania, the western coast of Scania and Halland, the eastern coast of Scania, and Blekinge. In the time span 850-950 AD, it is mainly southernmost Scania that regains its large landscape openness, while northern Småland becomes for the first time as deforested as southern Scania. Deforestation of Småland at that time is explained by the development of iron production (e.g. Lindström, 2022). For the purpose of the present study, we reexamined the pollen data from Skåne (S7.3, Table S7.3.1 and Figures S7.3.1 and 7.3.2) looking at the pollen percentage changes of broadleaved trees (as a group of taxa) and four indicators of deforestation and agriculture. It shows a clear decrease in agriculture in lowland Scania (southern and northern hummocky regions) ca. 550-650 AD while the more marginal sites at higher locations in the inland (Bökesjön), in NW Scania, and at the northwestern coast show a decrease in grazing land already from 150-250 AD and very little or no cultivation of cereals all through the studied period until 800 AD. In lowland Scania, agriculture increased again from ca. 650 AD (southern hummocky area; e.g. Gaillard et al., 1991) or slightly later, and the same is valid for grazing in the marginal areas. It is clear that land-use both decreased and underwent reorganisation from the time of the large "dynasty shift" for Scania, from being an independent political community since 400 AD to belonging to the realm of the Danes from 540 AD (Lindström, 2022). From 650 AD lowland areas were re-organized or reoccupied, and marginal areas were settled for the first time ca. 800 AD even though they were used for grazing earlier. The study by Vinogradova et al. (2024) using 5 pollen records and the Landscape Reconstruction Algorithm to quantify land-cover changes along the western coast of Småland N of Kalmar also indicates a decrease in landscape

openness and a reorganisation of land use during the migration period. Without estimates of population change, it is assumed that the pollen records are primarily showing a major change in the land-use strategies and land management, as well as in the location of settlements, rather than changes in the intensity of human impact on the landscape, in accordance with assumptions made earlier by e.g. Berglund et al. (1991).

S7.2. The impact of the volcanic double event in AD 536 and AD 539/540 on tree-ring growth

Hanne Larsen, Morten Fischer Mortensen

Dendrochronology provides a year-to-year record of growth conditions and hence, constitutes a unique method to reveal past climate variation. The method's ability to capture the speed and regional impact of these changes makes it a powerful tool for understanding environmental shifts over time.

We examined dendrochronological data consisting of tree-ring measurements from 654 wood samples of *Quercus* sp. from 42 archaeological sites in Denmark covering the late Iron Age. Most sites are located in Jutland. The samples collected from various types of wooden constructions mainly originated from wells (54%) and bridge constructions (26%) but also pole bars from the seabed, timber, house constructions, canals and caissons. The dendrochronological data is believed to cover most of all wooden material from Denmark between AD 300 and AD 800. This time interval of 500 years (Figure 1) was chosen to emphasize that no such remarkable decrease in tree-ring formation seemed to lie within the normal range of growth variation over an extended timeframe. The tree-ring measurements were combined into an average growth curve for Denmark by use of TSAP-Win^{TM} Rinntech (version 4.82b2)

The average growth curve of the 654 samples of *Quercus* sp. showed a marked decrease in growth from AD 535 to AD 536 with a total reduction of 33% (Figure 2), which was also expressed by very narrow growth rings on the wood samples (Figure 3). Apart from a growth increment in AD 537, the growth continued to decrease and reached a minimum in AD 539 with a total growth reduction of 53% compared with AD 535.

The average growth curve indicates that the climate changes towards much colder conditions. Climate modelling from Southern Norway indicates a temperature reduction up to 3.5 C^{o} degrees during the mid-sixth century (van Dijk et al. 2023). The marked reduction in tree-ring growth in the years around AD 536 and AD 539/540 is known from all over the Northern Hemisphere and are associated with major volcanic eruption followed by a global volcanic dust veil that reduced the solar radiation and lowered the global temperature (Gundersen 2021)

