Steppe Ancestry in Western Eurasia and the Spread of the Germanic Languages

Hugh McColl[1,2]*, Guus Kroonen[3,4]*, J. Víctor Moreno-Mayar[1], Frederik Valeur Seersholm[1], Gabriele Scorrano[1,5], Thomaz Pinotti[1,6], Tharsika Vimala[1], Søren M. Sindbæk[7], Per Ethelberg[8], Ralph Fyfe[9], Marie-José Gaillard[10], Hanne M. Ellegård Larsen[11], Morten Fischer Mortensen[11], Fabrice Demeter[1,12], Marie Louise S. Jørkov[13], Sophie Bergerbrant[14], Peter de Barros Damgaard[1], Morten E. Allentoft[15,1], Lasse Vinner[1], Charleen Gaunitz[1], Abigail Ramsøe[1], Isin Altinkaya[1], Rasmus Amund Henriksen[1], Evan K. Irving-Pease[1,16], KG Sjögren[2], Serena Sabatini[2], Anders Fischer[17,18], William Barrie[19], Andrés Ingason[20,1], Anders Rosengren[20,1], Andrew Vaughn[21], Jialu Cao[1], Jacqueline Staring[1], Jesper Stenderup[1], Fulya Eylem Yediay[1,18], Torbjörn Ahlström[22], Irakli Akhvlediani[23], Sofie Laurine Albris[24], Biyaslan Atabiev[25], E.V. Balanovskaya[26], Pernille Bangsgaard[1], Maria Giovanna Belcastro[27], Nick Card[28], Philippe Charlier[29,30], Elizaveta Chernykh[31], Torben Trier Christiansen[32], Alfredo Coppa[33], Maura De Coster[34], Sean Dexter Denham[35], Sophie Desenne[36,37], Jane Downes[28], Karin Margarita Frei[38], Olivér Gábor[39], U.B. Gadiev[40], Johan Zakarias Gårdsvoll[24], Zanette Tsigaridas Glørstad[41], Jesper Hansen[42], Stijn Heeren[43], Merete Henriksen[44], Volker Heyd[45], Mette Høj[46], Mads Kähler Holst[47], Rimantas Jankauskas[48], Henrik Janson[2], Mads Dengsø Jessen[24], Jens Winther Johannsen[49], Torkel Johansen[50], Ole Thirup Kastholm[49], Anton Kern†[51], Ruslan Khaskhanov[52], Katrine Ipsen Kjær[49], Vladimir Kolosov[53], Lisette M. Kootker[34], Klaudia Kyselicová[54], Anne Christine Larsen[55], Thierry Lejars[56], Mette Løvschal[57,47], Niels Lynnerup[58], Yvonne Magnusson[59], V. Yu. Malashev[60], Kristiina Mannermaa[45], Vyacheslav Masyakin[61], Anne Lene Melheim[41], Inga Merkyte[62], Vyacheslav

24   Moiseyev[63], Stig Bergmann Møller[64], Erika Molnár[65], Nadja Mortensen[66], Eileen Murphy[67],

25   Bjarne Henning Nielsen[68], Doris Pany-Kucera[51], Bettina Schulz Paulsson[2], Gertjan Plets[69],

26   Marcia S Ponce de León[70], Håkon Reiersen[71], Walter Reinhard[72], Antti Sajantila[73,74], Birgitte

27   Skar[44], Vladimir Slavchev[75], Václav Smrčka[76], Lasse Sørensen[77], Georg Tiefengraber[51], Otto

28   Christian Uldum[78], Helle Vandkilde[57], Jorge Vega[79], Daniele Vitali[80], Alexey Voloshinov[81],

29   Sidsel Wåhlin[68,82], Holger Wendling[83,84], Anna Wessman[85], Helene Wilhelmson[86,87], Karin

30   Wiltschke[51], João Zilhão[88], Christoph PE Zollikofer[70], Thorfinn Sand Korneliussen[1], Bruno

31   Chaume[89], Jean-Paul Demoule[90,91], Thomas Werge[20,1,92], Line Olsen[1], Rasmus Nielsen[93,1],

32   Lotte Hedeager[1], Kristian Kristiansen[1,18]* Martin Sikora[1], Eske Willerslev[1,19,94]*

33

34

35   [1]Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen,

36   Copenhagen, Denmark, [2]Department of Historical Studies, University of Gothenburg,

37   Gothenburg, Sweden, [3]Leiden University Centre for Linguistics, Leiden University, Leiden,

38   The Netherlands, [4]Department of Nordic Studies and Linguistics, University of Copenhagen,

39   Copenhagen, Denmark, [5]Center for Molecular Anthropology for the Study of Ancient DNA,

40   Department of Biology, University of Rome Tor Vergata, Rome, Italy, [6]Laboratório de

41   Biodiversidade e Evolução Molecular (LBEM), Universidade Federal de Minas Gerais, Belo

42   Horizonte, Brazil, [7]Centre for Urban Network Evolutions (UrbNet), School of Culture and

43   Society, Aarhus University, Aarhus, Denmark, [8]Museum Sønderjylland, Arkæologi,

44   Denmark, [9]School of Geography, Earth and Environmental Sciences, University of Plymouth

45   UK, [10]Department of Biology and Environmental Sciences, Linnaeus University, Kalmar,

46   Sweden, [11]National Museum of Denmark. Environmental Archaeology and Materials

47   Science, Denmark, [12]Eco-anthropologie (EA), Dpt ABBA, Muséum national d'Histoire

48   naturelle, CNRS, Université Paris Cité, Musée de l'Homme 17 place du Trocadéro, 75016

49   Paris, France, [13]Laboratory of Biological Anthropology, Globe Institute, University of

50   Copenhagen, Denmark, [14]Department of Archaeology and Ancient History, Campus Gotland,

51   Uppsala University, Sweden, [15]Trace and Environmental DNA (TrEnD) Laboratory, School

52   of Molecular and Life Sciences, Curtin University, Perth, Australia, [16]Section for Molecular

53   Ecology and Evolution, Globe Institute, University of Copenhagen, Copenhagen, Denmark,

54   [17]Sealand Archaeology, Kalundborg, Denmark, [18]Department of Historical Studies,

55   University of Gothenburg, Sweden, [19]Department of Genetics, University of Cambridge,

56   Cambridge, UK, [20]Institute of Biological Psychiatry, Mental Health Services, Copenhagen

57   University Hospital, Roskilde, Denmark, [21]Center for Computational Biology, University of

58   California Berkeley, Berkeley, California, United States of America, [22]Department of

59   Archaeology and Ancient History, Lund University, [23]Georgian and Caucasus YDNA,

60   Family Tree DNA, Houston, Texas, USA, [24]National Museum of Denmark, Prehistory,

61   Middle Ages and Renaissance, [25]Institute for Caucasus Archaeology, Russia, [26]Medical

62   Genetic Research Center, Moscow, Russia, [27]Department of Biological, Geological and

63   Environmental Sciences - BiGeA, University of Bologna, Via Selmi 3 40126 Bologna, Italy,

64   [28]Archaeology Institute, University of Highlands and Islands, UK, [29]Laboratory

65   Anthropology, Archaeology, Biology (LAAB), UFR of Health Sciences (UVSQ / Paris

66   Saclay University), 2 avenue de la source de la Bièvre, 78180 Montigny-Le-Bretonneux,

67   France, [30]Department of Epidemiology and Public Health, University Hospital Raymond

68   Poincaré (APHP), Garches, France, [31]Udmurt State University, [32]The Historical Museum of

69   Northern Jutland, [33]Department of Environmental Biology, Sapienza University of Rome,

70   [34]Geology & Geochemistry cluster, Department of Earth Sciences, Vrije Universiteit

71   Amsterdam, The Netherlands, [35]Museum of Archaeology, University of Stavanger, [36]Institut

72   national de recherches archéologiques préventives, [37]UMR 8215, CNRS, [38]National Museum

73   of Denmark, [39]Janus Pannonius Muzeum,Pécs, Hungary, [40]E.I. Krupnov Archaeological

74   Center, Ingushetia, Russia, [41]Museum of Cultural History, University of Oslo, [42]Svendborg

75   Museum, Denmark, [43]Vrije Universiteit Amsterdam, the Netherlands, [44]Institute of

76   archaeology and cultural history, NTNU University Museum, Norway., [45]Department of

77   Cultures / Archaeology, University of Helsinki, Finland, [46]The Museum Organisation

78   ROMU, Collections, [47]Moesgaard Museum, [48]Vilnius University, Lithuania, [49]The Museum

79   Organisation ROMU, Archaeology, [50]NTNU University Museum, Department of

80   Archaeology and Cultural History, [51]Natural History Museum Vienna, Austria, [52]Complex

81   Research Institute of the Russian Academy of Sciences, Grozny, Russia, [53]The State

82   Hermitage Museum, Russia, [54]Faculty of Medicine, Institute of Physiology, Comenius

83   University Bratislava, Slovakia, [55]National Museum of Denmark, Trelleborg, Denmark,

84   [56]CNRS, Umr 8546, ENS, PSL, Paris, France, [57]Department of Archaeology and Heritage

85   Studies, School of Culture and Society, Aarhus University, Denmark, [58]Department of

86   Forensic Medicine, University of Copenhagen, Denmark, [59]Malmö museum, 205 80 Malmö,

87   Sweden/Regionmuseet Skåne, Box 134, 291 22 Kristianstad, [60]Institute of Archeology of the

88   Russian Academy of Sciences, Moscow, Russia, [61]National Academy of Sciences, Ukraine,

89   [62]Saxo-Instituttet, Copenhagen, [63]Peter the Great Museum of Anthropology and Ethnography

90   (Kunstkamera), Russian Academy of Sciences, Universitetskaya Nab. 3, St. Petersburg,

91   Russia, [64]Nordjyske Museer, Aalborg, Denmark, [65]Department of Biological Anthropology,

92   University of Szeged, Hungary, [66]The Museum Organisation ROMU, Archaeology,

93   Denmark, [67]Archaeology and Palaeoecology, School of Natural and Built Environment,

94   Queen's University Belfast, Northern Ireland, [68]Vesthimmerlands Museum, Denmark,

95   [69]Universiteit Utrecht, The Netherlands, [70]Department of Informatics, University of Zurich,