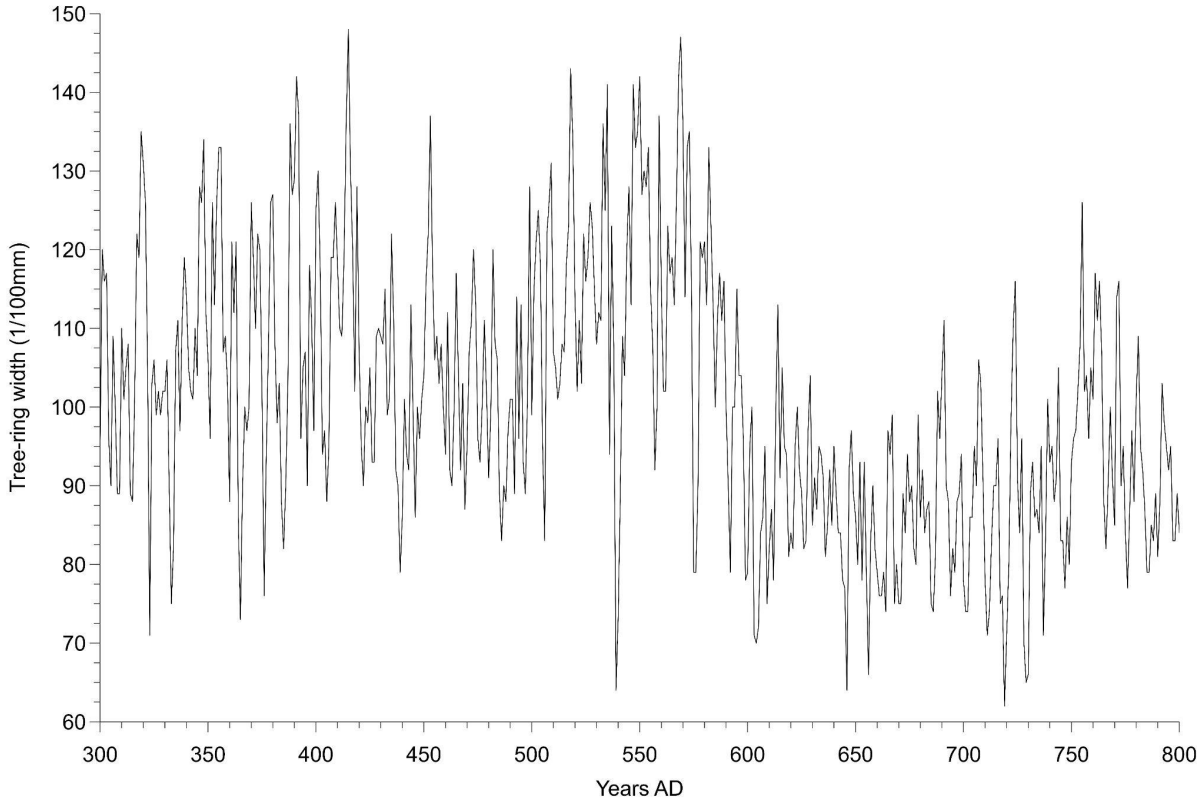


Figure 1. Average growth curve of *Quercus* sp. in Denmark from 300 - 800 CE (1650 - 1150 BP)

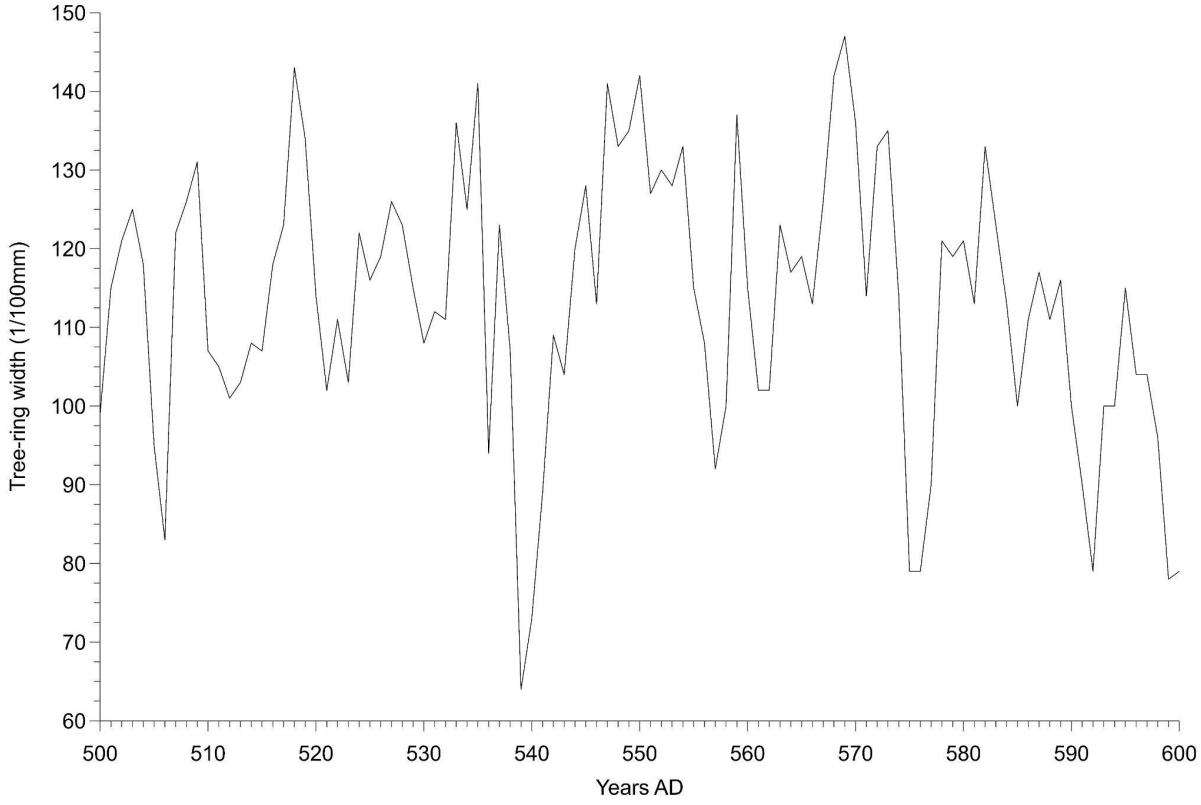
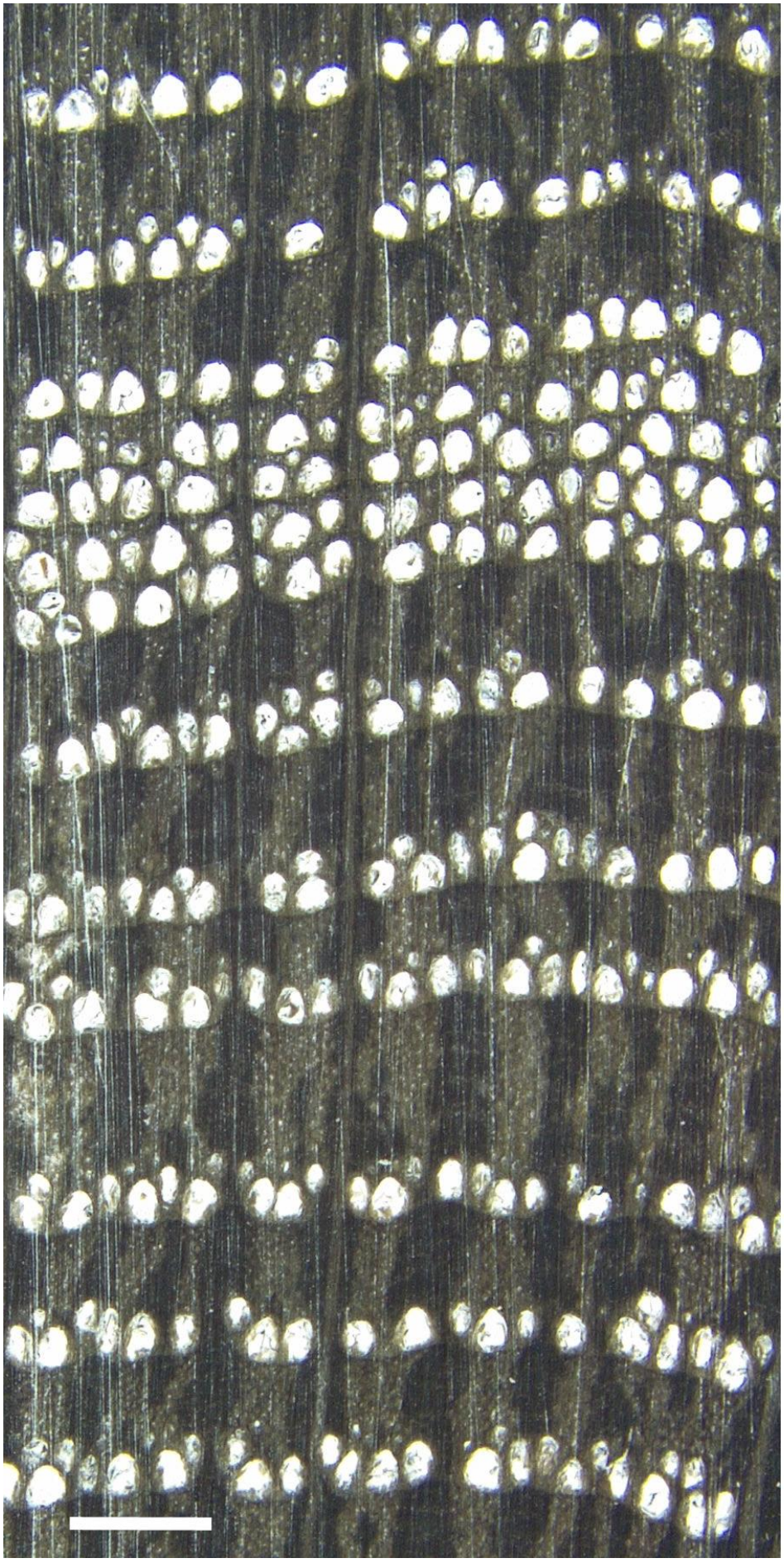