96   Switzerland, [71]Museum of Archaeology, University of Stavanger, Norway, [72]Staatliches

97   Konservatoramt, Saarbruecken, Germany, [73]University of Helsinki, Finland, [74]Finnish

98   Institute of Health and Welfare, [75]Museum of Archaeology, Varna, Bulgaria, [76]Institute for

99   History of Medicine, First Faculty of Medicine, Charles University, Prague, Czech Republic,

100   [77]The National Museum of Denmark, Prehistory, Middle Ages and Renaissance, Denmark,

101   [78]The Viking Ship Museum, Roskilde, Denmark, [79]Argea Consultores SL, Madrid, Spain,

102   [80]Département de Histoire des Arts et Archéologie, University of Burgundy, France,

103   [81]Institute of Archeology, Moscow, Russia, [82]Vendsyssel Historiske Museum, [83]Salzburg

104   Museum, Austria, [84]Archäologische Staatssammlung, [85]University of Bergen, Department of

105   Cultural History, Norway, [86]Lund University, Sweden, [87]Sydsvensk arkeologi, Malmö,

106   Sweden, [88]University of Lisbon - UNIARQ (Centro de Arqueologia da Universidade de

107   Lisboa), Faculdade de Letras, Alameda da Universidade, 1600-214 Lisboa, Portugal,

108   [89]UMR6298, ARTEHIS, CNRS (Centre National de la Recherche Scientifique), Dijon,

109   France, [90]University of Paris I - Panthéon-Sorbonne, France, [91]UMR 8215 CNRS,

110   [92]Department of Clinical Medicine, Faculty of Health Sciences, University of Copenhagen,

111   Copenhagen, Denmark, [93]Departments of Integrative Biology and Statistics, UC Berkeley,

112   USA, [94]MARUM, University of Bremen, Germany

113

114

115   *Corresponding authors; email: ewillerslev@sund.ku.dk, hugh.mccoll@sund.ku.dk,

116   guus@hum.ku.dk, kristian.kristiansen@archaeology.gu.se

117

118

Summary

Today, Germanic languages, including German, English, Frisian, Dutch and the Nordic languages, are widely spoken in northwest Europe. However, key aspects of the assumed arrival and diversification of this linguistic group remain contentious[1–3]. By adding 712 new ancient human genomes we find an archaeologically elusive population

124 **entering Sweden from the Baltic region by around 4000 BP. This population became**

125 **widespread throughout Scandinavia by 3500 BP, matching the contemporaneous**

126 **distribution of Palaeo-Germanic, the Bronze Age predecessor of Proto-Germanic[4–6].**

127 **These Baltic immigrants thus offer a new potential vector for the first Germanic**

128 **speakers to arrive in Scandinavia, some 800 years later than traditionally assumed[7–12].**

129 **Following the disintegration of Proto-Germanic[13–16], we find by 1650 BP a southward**

130 **push from Southern Scandinavia into presumed Celtic-speaking areas, including**

131 **Germany, Poland and the Netherlands. During the Migration Period (1575–1375 BP),**

132 **we see this ancestry representing West Germanic Anglo-Saxons in Britain, and**

133 **Langobards in southern Europe. We find a related large-scale northward migration**

134 **into Denmark and South Sweden corresponding with historically attested Danes and the**

135 **expansion of Old Norse. These movements have direct implications for multiple**

136 **linguistic hypotheses. Our findings show the power of combining genomics with**

137 **historical linguistics and archaeology in creating a unified, integrated model for the**

138 **emergence, spread and diversification of a linguistic group.**

139

140 **Main (531 words)**

141 The arrival of Steppe-related groups in Europe around 5000 BP (calibrated years before

142 present/1950 CE) is commonly known as the last major prehistoric migration into the

143 region[17,18]. Archaeologically, two major European groupings of this time are the Corded

144 Ware and the Bell Beaker complexes. Linguistically, these are widely accepted as being

145 connected with the dispersals of multiple Indo-European language groups[19]. In Scandinavia,

146 both complexes have been specifically associated with the introduction of the Germanic

147 language group[7–10,20]. However, this association is complicated by the significant time gap of

148   ~2-3 millennia that exists between these first waves of Steppe-related ancestry (~5000–4500

149   BP) and the appearance of the earliest Germanic written evidence (~2000–1800 BP)[21,22].

150

151   In Scandinavia, the arrival of Steppe ancestry (~4800 BP) coincides with the archaeological

152   transition from the Funnel Beaker Culture to the Corded Ware Complex[23]. By ~4600 BP,

153   cultural boundaries emerged separating the Scandinavian Peninsula's Battle Axe culture from

154   the Jutlandic Single Grave culture, with these showing connections respectively to the east

155   and the south[24]. In addition, Scandinavia had cultural links associated with Bell Beaker

156   groups of western and central Europe[25]. At this time, no linguistic data is available on the

157   distribution of the Germanic language group. Only by the Late Bronze Age do linguistic

158   interactions with Celtic-speaking groups in the south[4,5,26] and Finno-Saamic-speaking groups

159   in the east[6,27] suggest that an early form of Germanic was present in the area between

160   northern Germany and the East Baltic. Despite some preliminary evidence of genetic

161   structure at this time in Scandinavia[28,29], its formation as well as its potential for tracing the

162   spread of Germanic remain unexplored.

163

164   By ~2000 BP, the common Germanic linguistic ancestor was diverging into several

165   subgroups, probably involving an initial East vs Northwest Germanic split[13–16,30]. This

166   linguistic process was associated with various migrations during the Roman Iron Age (1950–

167   1575 BP) and the Migration Period (1575–1375 BP). However, at present, our understanding

168   of the relevant population dynamics is limited. Genetic studies have detected a northern

169   European origin for Late Iron Age and Migration Period individuals often ascribed by post-

170   Classical authors as Goths, Anglo-Saxons and Langobards[31–35], consistent with Late Antique

171   historical sources. However, these studies have not confirmed their specific regions of origin.

172   Consequently, no comprehensive model for the spread and diversification of the Germanic

173   languages currently exists.

174

175   Ancient genomics have proven a means to address historical linguistic hypotheses[19] but have

176   been constrained by limited sample sizes and an inability to distinguish between closely

177   related populations. However, recent studies reveal that with dense ancient DNA sampling, at

178   sufficient sequencing depth for imputation, the detection of fine-scale population structure is

179   now possible[28,29,36].

180

181   To investigate the formation and diversification of Germanic-speaking populations and link

182   the findings to historical linguistic theory, we sequenced the genomes of 712 ancient

183   individuals to above 0.01X (Supplementary Note S4, Supplementary Table S1.1), with a

184   focus on the Northern European Iron Age and the bordering Celtic-speaking region of

185   western Europe[26] (Fig. 1). Together with published ancient genomes from around the world,

186   we selected samples with suitable average depth of coverage for imputation (~0.1X for whole

187   genomes)[28,29,37]. After overlapping with publicly available ancient SNP capture data suitable

188   for imputation (~1X on targeted SNPs), removing close relatives and applying quality filters,

189   the final imputed dataset contained 578 new and 4,009 published individuals covering

190   690,211 SNPs (Supplementary Note S4).

191
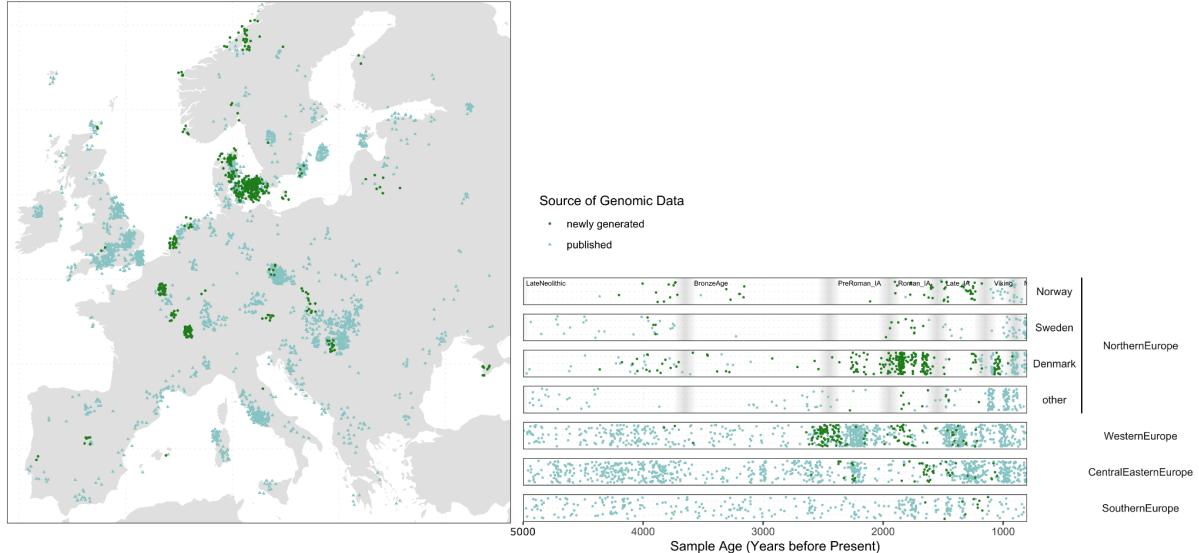
Fig. 1. Geographic and temporal sampling of a subset of ancient individuals included in the final dataset. Newly generated (green) and published (light blue) ancient individuals from the Late Neolithic / Early Bronze Age to the Viking Age. Grey bars on the timeline represent the boundary between historical periods denoted in the top panel.