Figure 2. Average growth curve of *Quercus* sp. in Denmark from 500-600 CE (1450-1350 BP)



2080

Figure 3. The impact of the climate changes is shown by the formation of narrow growth rings between 539-542 CE (1411-1408 BP)

S7.3 Vegetation change 4000-800 cal BP in Scania, southern Sweden

Ralph Fyfe and Marie-José Gaillard

Datasets and analytical approach

Twenty sequences were drawn from the LANDCLIMII project data archive for Scania (Githumbi et al 2022; Table S7.3.1), which was compiled using sequences available within the European Pollen Database (Fyfe et al 2009; Giesecke et al 2014) and provided by individual data contributors. The sequences were grouped into four geographical regions: southern hummocky landscapes, northern hummocky landscapes, and upland sites (i.e., marginal areas at slightly higher altitudes on horst ridges or Precambrian bedrock). In this way we can identify major changes for three distinctive areas of Scania in terms of geology and soils, akaline soils for the S hummocky area, slightly acidic soils or acidic soils for the N hummocky area, and acidic soils on eskers and NW Scania. Sequences were prepared for analysis by standardizing the pollen nomenclature to a set of 87 of the most common taxa (those that were present in each sequence at at least 2% total land pollen), and samples dated to 4000 – 800 cal BP (2050 cal BC – cal AD 1150) extracted using the chronologies established in the LANDCLIMII project. A simplified indicator-species approach was used to explore intensity of land-use practice over time, by summarising the proportion of pollen types directly associated with cereal production (Cerealia types, *Secale t.* (rye)), open ground indicators including grasses (Poaceae) and heather (*Calluna vulgaris*) and deciduous trees (the sum of *Corylus*, *Betula*, *Quercus*, *Tilia* and *Ulmus*). In some regions of southern Sweden heather can reflect the development of cultural heathlands used for grazing. Pollen proportions were summarised in 100-year time intervals between 4000-800 cal BP.

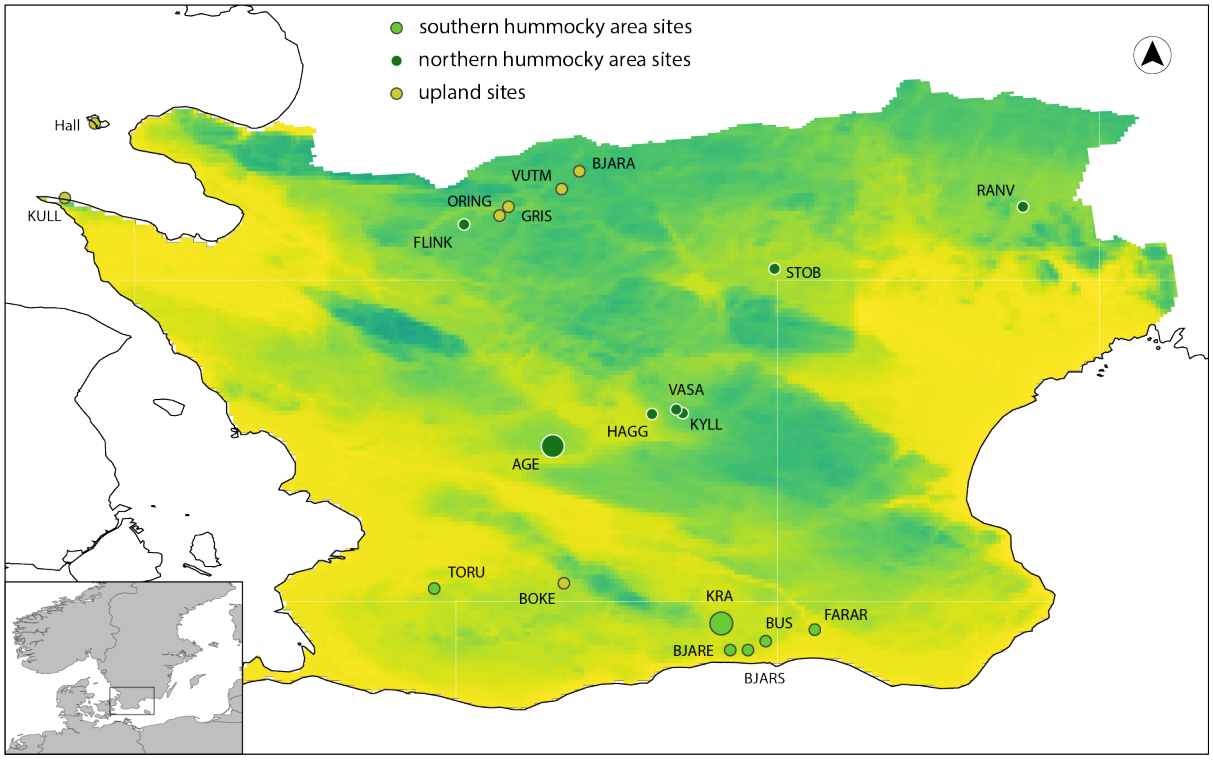


Figure S7.3.1: location of pollen sequences used to assess intensity and nature of land use in southernmost Sweden, 4000-800 BP. Larger circles indicate sites with a more regional character; small circles indicate local pollen sequences. Background colour indicates general pattern of elevation. Details of sites are on Table S7.3.1.

Table S7.3.1: Details of pollen sequences from Scania, taken from Githumbi et al (2022)

Site name	Code	Grouping	References
Bjäresjösjön	BJARE	South hummocky	Gaillard, M.-J. and Berglund, B. E. 1988; Gaillard, M.-J. and Göransson, H. 1991; Gaillard, M.-J. et al. 1991a, b
Bjärsjöholmssjön	BJARS	South hummocky	Göransson, H. 1991; Gaillard, M.-J. and Göransson, H. 1991
Bussjösjön	BUS	South hummocky	Regnéll, J. 1989

Torup	TORU	South hummocky	Hultberg, T. et al. 2010
Fararpsmosse	FARAR	South hummocky	Berglund, B. E. et al., 1991
Krageholmssjön	KRA	South hummocky	Gaillard, M.-J. 1984a, b, c
Flinkasjön	FLINK	North hummocky	Björkman, L. 2004
Fulltofta/Häggenäs	HAGG	North hummocky	Lindbladh, M. et al 2007, Lindbladh, M. and Foster, D. 2010
Fulltofta/Kyllingahus	KYLL	North hummocky	Lindbladh, M. et al 2007, Lindbladh, M. and Foster, D. 2010
Fulltofta/Vasahus	VASA	North hummocky	Lindbladh, M. et al 2007, Lindbladh, M. and Foster, D. 2010
Ranviken	RANV	North hummocky	Digerfeldt, G. 1973
Stoby	STOB	North hummocky	Lagerås, P. 2002
Ageröds Mosse	AGE	North hummocky	Nilsson, T. 1964
Kullaberg	KULL	Upland	Björkman, L. 2001
Hälledammen	Hall	Upland	Molinari C. 2002; Lindbladh, M. and Foster, D. 2010
Bjärabygget	BJARA	Upland	Lagerås, P. 2007
Grisavad	GRIS	Upland	Lagerås, P. 2007

Östra Ringarp	ORING	Upland	Lagerås, P. 2007
Värsjö Utmark	VUTM	Upland	Lagerås, P. 2007
Bökesjön	BOKE	Upland	Gaillard, M.-J. unpublished

Results and interpretation

The compilation of datasets from Scania shown in Figure 7.3.1 shows the impact of agricultural groups since the early Bronze Age. Taking the region as a whole (Figure 7.3.2A), landscape openness across the region increases around 3000 BP (reflected in both the Poaceae and *Calluna vulgaris* values, with a second increase (Poaceae and cereals) at 2600 BP indicating further opening of the landscape and increase in agriculture. Although trends in landscape openness can be established using percentage pollen data, the degree of openness is likely strongly under-estimated in these data. Quantified vegetation openness, using the pollen record from Krageholmssjön (Fig. 7.3.1) and the REVEALS model (Sugita, 2007a), confirms the broad trends in changes, and indicates that tree cover drops from 75% to 35% in Scania between 2800 and 2400 BP (Sugita et al., 2008; Gaillard et al., 2010; O'Dwyer et al. 2021). The grouping of sites into upland, northern and southern hummocky areas demonstrates regional variability in land-cover openness. Sites to the south (Fig. 7..2C) show earlier increases in openness (around 3000 cal BP), and higher levels of openness, reflecting the earlier increase in agriculture with very large areas used for grazing (Berglund 1991). Opening of the landscape in the northern hummocky landscape does not occur until several hundred years later (Fig. 7.3.2B). Cereal pollen types are recorded from as early as 4000 cal BP in the southern hummocky landscape; quantified regional vegetation estimates suggest that just under 1% of the landscape was under cereals between 2700-2200 BP (Githumbi et al. 2022). Evidence for cereals is scattered in the northern hummocky region, and very limited or absent from the uplands.