Fine-scale resolution of early Steppe ancestry in Europe (5000–4000 BP)

We explored the genomic affinities between all individuals in the dataset using the identity-by-descent (IBD) hierarchical clustering method (Supplementary Note S5.2) and mixture modelling (Supplementary Note S5.3) to discern the closely related genomic ancestries[28]. Here, clusters form on the basis of the long shared genomic segments between all pairs of individuals within the dataset, rather than by proportions of the deeply diverging ancestries they carry. As discrete clustering does not display the complexities of admixture, potentially giving false impressions of continuity, we applied IBD mixture modelling to assess the genetic structure within the clusters. In brief, we created a 'palette' for every individual, based on the total length of IBD segments shared between that individual and all 386 clusters in the dataset. We then define sets of individuals from specific clusters as 'sources', and

210   modelled the palettes of 'target' individuals as a mixture of all possible source palettes, using

211   an non-negative least squares approach, similar to chromosome painting[38]. As a result, even

212   though a temporally distant source and target individual may share very few segments, the

213   shared segments with common intermediate populations are informative in the modelling. We

214   find this method to be effective for the sampling density and temporal distances relevant to

215   this study (Supplementary Note S5.3). To understand the general trends through time, we

216   applied ordinary spatio-temporal kriging[39] to the mixture modelling results (Supplementary

217   Note S5.3, Extended Data Figures 1-7).

218

219   We focused our efforts within the last ~5000 years, commonly considered the hallmark for

220   introduction of Steppe ancestry across Europe and widely acknowledged as a likely terminus

221   post quem for the spread of the Indo-European language family[17,18]. We find the majority of

222   European individuals over the last 5,000 years fall within four main IBD clusters, with a

223   varying geographical and cultural distinction for each (Supplementary Note S5.2.1). Based on

224   this close correspondence with individuals assigned to various archaeological groups, we

225   refer to these clusters as Yamnaya-, Corded Ware (East)-, Corded Ware (North)-, and Bell

226   Beaker-related. Notably, individuals from each cluster are placed adjacent to each other in a

227   standard western Eurasian PCA (Fig. 2, Supplementary Note S5.1), and each cluster occupies

228   different positions along the well-established cline of Steppe-Farmer ancestry that formed in

229   Europe from the Bronze Age.

230
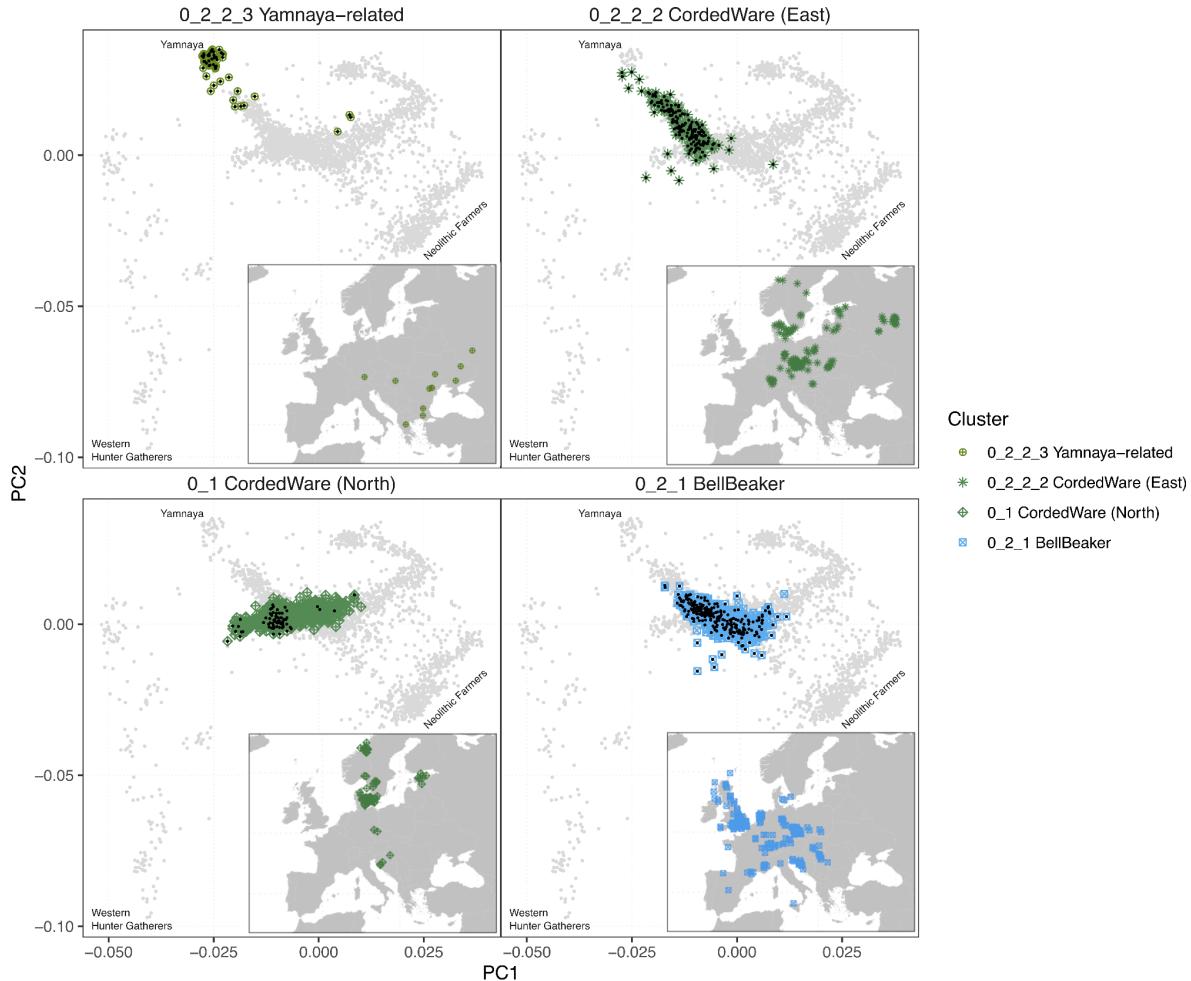
Fig. 2. Distribution of the main four Steppe-related clusters in Geographical (inset) and western Eurasian PCA space. Samples older than 2800 BP are indicated with a black point on the PCA. On the map, only samples older than 2800 BP are shown.

Importantly, our mixture modelling results show that only samples that are modelled with a high proportion of Steppe ancestry fall within these clusters (Supplementary Fig. S5.16, Supplementary Note S5.2.3), meaning the Steppe ancestry of many individuals is not represented by the clustering. Of particular importance are many Bronze Age and later southern European individuals, whose clustering is informative of their Farmer rather than their Steppe ancestry. To understand how the Steppe ancestry of individuals of primarily Farming ancestry are related to the structure mentioned above, we applied IBD mixture

243   modelling with sources representative of the Yamnaya, Corded Ware and Bell Beaker

244   clusters (Supplementary Note S5.3).

245

246   We find the Steppe ancestry of the majority of the more southern individuals to be modelled

247   as Bell Beaker-related (Fig. 3, Supplementary Note S5.2.3, Supplementary Fig. S5.17),

248   together with a series of previously unidentified admixture clines (Fig. 4, Extended Data Fig.-

249   8, Supplementary Note S5.3.3). For the Yamnaya, Corded Ware and Bell Beaker-related

250   source clusters, we found the difference in the mean proportion of Steppe ancestry to be

251   significantly higher for other individuals archaeologically assigned to the corresponding

252   culture, than to the other cultures (Supplementary Note S5.3.7). For many samples, an

253   archaeological assignment was not present, however we found that samples with Steppe

254   ancestry modelled primarily as Corded Ware-, Bell Beaker- or Yamnaya-related tend to

255   correspond closely with the regions ascribed to each culture in the archaeological record (Fig.

256   3).

257

258   During the Late Bronze Age (3150–2450 BP), the linguistically inferred presence of Palaeo-

259   Germanic in Southern and Eastern Scandinavia falls within the geographic boundaries of the

260   Corded Ware-related ancestries and cultures (Fig. 3). By the Roman Iron Age, Germanic-

261   speaking regions continue to fall within the broader range of individuals of Corded Ware-

262   related ancestry, bordered to the south by Celtic-speaking regions and Bell Beaker-related

263   ancestries (Fig. 3). For Iron Age regions assigned as likely Germanic- or Celtic-speaking, we

264   similarly find in all instances highly significant correlation with Corded Ware and Bell

265   Beaker ancestries (Supplementary Note S5.3.7). This correspondence and continuity suggest

266   that Germanic emerged and primarily continued to be spoken by populations of Corded

267   Ware, rather than Bell Beaker-related ancestry.
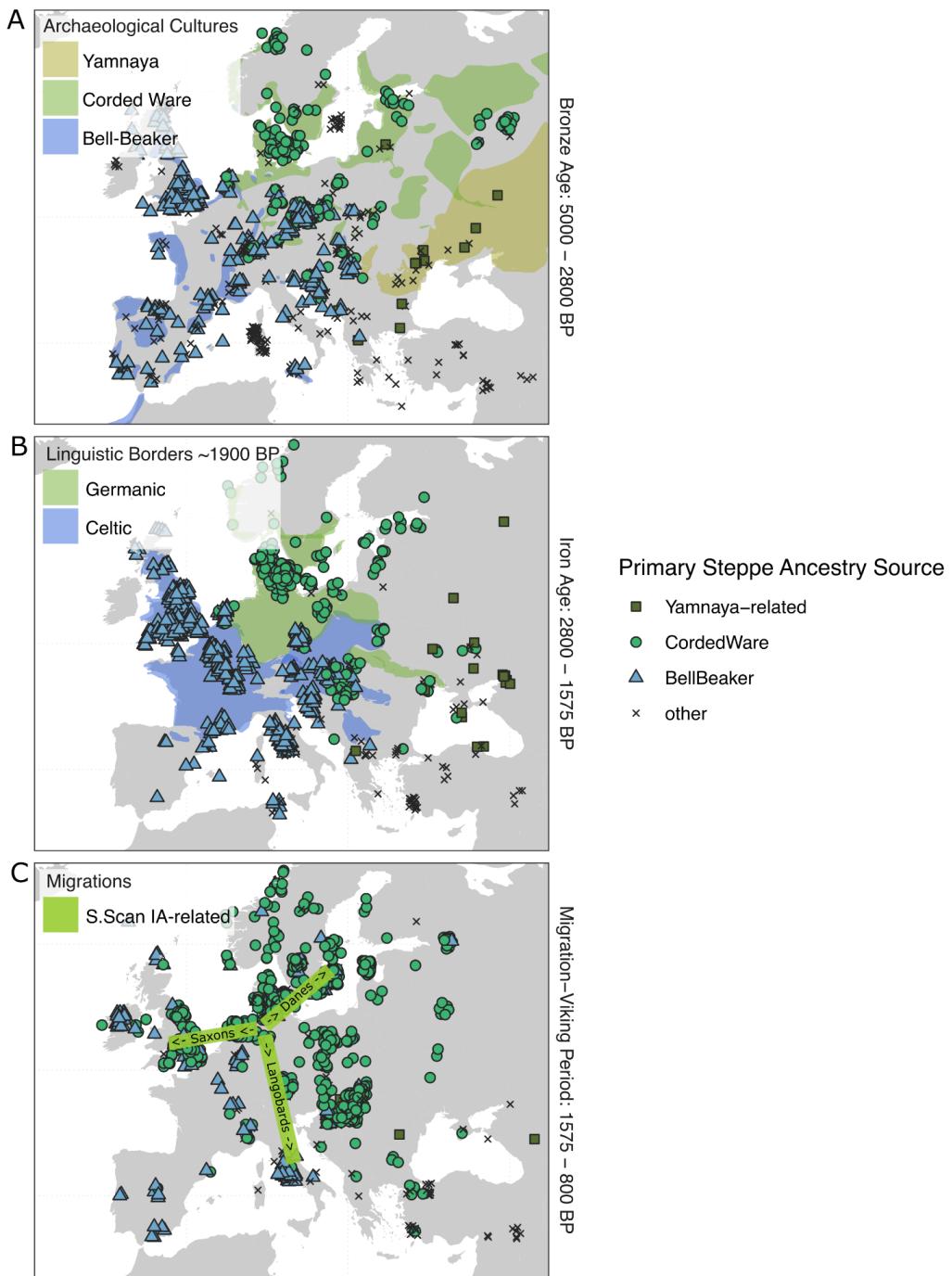
Fig. 3. Geographical distributions of major archaeological cultures, language families and Steppe ancestry sources. Individuals modelled with less than 10% Steppe ancestry or less than 66% from one of the source groups are indicated with an 'x'. Archaeological boundaries reproduced from Furholt et al.[40].