The period between 1600-1200 BP (400-800 CE) is of particular interest for this study, and the evidence from the pollen record can be used to assess the changes in extent of forest recovery, and cultivation, across southern Sweden through this period. Figure 7.3.2 demonstrates sub-regional trends within Scania, based on pollen proportion data. Cover of broadleaf trees increases in marginal landscapes at 1600 BP (Fig. S7.3.2D) and pollen proportions remain

higher until 1300 BP. There is no evidence of cultivation throughout the period in the uplands. In the northern hummocky area broadleaf tree pollen proportions initially increase (from, on average, 52 to 58% total land pollen); however, this woodland recovery is short lived, with tree pollen proportions dropping again in the interval 1500-1400 BP. There is evidence for continued cereal cultivation during the period, but *Secale* is cultivated first from ca. 1200 BP. In the southern hummocky region tree pollen proportions declined across the period 1600-1200 BP, with no evidence of woodland regeneration. The southern hummocky sites also show a gradual increase in intensity of land use between 1600-1200 BP, characterized by increased representation of cereals (Cerealia type and *Secale*).

The sub-regional picture that emerges from the careful examination of the pollen sequences is thus one of decrease in grazing of marginal land at 1600 cal BP, as reflected in the upland sites, decrease in land use (both grazing and agriculture) in the northern and southern hummocky areas, however rather a stagnation in the South ca. 1600-1400 BP with no regeneration of broadleaved trees. Recovery in the northern hummocky regions is short-lived, with apparent reoccupation seen after around 100 years (by 1500 BP). Cultivation of cereals at the upland sites does not occur until after 1300-1200 BP.

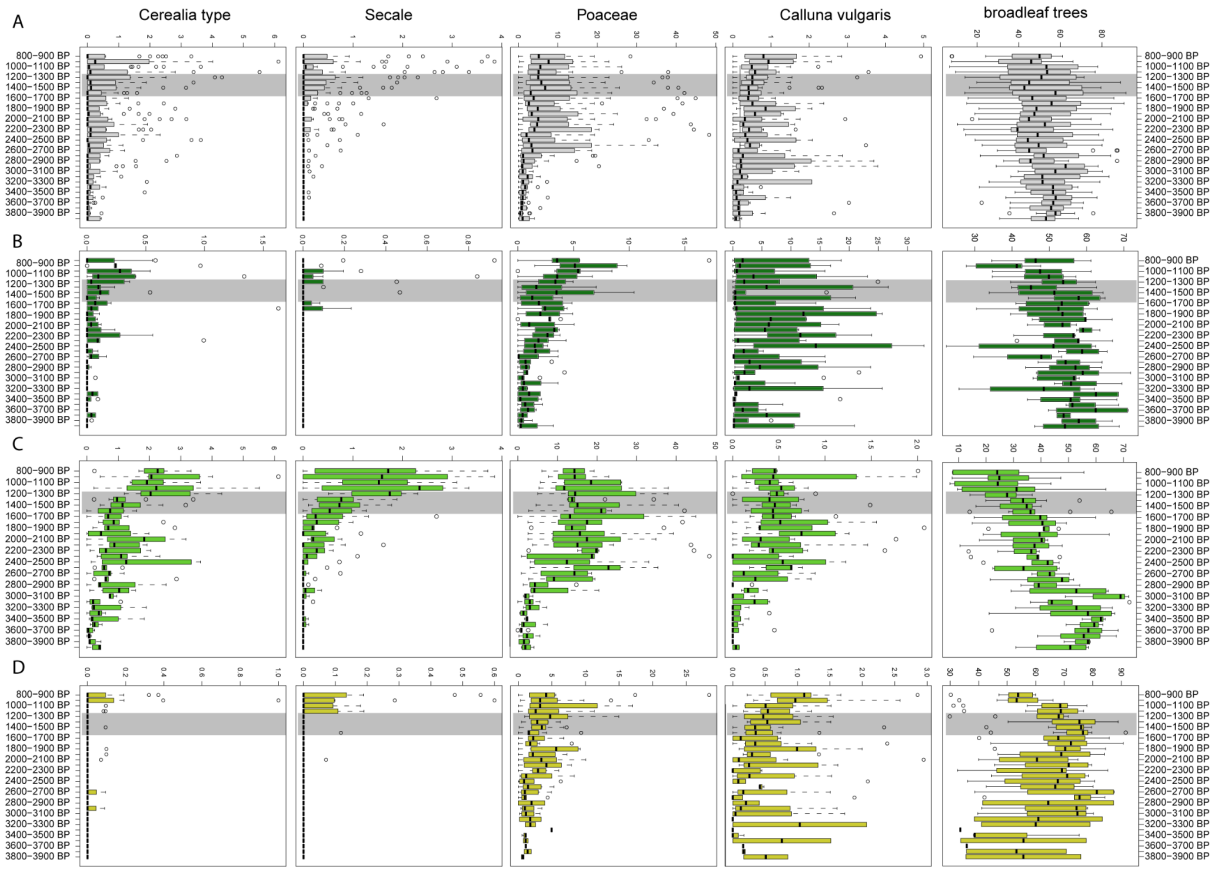


Figure S7.3.2: Summary of key pollen indicator types, southernmost Sweden. Pollen data have been aggregated into 100-yr time intervals. Broadleaf trees include *Corlyus*, *Betula*, *Quercus*, *Tilia*, and *Ulmus*. Panel A: all sites (n=20); Panel B: northern hummocky region sites (7); Panel C: southern hummocky region sites (6); Panel D: upland sites (7). Locations of sites are shown on Figure S7.3.1.