274

275
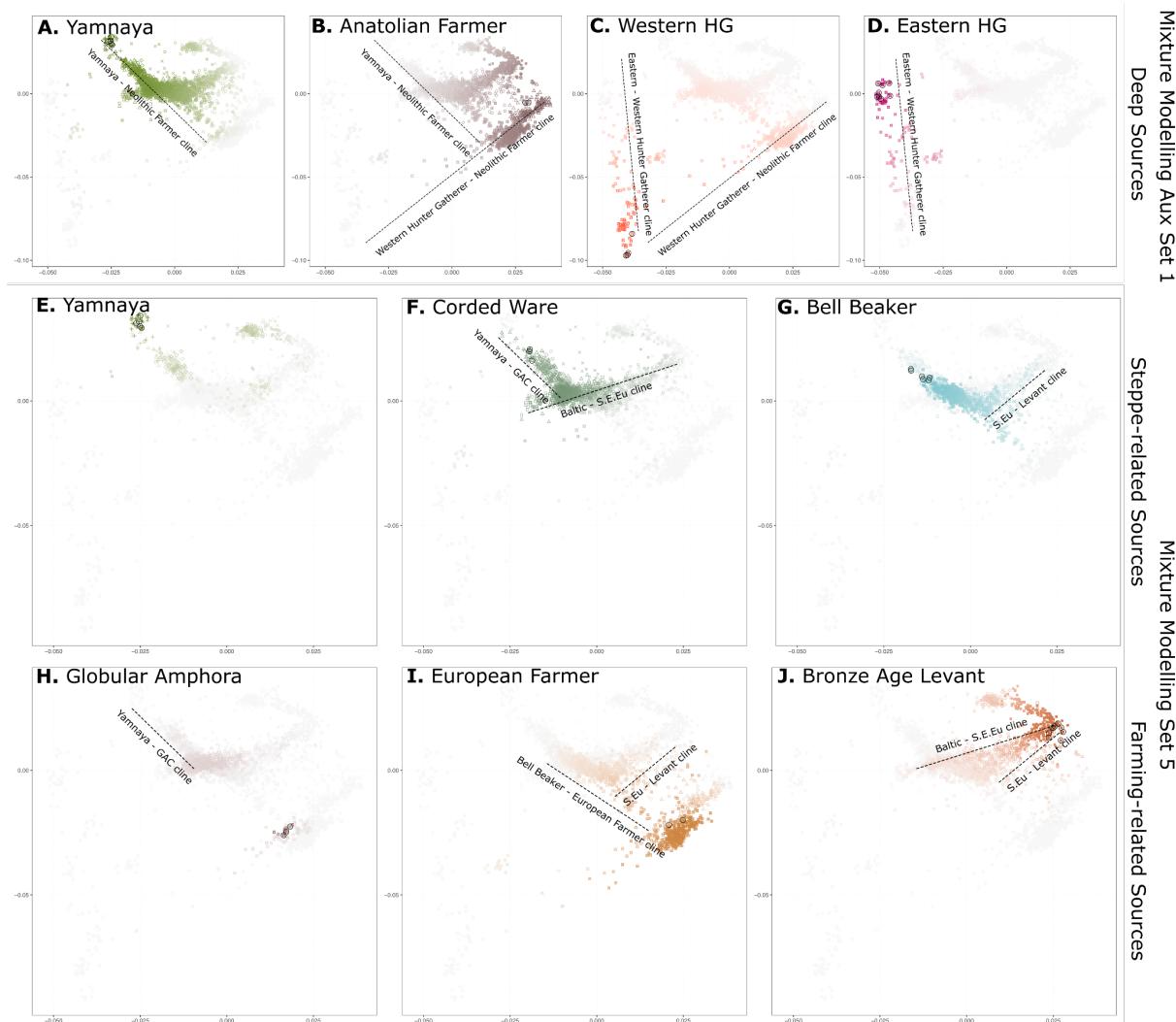
Fig. 4. A subset of mixture modelling results highlighting admixture between various Steppe and Farming-related sources displayed on the western Eurasian PCA (Supplementary Note S5.3.1, Supplementary Fig. S5.20). Auxiliary Set 1 highlights the well-established clines representing the diversity of western Eurasian Hunter-Gatherers (C and D), the arrival of Neolithic Farmers (B) in Europe admixing with the local Hunter-Gatherers (C), and the arrival of Yamnaya-related ancestry (A) admixing with European Farmers. Mixture modelling Set 5 further highlights the clines between various Steppe-related ancestry sources: E) Yamnaya-related, F) Corded Ware-related and G) Bell Beaker-related, and Farmer ancestry-related sources: H) Globular Amphora Culture-related, I)

285   European Farmer-related and J) Levant/Bronze Age Anatolia-related. Source individuals are

286   circled here and detailed in Supplementary Table S5.3 (Set 5) and S5.4 (Aux Set 1).

287   Admixture proportions follow a gradient from full colour (100%) to grey (0%).

288

289   From the Late Bronze Age onwards, irrespective of clusters, the Steppe ancestry in almost all

290   Europeans is modelled by either Corded Ware (East) or the Bell Beaker sources (Fig. 3,

291   Extended Data Fig.-4). Notable exceptions appear in the coastal region of the Netherlands

292   and in Bohemia, where the two complexes overlap (Supplementary Note S5.7.2). In the

293   Netherlands, we identify an 'Eastern North Sea' cluster that persists from 3700 BP to 1700

294   BP and is modelled with equal proportions of Corded Ware- and Bell Beaker-related ancestry

295   (Supplementary Fig. S5.172, Supplementary Note S5.7.2). The existence of this genetic

296   outlier has a linguistic analogue in the so-called 'Nordwestblock' hypothesis, which claims

297   that the Netherlands harboured an Indo-European language distinct from both Celtic and

298   Germanic[41].

299

300

301   **Scandinavian population dynamics from the Late Neolithic/Bronze Age to**

302   **the Iron Age (4000–2500 BP)**

303   Although Corded Ware-related ancestry persisted in northern Europe from the Bronze Age to

304   the Iron Age (Fig. 3), the extent to which migrations occurred within this region is poorly

305   understood. Linguistically, this period encompasses the prehistoric evolution of the Germanic

306   branch following its split from the Indo-European proto-language and comprises two phases:

307   Palaeo- and Proto-Germanic. Lexical exchanges with Celtic in the south[4,5] and Finno-Saamic

308   in the east[6,27] indicate that Palaeo-Germanic was present between northern Germany and the

309   East Baltic by the Late Bronze Age. The timing of these lexical exchanges is evident in loans

310 predating the so-called Germanic sound shifts, which had happened by the beginning of the

311 Iron Age, which in Scandinavia started around 2500 BP. The entry of Germanic loans into

312 the Finnic Uralic group cannot be earlier than the putative arrival of their speakers in the

313 Baltic, which according to the latest studies coincides with the arrival of Siberian ancestry

314 after 3500 BP and likely with the cist graves from around 3150-3050 BP[42]. Some early Celtic

315 loans likewise predate the Germanic sound shifts. As a result Palaeo-Germanic can be said to

316 have been in contact with both Finnic in the east and Celtic in the south during the period

317 3200-2500 BP, a period corresponding to the Late Bronze Age. To investigate relevant

318 demographic shifts at this time, we first established the population substructure within

319 Bronze Age Scandinavia by reclustering all ancient samples older than 2800 BP

320 (Supplementary Note S5.2.2, Supplementary Table S5.2).

321

322 Within Scandinavia, three clusters are apparent (Extended Data Fig.-9): 1) an Early

323 Scandinavian cluster, including individuals from the Swedish Battle Axe Culture, some of the

324 oldest Danish individuals, and nearly all Norwegians; 2) a later 'Southern Scandinavian'

325 cluster restricted to Denmark and the southern tip of Sweden; and 3) an even later 'Eastern

326 Scandinavian' cluster, spanning Sweden and geographically overlapping with the Southern

327 Scandinavia cluster. In all three clusters, we see a striking correspondence between Y-

328 haplogroups and IBD clusters (Extended Data Fig.-9A), largely driven by different

329 frequencies of haplogroups I1a (I1a-DF29), R1a1a1b1a3a (R1a-Z284) and R1b1a1b1a1a1

330 (R1b-U106), all being strongly associated with Scandinavian ancestry (Supplementary Note

331 S5.2.2). We find a large degree of geographic overlap between the Early, Southern and

332 Eastern Scandinavian clusters of the pre-2800 BP individuals and three subclusters detected

333 in the original northern Corded Ware cluster: Western, Southern and Eastern Scandinavian,

334 respectively (Extended Data Fig.-9B, Supplementary Note S5.2.2). The presence of genetic

335   structure in the Bronze Age has previously been detected[28], however, its formation and

336   relation to the Bronze Age distribution of Palaeo-Germanic, and to the genetic structure of

337   Iron Age Scandinavia, have not been addressed.

338

339   The three pre-2800 BP clusters are found across the region in which Palaeo-Germanic is

340   assumed to have been spoken, and understanding their formation may be informative for the

341   introduction of the Germanic branch of the Indo-European language family to Scandinavia.

342   To investigate this, we applied mixture modelling, using a variety of deep Hunter-Gatherer

343   sources. We find the Yamnaya-related ancestry in the three Scandinavian clusters to be

344   modelled as Eastern Hunter-Gatherer and Caucasus Hunter-Gatherer (Supplementary Fig.

345   S5.27, Supplementary Fig. S5.23: Set 2). This is in agreement with expectations for Steppe

346   ancestry in Europe[43]. However, the early (4100–2800 BP) Eastern Scandinavians are distinct

347   in their relatively high proportion of Eastern Hunter-Gatherer ancestry, compared to Western

348   and Southern Scandinavians (Supplementary Note S5.3.4). To identify the specific source of

349   this Hunter-Gatherer ancestry, we included additional Hunter-Gatherer sources from the

350   region (Norway, Sweden, Denmark, Latvia and Lithuania) together with Yamnaya. We find

351   that the Eastern Scandinavians' Hunter-Gatherer ancestry is modelled entirely by the

352   Latvian/Lithuanian Hunter-Gatherer source from across the Baltic, rather than the local

353   Scandinavian Hunter-Gatherers (Supplementary Note S5.3.4, Supplementary Fig. S5.23: Set

354   3). In contrast, the Southern and Western Scandinavians are modelled with additional

355   Western Hunter-Gatherer ancestry. These mixture modelling results are consistent with the

356   subtle differences in the distribution of individuals older than 2800 BP in the Scandinavian

357   clusters in the western Eurasian PCA (Supplementary Fig. S5.7, Supplementary Fig. S5.8,

358   Supplementary Fig. S5.9). To confirm these patterns, we ran rotating qpAdm models with 10

359   different Hunter-Gatherer clusters as sources, including representatives of the previously

360   identified genetic structure of Mesolithic Scandinavia[44–46]. While an early Corded Ware

361   cluster could be modelled as only Yamnaya and Globular Amphora, the Southern and

362   Western Scandinavians each required a contribution by Hunter Gatherers from continental

363   Europe and Scandinavia respectively. Only when including a second Hunter-gatherer source

364   could the Eastern Scandinavians be well modelled, in many combinations, suggesting the true

365   source has not yet been sampled. The presence of a non-local Hunter-Gatherer ancestry

366   suggests the three clusters did not diversify from a single migration into the region. Although

367   individuals from all three clusters likely spoke Indo-European languages, the Corded-Ware

368   related Battle Axe Culture of the Scandinavian Peninsula and the Single Grave Culture of the

369   Jutlandic Peninsular are traditionally assumed to be the vectors for Indo-European dialects

370   ancestral to Germanic (Supplementary Note S7.1.3), fitting closely with the early distribution

371   of Western and Southern Scandinavian clusters respectively. However, the Eastern

372   Scandinavian cluster must also be considered as a potential vector for the arrival of

373   Germanic.

374

375   The Eastern Scandinavians first detection 6-800 years after the earliest Corded Ware

376   populations in Scandinavia (Extended Data Fig.-9A), and the presence of a Hunter-Gatherer

377   ancestry, not well represented by the three waves of Hunter-Gatherers previously identified in

378   Scandinavia, points to an additional, late arrival into Scandinavia by the ancestors of the

379   Eastern Scandinavians. The Hunter-Gatherer ancestry suggests a link across the Baltic or

380   from the northeast along the Baltic coastline. With regards to their Steppe-related ancestry,

381   the Corded Ware individuals from Lithuania (4842–4496 BP), Latvia (4833 BP) and Estonia

382   (4638–4400 BP) are not well-modelled by the Eastern Scandinavians, suggesting a source

383   region further north (Supplementary Fig. S5.28). Notably, Corded Ware and Hunter-Gatherer

384   genomes from Finland and the northeast coast of Sweden are not represented in the dataset,

385   and may also be a suitable source for the Hunter-Gatherer ancestry in Eastern Scandinavians.

386   Such a location would be consistent with strontium isotopes of the Late Neolithic Swedes of

387   Central Sweden which link with eastern and northern Sweden, Finland and possibly Karelia

388   and with similarities in pottery styles between Late Neolithic Sweden and the Kiukainen

389   culture (4500–3800 BP) of southwestern Finland and the Åland Islands[47–50]. Combined, the

390   results point to the presence of an unsampled hunter-gatherer population, likely carrying I1

391   haplogroups, admixing with a Corded Ware-related population, similar to those of

392   Scandinavia, to form the Eastern Scandinavians somewhere between Finland and Northeast

393   Sweden. At present, no genomes from this region and time period exist.

394

395   Interestingly, the arrival of this first Eastern Scandinavian ancestry coincides with a new

396   burial rite of gallery graves in South Sweden, the Danish Isles[25] and Norway[51], a new house

397   type[52–54] and the first durative bronze networks[55]. These events culminated at the onset of the

398   Nordic Bronze Age (3550 BP) with cultural homogenisation and the end of the east-west

399   divide in northern Europe[53].

400

401
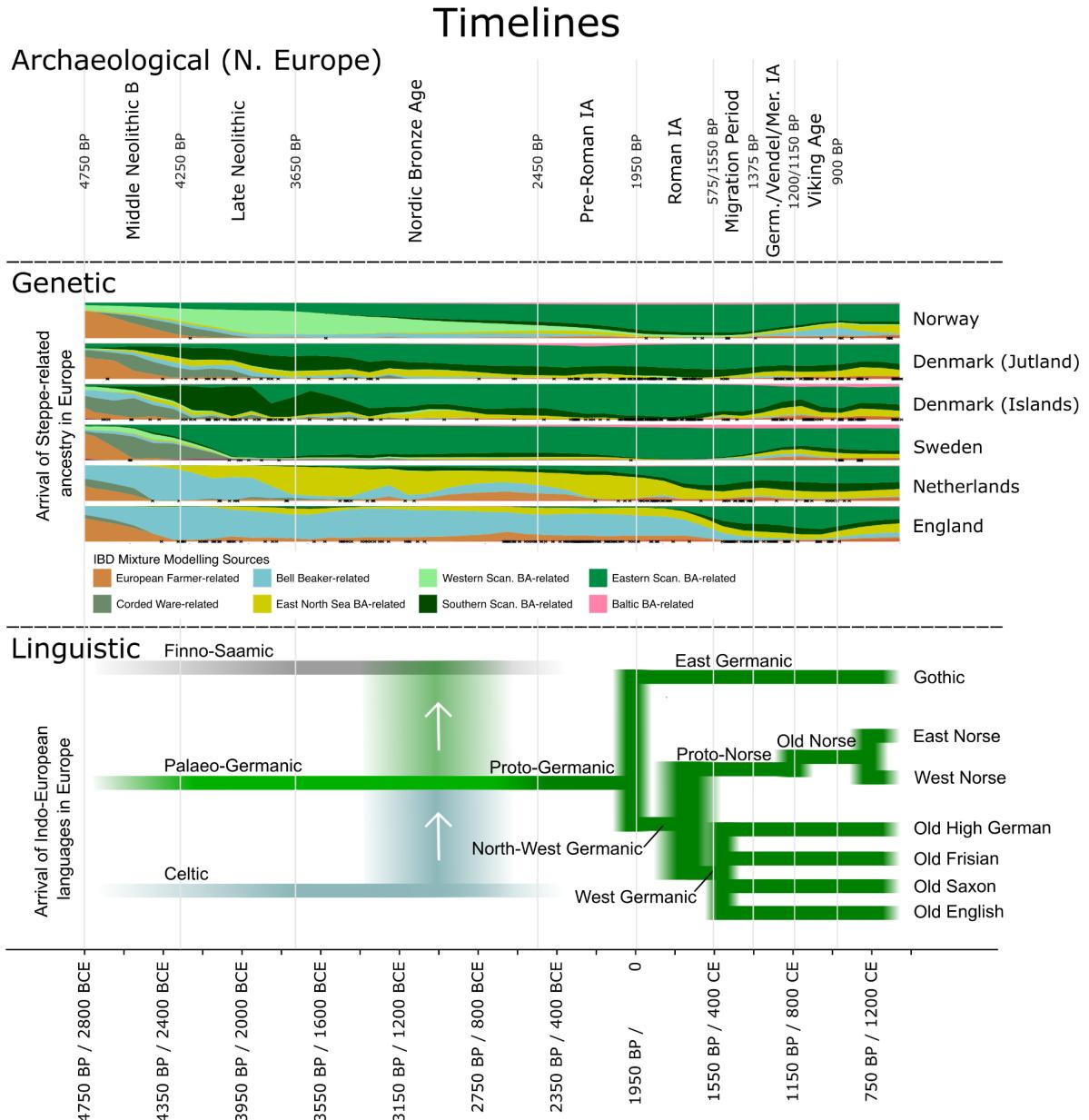
Fig. 5. A timeline showing Archaeological Periods, Spatiotemporal Kriging results with the general trends of the proportion of ancestries using Bronze Age European sources (Set 5), and relevant events in the formation of Germanic languages. During the Late Neolithic, the highest proportions of Eastern, Western and Southern Scandinavian Bronze Age ancestry are in the local region, however some ancestry remains modelled as Corded Ware (East)-related. Throughout the Bronze Age and into the Iron Age, Eastern Scandinavian ancestry becomes widespread, coinciding with loan words passing from Celtic to Germanic,

410   and Germanic to Finno-Saamic (indicated by the arrows). Samples within 250 km of the

411   coordinates assigned to each country are indicated with an 'x', providing insight into the

412   confidence with which we can assess the timing of the transition at these coordinates. The

413   coordinates representing each country can be found in Supplementary Fig. S5.134.