References

Berglund BE (ed.) (1991) *The cultural landscape during 6000 years in southern Sweden: the Ystad project*. Copenhagen, Denmark: Munksgaard

Berglund BE, Hjelmroos M and Kolstrup E (1991) The Köpinge area. Vegetation and landscape through time, in Berglund BE (ed.) *The cultural landscape during 6000 years in southern Sweden: the Ystad project*, 109-112. Copenhagen, Denmark: Munksgaard

Berglund BE, Larsson L, Lewan N, Olsson GA and Stensjö S (1991) Ecological and Social factors behind the landscape changes. Cultural landscape - patterns and processes. In Berglund BE (ed.) *The cultural landscape during 6000 years in southern Sweden: the Ystad project*, 167-174. Copenhagen, Denmark: Munksgaard

Björkman L (2001) The role of human disturbance in Late Holocene vegetation changes on Kullaberg, southern Sweden. *Vegetation History and Archaeobotany* 10, 201–210

Björkman L (2004) Detaljerad pollenanalytisk undersökning av tre lagerföljder (Flinkasjön, E4:13, E4:16) från Örkelljungatrakten i nordvästra Skåne. *LUNDQUA Uppdrag* 54, 1–20

Digerfeldt G (1973) The Postglacial development of the Ranviken bay in Lake Immeln. *Geologiska Föreningen Stockholm Förhandlingar* 96, 3-32.

Ellegård Larsen HM, Baittinger C, Bonde N, Jensen JO, Søvsø M, Ulriksen J, Mortensen MF. Under review. The impact of the volcanic double event in AD 536 and AD 539/540 on tree-ring growth and felling activity in Danish oak trees. *Journal of Archaeological Science Reports*.

Fyfe RM, de Beaulieu JL, Binney H, Bradshaw RHW, Brewer S, Le Flao A, Finsinger W, Gaillard MJ, Giesecke T, Gil-Romera G, Grimm EC, Huntley B, Kunes P, Kuhl N, Leydet M, Lotter AF, Tarasov PE and Tonkov S (2009) The European Pollen Database: past efforts and current activities. *Vegetation History and Archaeobotany* 18, 417-424

Gaillard MJ (1984a) Water-level changes, climate and human impact: a palaeohydrological study of Krageholmssjön (Scania, southern Sweden). In Mörner NA and Karlen W (eds.) *Climate changes on a yearly to millennial basis: geological, historical, and instrumental records*, 147-154. Dordrecht, Reidel Publishing Company

Gaillard MJ (1984b) A palaeohydrological study of Krageholmssjön (Scania, South Sweden): Regional vegetation history and water-level changes. *LUNDQUA Report* 25, Lund University, ISSN 0281-3076

Gaillard MJ and Berglund BE (1988) Land-use history during the last 2700 years in the area of Bjäresjö, South Sweden. In Birks HH, Birks HJB, Kaland PE and Moe D (eds) *The cultural landscape- past, present and future*, 409-428. Cambridge University Press, Cambridge, UK

Gaillard MJ, Olausson D and Skansjö S (1991a) The Bjäresjö area. Conclusions. In Berglund BE (ed.) *The cultural landscape during 6000 years in southern Sweden: the Ystad project*, 167-174. Copenhagen, Denmark: Munksgaard

Gaillard MJ, Dearing JA, El-Daoushy F, Enell M and Håkansson H (1991b) A multidisciplinary study of lake Bjäresjö (S Sweden): land-use history, soil erosion, lake trophy and lake-level fluctuations during the last 3000 years. *Hydrobiologia* 214, 107–114

Gaillard MJ, Dearing JA, El-Daoushy F, Enell M and Håkansson H (1991c) A late Holocene record of land-use history, soil erosion, lake trophy and lake-level fluctuations at Bjäresjösjön (South Sweden). *Journal of Palaeolimnology* 6, 51-81

Gaillard MJ, Sugita S, Mazier F, Trondman A, Broström A *et al*. (2010). Holocene land-cover reconstructions for studies on land cover-climate feedbacks. *Climate of the Past* 6. 483-499

Giesecke T, Davis B, Brewer B, Finsinger W, Wolters S, Blaauw M, de Beaulieu J-L, Fyfe RM, Gaillard M-J, Gil-Romera G, van der Knaap WO, Kuneš P, Kühl N, van Leeuwen JFN, Leydet M, Lotter AF, Semmler M, and Bradshaw RHW (2014) Towards mapping the late Quaternary vegetation change of Europe. *Vegetation History and Archaeobotany* 23, 75-86