414

415   The Nordic Bronze Age is key in linking the three genetic clusters of the Late Neolithic with

416   the demographic structure of the Iron Age. Archaeologically, the Nordic Bronze Age is a

417   period of strong cultural homogenisation in southern Scandinavia, starting around 3600 BP,

418   forming the so-called Nordic Cultural Zone, which lasted until the start of the Iron Age (2450

419   BP)[56,57]. This homogenisation is mirrored genetically. At the beginning of the period, we

420   observe a number of admixed Norwegian and Danish Bronze Age outliers, each carrying

421   local and Eastern Scandinavian ancestry (Supplementary Fig. S5.29). By the Iron Age,

422   however, the population structure has become more homogenised by the widespread presence

423   of Eastern Scandinavian ancestry, the trend of which is visible in the spatio-temporal kriging

424   results produced from the mixture modelling output (Fig. 5, Supplementary Note S5.3.8). The

425   impact of this expansion is most apparent on the Danish Isles, followed by Norway and

426   finally the Jutlandic Peninsula. The least impact is seen in Jutland, where a large degree of

427   genetic structure distinguishes it from the rest of Scandinavia (Fig. 5).

428

429   To assess whether the spread of Eastern Scandinavian ancestry is compatible with being a

430   vector for Palaeo-Germanic, understanding the timing of this migration is crucial.

431   Linguistically, the Late Bronze Age is the period during which Palaeo-Germanic donated

432   vocabulary to Finno-Saamic in the east[6,27] and adopted vocabulary from Celtic in the south[4,5]

433   suggesting that Palaeo-Germanic was spoken within Sweden and Denmark[4,5]. However, no

434   demographic vector has yet been identified to match this distribution. Although genomes are

435   scarce for this period due to prevalent cremation practices, we used DATES[58] to date the

436   admixture time between the Eastern Scandinavians and the Southern Scandinavians for

437   various Bronze Age and Iron Age clusters in Denmark, to confirm the formation of the Iron

438   Age populations matched that of the earliest admixed individuals we detect directly

439   (Supplementary Note S5.5). The timing of the earliest admixed individuals fits closely with

440   the point estimates from the DATES results, suggesting an admixture time of 3750–3250 BP,

441   consistent with widespread Eastern Scandinavian ancestry within the spatio-temporal range

442   proposed for Palaeo-Germanic. The admixture time for Eastern and local Bronze Age

443   ancestry in Norway (4200–2600 BP) presents a similar range, with point estimates again

444   closely matching the first direct evidence of admixed Western Scandinavian and Eastern

445   Scandinavian individuals. Conversely, at the time of the incorporation of loanwords (3200–

446   2500 BP), Western Scandinavian and Southern Scandinavian Bronze Age ancestry had

447   decreased both in geographical range and proportion. Even at its most expansive distribution,

448   Southern Scandinavian ancestry did not border the region in which Finno-Saamic was

449   spoken, and by 4200-4000, Western Scandinavian bordered neither Finno-Saamic nor Celtic.

450   Combined we find Eastern Scandinavian to be the most parsimonious vector for Palaeo-

451   Germanic.

452

453   The link between Eastern Scandinavian ancestry and Palaeo-Germanic, and the arrival of

454   Eastern Scandinavian ancestry in Norway generates the hypothesis of a dialect of Palaeo-

455   Germanic being introduced to Norway, however we note that no linguistic evidence of a

456   presence in Norway at this time exists. Moreover, the apparent formation of Eastern

457   Scandinavians from the Baltic reigon potentially sheds new light on the well-known problem

458   of the similarities shared between the Germanic and Balto-Slavic branches of the Indo-

459   European language family, pointing to prehistoric borrowing, a linguistic subclade or a

460   combination of both[59].

461

462

463   **Early Iron Age Scandinavia and the formation of Proto-Germanic (2450–**

464   **1950 BP)**

465   By the onset of the Northern European Iron Age (2450 BP), Palaeo-Germanic had

466   transitioned into Proto-Germanic, the last common ancestor of the Germanic languages,

467   through a series of defining sound changes[60–63]. This linguistic transition has been ascribed to

468   the assimilation of speakers of an unknown language[21,64], potentially indicative of admixture.

469   The Proto-Germanic speech community is assumed to have been present in southern

470   Scandinavia and northern Germany throughout the Pre-Roman Iron Age (2500–1950 BP)[65,66],

471   with the Nordic Iron Age and the Jastorf culture being likely archaeological contexts[3,67]. To

472   date, it remains untested whether such admixture occurred.

473

474   Based on IBD mixture modelling using the admixed Bronze Age individuals as sources (Fig.

475   5, Supplementary Fig. S5.32) and the DATES results (Supplementary Note S5.5), we find

476   evidence of genetic continuity in Jutland throughout this period. In the Danish Isles,

477   additional Eastern Scandinavian ancestry is present relative to Jutland; however, the DATES

478   analysis indicates that the formation of the structure of the two regions occurred at a similar

479   time. The absence of major admixture events around the supposed transition from Palaeo- to

480   Proto-Germanic suggests this linguistic phase shift may have been induced by other factors at

481   the onset of the Iron Age, such as changes in social organisation or language-internal

482   developments.

483

484 Between 1950 and 1550 BP, the earliest runic carvings attest to relatively undiverged

485 Germanic dialects being spoken across Denmark, Norway and Sweden. At this time, although

486 Eastern Scandinavian BA ancestry was widespread across the region, the genetic structure

487 that formed during the Bronze Age still persisted between Jutland, the Danish Isles, Sweden

488 and Norway. In Iron Age Jutland, and to a lesser degree the Danish Isles and Norway,

489 individuals continue to be modelled with a minor proportion of the local Scandinavian

490 Bronze Age ancestry (Fig. 5, Supplementary Fig. S5.29: Set 5). The combined genetic and

491 linguistic insights reveal that these genetically distinct Iron Age populations were united by a

492 Germanic dialect continuum.

493

494 **Roman Iron Age and the Migration Period expansions from Scandinavia**

495 For the Roman Iron Age and Migration Period (1950–1375 BP), Late Antique historical

496 sources suggest that multiple migrations of Germanic-speaking groups occurred out of

497 Scandinavia, with precise but often contradictory homelands[68]. While genetic studies have

498 confirmed a northern European origin for some individuals associated with these

499 populations[31–34], ascertaining precise source regions has proven challenging. To obtain more

500 precise geographical sources of migration, we included a series of the admixed Iron Age

501 clusters in the mixture modelling. These included subclusters from northern Jutland, the

502 Danish Isles, Sweden, and Norway (Supplementary Note S5.3.1, Set 6). At a broader scale,

503 we find a pattern south of the Nordic region that indicates independent migrations from

504 Jutland and the Scandinavian Peninsula. In western Europe, including present-day Germany,

505 the Netherlands and England, we find the Jutlandic Iron Age source to be the primary

506 Scandinavian ancestry (Fig 6A, Extended Data Fig.-10). In contrast, in the coastal region

507 south and east of the Baltic Sea, we find that populations are primarily mixtures of Swedish

508 Eastern Scandinavian and Baltic Bronze Age ancestries (Extended Data Fig.-10).

509

510
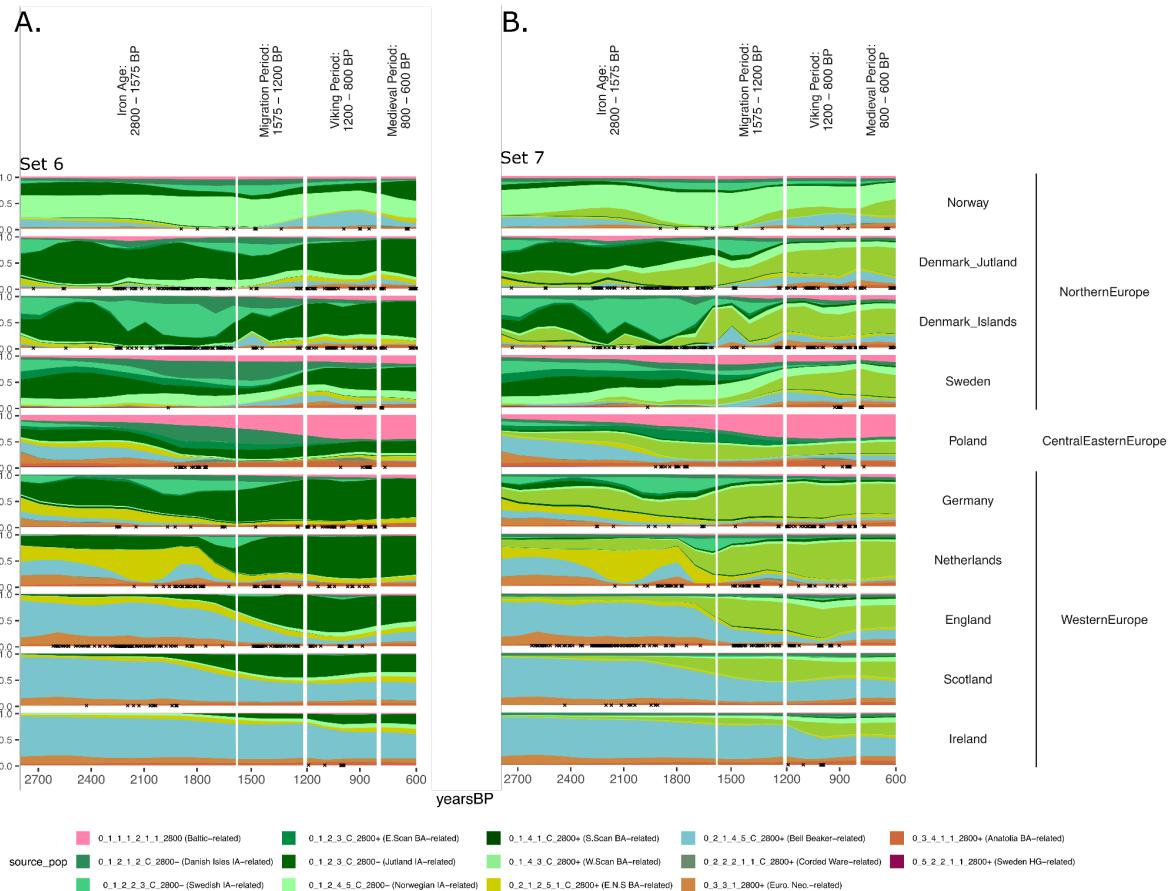
Fig. 6. Spatiotemporal Kriging results showing the general trends of the proportion of ancestries modelled for each Iron Age individual from across Europe, including Iron Age European sources from Jutland, the Danish Isles, Sweden and Norway (Panel A, Set 6), and also including northern Germany (Panel B, Set 7). Samples with 250 km of these coordinates assigned to each country are indicated with an 'x'. The coordinates representing each country can be found in Supplementary Fig. S5.134.

Historical and archaeological evidence suggest various regions along the Eastern North Sea coast, from the Netherlands to Jutland, as the sources of the migrations bringing the West Germanic predecessor of Old English to Britain[69]. Although genetic studies have indicated that a large-scale migration indeed occurred at this time, the precise source region has

523   remained unclear, with candidates ranging from the Netherlands to southern Sweden[33]. We

524   therefore added an additional Iron Age source from Mecklenburg, Northern Germany

525   (Supplementary Note S5.3.1, Set 7), to the mixture modelling. This showed ancestry in

526   Anglo-Saxons from England, Frisians from the Netherlands and Iron Age Germans to be

527   modelled as the northern German source (Fig. 6B, Extended Data Fig.-11). The observed

528   heterogeneity is consistent with the various homelands of the Angles, Saxons and Jutes along

529   the Eastern North Sea coast migrating to Britain during this period (Extended Data Fig.-12,

530   Supplementary Note S5.7.5). Prior to these expansions, individuals from northern Jutland, the

531   Danish Isles and Sweden on the contrary, are modelled primarily as the local Iron Age

532   ancestry. Therefore, we can reject Northern Jutland, the Danish Isles and Sweden as the main

533   source areas for the Anglo-Saxon migrations to Britain.

534

535   A similar cultural transition to that in Britain occurred in the Netherlands. Following the

536   Roman period, Dutch coastal areas saw a habitation hiatus around 1600 BP and the

537   subsequent appearance of a new material culture that is often referred to as Anglo-Saxon in

538   nature, and which may be related to the emergence of the Frisians[70]. As in Britain, we see a

539   genetic transition occurring from local to northern German/southern Jutlandic ancestry. Here,

540   the incoming ancestry becomes dominant in the area previously dominated by the distinct

541   Eastern North Sea population (Supplementary Note S5.7.2). Although no unadmixed Eastern

542   North Sea populations are found during the Migration Period, the population replacement

543   was not complete: many later individuals from Northern and Western Europe from this time

544   onward carry small proportions of Eastern North Sea ancestry.

545

546   A third West Germanic-speaking population, the Langobards[71], are historically documented

547   in the regions of today's Czech Republic, Hungary and Italy and have genetically been linked

548   to northern Europe[32]. However, according to post-classical origin legends, their homeland is

549   in Southern Scandinavia specifically[72]. We find that the majority of the Langobards are

550   indeed modelled by the Jutlandic Iron Age source (set 6), supporting a homeland around

551   Jutland or northern Germany (Supplementary Note S5.7.4). Notably, there are three outlier

552   Langobards from the Czech Republic and Hungary being of Eastern Scandinavian origin, and

553   a number of individuals who genetically appear to be local. This close link between Southern

554   Scandinavian ancestry and linguistically West Germanic groups is also seen in Germany

555   (Extended Data Fig.-14). Here, 56 of 76 individuals between 1650 and 775 BP fall within the

556   0_1_3_x clusters, who tend to be modelled by the Northern German source (Supplementary

557   Table S5.1).

558

559   While West Germanic groups appear closely linked with Northern German/Southern Jutland

560   ancestry, consistent with historical evidence, the genetic affiliation of East Germanic groups

561   is currently unknown. A 6th-century origin legend of the East Germanic Goths claims a

562   homeland on the Scandinavian Peninsula[73]. The earliest evidence of plausibly East Germanic-

563   speaking populations out of Scandinavia are the individuals from the Wielbark culture of

564   present-day Poland, dated to around 1950 BP[34]. We find these individuals to be modelled

565   primarily by Eastern Scandinavian BA and Swedish Iron Age ancestry (Supplementary Note

566   S5.7.4), supporting a migration from a region and population distinct from those of the West

567   Germanic groups – a scenario consistent with Gothic oral history[68]. In contrast, later

568   individuals associated with the East Germanic-speaking groups, the Ukrainian Ostrogoths

569   and the Iberian Visigoths, are modelled with local ancestries (Supplementary Note S5.7.4).

570   Two exceptions are the Iberian Visigoths, who genetically fall on the Baltic cline, suggesting

571   an origin in North East Europe, but not specifically in Eastern Scandinavia. The genetic

572   evidence here is consistent with a two-step model on the ethnogenesis of the Goths: an initial

573   migration from Scandinavia, followed by the acculturalization of others, including Baltic and

574   Slavic groups[74].

575

Migration Period expansions into Denmark and Sweden

577   Unlike West and East Germanic, the distribution of North Germanic across Scandinavia is

578   not typically associated with population migrations. Runic inscriptions from across this area

579   testify to Germanic dialects that remained relatively similar to Proto-Germanic between 2000

580   and 1500 BP. However, from the beginning of the Migration Period to the Viking Period

581   (1575–1200 BP), accelerated linguistic change led to the formation of Old Norse[75], the

582   common language of Viking-Age Scandinavians. Following this linguistic transition, the

583   original Germanic runic script, known as the Elder Futhark, was fundamentally remodelled,

584   giving rise to the Younger Futhark, a script exclusively used to document Old Norse[76,77].

585

586   Interestingly, we observe a striking genetic transition in the Danish Isles and southern

587   Sweden around the time of this linguistic transition. Prior to 1575 BP, individuals in these

588   areas are modelled almost exclusively with Eastern Scandinavian Bronze Age ancestry,

589   whereas by 1250 BP, i.e. approximately 100 years before the Viking Period, and increase in

590   Southern Scandinavian and ENS ancestry is detected (Fig. 6, Extended Data Fig.-13). By

591   using the Iron Age Scandinavian sources, the extent of the impact can be seen (Extended

592   Data Fig.-10). On the Danish Isles, the local ancestry that was widespread from 2200 to 1600

593   BP essentially disappears by 1250 BP, however the rate at which this transition occurs is not

594   clear . This contrasts with nearby regions, where Jutlandic Iron Age Southern Scandinavian

595   and Swedish Iron Age Eastern Scandinavian ancestry persists through to the Viking Age,

596   despite being less prominent than before.

597

598   To understand the source of these migrations, we turned to the period from 1250 BP onwards,

599   when sampling improves as inhumation burials become more common. Here, most

600   individuals can be modelled with small proportions of ancestries that prior to 1575 BP were

601   primarily found south of Scandinavia: Eastern North Sea coast ancestry, Celtic ancestry from

602   Britain, Ireland and France, Slavic-related ancestry from Northeastern Europe

603   (Supplementary Note S5.7.6), and European Farming ancestry from western Europe

604   (Extended Data Fig.-10, Extended Data Fig.-14). While previous studies have documented an

605   influx of these more southern ancestries[78,79], continuity of the local Scandinavian ancestry has

606   generally been assumed.

607

608   To further investigate the transition within Denmark and southern Sweden, we again applied

609   the mixture modelling with the Southern Scandinavian Iron Age sources from northern

610   Jutland and from northern Germany (Extended Data Fig.-11). We find the Scandinavian

611   ancestry of the later Danish Isles and southern Sweden individuals to be modelled as the

612   northern German IA source ancestry rather than the more local northern Jutland IA source.

613   This suggests large-scale population movements into Denmark and southern Sweden. In

614   Jutland, we also see an influx of northern German IA ancestry. However, the impact is much

615   less pronounced than for the Danish Isles. Due to the absence of samples between northern

616   Jutland and northern Germany, we cannot rule out a southern Jutland origin as a source.

617

618   The above findings of a large-scale demographic shift have potential relevance for

619   understanding the linguistic dynamics of Iron Age Scandinavia, suggesting that the

620   emergence of Old Norse coincides with a period of social and demographic instability[80]. In

621   addition, the spread of Southern Scandinavian ancestry may partially explain why the Old

622   Norse language was referred to by its speakers as *dǫnsk tunga*, i.e. the Danish tongue, even in

623   Norway, Iceland and Sweden[81,82]. However, the close correspondence between Northern

624   German/Southern Jutlandic IA ancestry and Old Norse linguistic identity in Southern

625   Scandinavia is not seen in Norway. Here, similar social changes as in South Scandinavia

626   occurred during the same period, however, the Iron Age to Viking Age transition in Norway

627   is one of genetic continuity (Supplementary Note S5.7.3). In Norway, the cultural changes

628   occurred with limited genetic impact from populations from continental Europe. With the

629   exception of a single early Viking sample, the majority of Viking Age Norwegians display

630   either local ancestry or ancestry reflecting back-migrations from Celtic regions of Britain and

631   Ireland. Of note, the linguistic boundary between the East and West Norse dialect areas[75,83]

632   roughly corresponds to the border between Southern Scandinavian and Western Scandinavian

633   ancestry during the Viking Age (Figures S5.108 and S5.110).