Githumbi E, Fyfe RM, Gaillard MJ, Trondman AK, Mazier F, Nielsen AB, Poska A, Sugita S, Woodbridge J, Azuara J, Feurdean A, Grindean R, Lebreton V, Marquer L, Nebout-Combourieu N, Stančikaitė M, Tanţău I, Tonkov S, Shumilovskikh L and LandclimII data contributors (2022) European quantitative pollen-based land-cover reconstructions for the

2218 Holocene: methodology, mapping and potentials. *Earth System Science Data* 14, 1581-1619

2219 Göransson H (1991) Vegetation and man around Lake Bjärsjöholmssjön during prehistoric
2220 time. *LUNDQUA Report* 31, Lund University, 44 pp

2221 Gundersen IM. 2021. *Iron age vulnerability. The Fimbulwinter hypothesis and the archaeology*
2222 *of the inlands of eastern Norway*. Dissertation for the degree of Philosophae Doctor (PhD).
2223 Department of Archaeology, Conservation and History, Faculty of Humanities and Museum
2224 Cultural History, University of Oslo.
2225

2226 Hultberg T, Brunet J., Broström A, and Lindbladh M (2010) Forest in a cultural landscape –
2227 the vegetation history of Torup in southernmost Sweden. *Ecological Bulletins* 53, 141-153

2228 Lagerås P (2002) Skog, slåtter och stenröjning. Paleoekologiska undersökningar i trakten av
2229 Stoby i norra Skåne. In Carlie A (ed.) Skånska regioner. Riksantikvarieämbetet, Avdelningen
2230 för arkeologiska undersökningar. Skrifter No 40

2231 Lagerås P (2007) The ecology of expansion and abandonment.National Heritage Board,
2232 Sweden. Oxbow Books, Oxford, UK, 256 pp

2233 Lindbladh M and Foster D (2010) Dynamics of long-lived foundation species: the history of
2234 Quercus in southern Scandinavia. *Journal of Ecology* 98, 1330–1345

2235 Lindbladh M, Brunet J, Hannon G, Niklasson M, Eliasson P, Eriksson G and Ekstrand A (2007)
2236 Forest history as a basis for ecosystem restoration – a multi-disciplinary case-study in a south
2237 Swedish temperate landscape. *Restoration Ecology* 15, 284–295

2238 Lindström J (2022) *Sveriges långa historia*. Jonathan Lindström and Norstedts, Stockholm
2239

2240 Molinari C (2002) 2500 years of forest dynamics at So¨ndre Skog, a semi-natu[1]ral forest on
2241 Hallands Väderö island, Southern Sweden. MS Thesis, Turin University, Italy.

2242 Nilsson T (1964) Standardpollendiagramme und C14 Datierungen aus dem Agerödsmosse in
2243 mittleren Schonen. *Lunds Universitet Årskrift* 2, 59pp

2244 O'Dwyer, R, Marquer, L, Trondman, A-K, Jönsson, AM (2021) Spatially continuous land-
2245 cover reconstructions through the Holocene in southern Sweden. *Ecosystems* 24, 1450-1467

2246 Regnéll J (1989) Vegetation and land use during 6000 years: palaeocology of the cultural

landscape at two lake sites in southern Skåne, Sweden. *LUNDQUA Thesis* 27, 62pp. Lund University

Sugita S (2007a) Theory of quantitative reconstruction of vegetation I. Pollen from large lakes reveals regional vegetation. *The Holocene* 17, 229–41.

Sugita S (2007b) Theory of quantitative reconstruction of vegetation II: All you need is LOVE. *The Holocene* 17, 243–57.

Sugita S, Gaillard MJ, Hellman S and Brostrom A (2008) Model-based reconstruction of vegetation and landscape using fossil pollen. In Habelt R (ed.) *Layers of Perception, Kolloquien zur Vor- und Frühgeschichte* 10, 385–391. Bonn, Germany

van Dijk E, Gundersen IM, de Bode A, Høeg H, Loftsgarden K, Iversen F, Timmereck C, Jungclaus J, Krüger K. 2023. Climate and society impacts in Scandinavia following the 536/540 CE volcanic double event. *Climate of the Past*. 19:357-398. https://doi.org/10.5194/cp-19-357-2023

Vinogradova O, Gaillard, MJ, Andrén E, Palm V, Rönnby J, Dahl M, Almgren E, Karlsson J, Nielsen AB, Akesson, C and Andrén T (2024). 3000 Years of past regional and local land-use and land-cover change in the southeastern Swedish coastal area: Early human-induced increases in landscape openness as a potential nutrient source to the Baltic Sea coastal waters. *The Holocene* 34, 56-73.