634

635   Due to cremation practices, few genomes exist from the transition period (1575–1200 BP),

636   requiring us to rely on additional lines of evidence for population dynamics during this period

637   to assess when these major transitions in Denmark and Southern Sweden may have occurred

638   (Extended Data Fig.-15). Volcanic activity (1414 and 1411 BP) and the Justinian Plague

639   (1409 BP) caused a population decline in Scandinavia, from which the populations did not

640   fully recover until around 1300 BP (Supplementary Note S7.1.4)[84]. However, despite some

641   abandonment or depopulation of marginal subsistence areas[85], there was continuity in more

642   fertile and southern areas (Supplementary Note S6.4). For those who survived, the

643   subsequent improving conditions and relative abundance of resources due to a lower

644   population size would have created the opportunity for rapid expansion, as attested to in

645   historical sources in other areas (Supplementary Note S7.1.4). Despite these dramatic events,

646   the major cultural transitions began earlier, between 1550 and 1450 BP, and persisted through

647   the Viking Age (Supplementary Note S7.1.4), according to antique sources mentioning the

648   Danes living in South Scandinavia by 1450 BP[86,87] and oral histories indicating the South

649   Scandinavian royal lineage of the Danes, Swedes and Norwegians being initiated between

650   1550 and 1500 BP[88]. Based on the present archaeological and historical evidence

651   (Supplementary Note S7.1.4), we may thus conclude that the major population shift in South

652   Scandinavia between the Roman and the Viking periods was not solely driven by the climate

653   events or plague of 1450–1350 BP, but instead likely took hold between 1550 and 1450 BP,

654   coinciding with the Anglo-Saxon migrations and corresponding to a major cultural change

655   with the introduction of a new elite culture. Historical texts and legends suggest various

656   origins of the Danes [68,86–88], discussed in detail in Supplementary Note S7.1.4. Here we show

657   genetically that the Migration Period population influx in South Scandinavia had origins

658   around Northern Germany, with close similarity to that of the Anglo-Saxons.

659

660   This result has impacts on the interpretations of Gretzinger et al.[33] in relation to the source

661   region for Saxon migrations, and Margaryan et al.[78] in relation to the major influx of Danish

662   Viking Age ancestry into England (Supplementary Note S5.7.5.). Furthermore, the dense

663   sampling and high resolution demographic inference have allowed us to establish a baseline

664   ancestry for various regions and subsequently identify outliers (Supplementary Note S5.7.1).

665

666

Conclusions

668   Our findings constitute a fundamental revision to the formation of West Eurasian ancestry in

669   Europe. By additionally correlating these genetic findings with archaeological and linguistic

670   evidence, we complement previous interpretations of the archaeological record after the

671   Scandinavian Middle Neolithic B (4800 BP), identifying multiple populations movements,

672   and propose a new model for the emergence and spread of the Germanic languages.

673

674   Following the arrival of Steppe ancestry in Scandinavia during the Middle Neolithic B (4800

675   BP), we detect the arrival of the closely related Eastern Scandinavian ancestry during the

676   Late Neolithic with links to the North East Baltic region (4000 BP), offering a previously

677   unknown vector for the introduction of an Indo-European language to Scandinavia. We

678   cannot rule out more complex scenarios, for example, in which Palaeo-Germanic spread

679   through culture diffusion, with no genetic impact, in the opposite direction to the spread of

680   Eastern Scandinavian ancestry. However, the fact that Eastern Scandinavian ancestry was

681   widespread by the Nordic Bronze Age, encompassing the broad region where Palaeo-

682   Germanic is hypothesised to have been spoken, bordering Finno-Saamic in the east and Celtic

683   in the south, makes Eastern Scandinavians the most parsimonious vector for the spread of

684   Palaeo-Germanic. By 3500 BP, Eastern Scandinavian ancestry admixes with West and South

685   Scandinavians in Norway and Denmark, suggesting that the linguistic phase shift between

686   Palaeo- and Proto-Germanic postdated the admixture event. The genetic structure that forms

687   at this time persists until the end of the Roman Iron Age (1600 BP), encompassing the

688   Palaeo- to Proto-Germanic periods as well as the subsequent period of linguistic

689   disintegration into East, North and West Germanic. Thus, we reveal that by around 2000 BP,

690   Germanic dialects unified a region divided by clear genetic borders that had existed for a

691   thousand years. In the Polish Wielbark culture, we find support for the Late Antique Gothic

692   origin myth of an exodus from Scandinavia. At the same time, however, we observe that the

693   sampled southern European Goths carry little Eastern Scandinavian ancestry, suggesting a

694   largely cultural adoption of Gothic language and identity for these individuals. Consistent

695   with linguistic consensus on a Proto-Germanic homeland in northern Germany and Southern

696   Scandinavia during the Early Iron Age, we find the descendant North and West Germanic

697   languages across Europe to be closely linked to Southern Scandinavian ancestry, in support

698   of historical documentation of the West Germanic Anglo-Saxons and Langobards. Finally,

699   the linguistic changes leading to the formation of Old Norse appear to coincide with a back-

700   migration of Southern Scandinavians to the Danish Isles and southern Sweden, but not in the

701   more northern parts of Scandinavia, suggesting that a combination of demographic and

702   cultural mechanisms drove the evolution and transmission of this language. We later find the

703   dialectal division between Old West and Old East Norse to correspond to that of the Western

704   Scandinavian Vikings and the Southern Scandinavian Vikings.

705

706   These findings underline the potential of IBD approaches to unravel genetic, archaeological

707   and linguistic prehistories, at the same time revealing marked differences in the mechanisms

708   behind the proliferation of cultural and biological features. With the increasingly fine-scale

709   resolution of population genetics, additional sampling from along the Baltic coastline will

710   allow several unanswered questions to be addressed: 1) confirmation of the proposed Bronze

711   Age source of the Eastern Scandinavians along the Baltic coast; 2) identification of the Iron

712   Age genetic boundary between East Scandinavian and Jutlandic Iron Age ancestry in

713   continental Europe between Mecklenburg and Gdansk, potentially corresponding to the

714   linguistic border between East and Northwest Germanic; 3) determining whether the spread

715   of East Germanic dialects with Eastern Scandinavian ancestry extends beyond Poland; 4)

716   determining the more localised regions both along the Eastern North Sea coast and within

717   Britain representing each of the Angles, Saxons and Jutes; and finally 5) identifying the

718   regions in Northeast Europe related to the source of Baltic- and Slavic-speaking populations.

719

720

721   **References**

722   1.   Polomé, E. C. Who are the Germanic people? in *Proto-Indo-European: The*

723   *Archaeology of a Linguistic Problem : Studies in Honor of Marija Gimbutas* (ed.

724   Gimbutas, M.) 216–244 (Institute for the Study of Man, Washington, D.C., 1987).

725   2.   Mallory, J. P. *In Search of the Indo-Europeans: Language, Archaeology and Myth*.

726   (Thames and Hudson, London, 1989).

727   3.   Schmidt, K. H. The Celts and the Ethnogenesis of the Germanic People. *Hist.*

728   *Sprachforsch. Hist. Linguist.* **104**, 129–152 (1991).

729   4.   Koch, J. T. *Celto-Germanic: Later Prehistory and Post-Proto-Indo-European*

730   *Vocabulary in the North and West*. (University of Wales Centre for Advanced Welsh

731   and Celtic Studies, 2020).

732   5.   van Sluis, P., Jørgensen, A. R. & Kroonen, G. European Prehistory between Celtic and

733   Germanic: The Celto-Germanic Isoglosses Revisited. in *The Indo-European Puzzle*

734   *Revisited: Integrating Archaeology, Genetics, and Linguistics* (eds. Willerslev, E.,

735   Kroonen, G. & Kristiansen, K.) 193–244 (Cambridge University Press, Cambridge,

736   2023). doi:10.1017/9781009261753.018.

737   6.   Aikio, A. An essay on Saami ethnolinguistic prehistory. in *A Linguistic Map of*

738   *Prehistoric Northern Europe* (eds. Grünthal, R. & Kallio, P.) vol. 226 63–117 (Société

739   Finno-Ougrienne, Helsinki: Suomalais-Ugrilainen Seura, 2012).

740   7.   Childe, V. G. *The Aryans: A Study of Indo-European Origins*. (K. Paul, Trench, Trubner

741   & Company, Limited, 1926).

742   8.   Güntert, H. *Der Ursprung der Germanen*. (C. Winter, Heidelberg, 1934).

743   9.   Anthony, D. W. Pontic-Caspian Mesolithic and Early Neolithic societies at the time of

744   the Black Sea flood: a small audience and small effects. in *The Black Sea Flood*

745   *Question: Changes in Coastline, Climate and Human Settlement* (eds. Yanko-Hombach,

746   V., Gilbert, A. S., Panin, N. & Dolukhanov, P. M.) (Springer Science & Business

747   Media, 2006).

748   10. Kristiansen, K. *et al.* Re-theorising mobility and the formation of culture and language

749   among the Corded Ware Culture in Europe. *Antiquity* **91**, 334–347 (2017).

750   11. Iversen, R. & Kroonen, G. Talking Neolithic: Linguistic and Archaeological

751   Perspectives on How Indo-European Was Implemented in Southern Scandinavia. *Am. J.*

752   *Archaeol.* **121**, 511–525 (2017).

753   12. Glob, P. V. Studier over den jyske enkeltgravskultur. in *Fortid Og Nutid* 298–302

754   (Gyldendal, 1945).

755   13. Ringe, D. Proto-Germanic. in *From Proto-Indo-European to Proto-Germanic* (ed.

756   Ringe, D.) 0 (Oxford University Press, 2017).

757   doi:10.1093/oso/9780198792581.003.0004.

758   14. Salmons, J. 59. The evolution of Germanic. in *Handbook of Comparative and Historical*

759   *Indo-European Linguistics* vol. 2 1002–1027 (De Gruyter Mouton, 2017).

760   15. Hansen, B. S. S. & Kroonen, G. J. Germanic. in *The Indo-European Language Family*

761   (ed. Olander, T.) 152–172 (Cambridge University Press, Cambridge, 2022).

762   doi:10.1017/9781108758666.010.

763   16. Agee, J. R. Using Historical Glottometry to Subgroup the Early Germanic Languages. *J.*

764   *Ger. Linguist.* **33**, 319–357 (2021).

765   17. Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–

766   172 (2015).

767   18. Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European

768   languages in Europe. *Nature* **522**, 207–211 (2015).

769   19. *The Indo-European Puzzle Revisited: Integrating Archaeology, Genetics, and*

770   *Linguistics*. (Cambridge University Press, 2023).

20. Prescott, C. Dramatic beginnings of Norway's history? Archaeology and Indo-Europeanization. in *Language and Prehistory of the Indo- European Peoples: A Cross-disciplinary Perspective* (ed. Hyllested, A.) (Museum Tusculanum Press, 2017).

21. Kroonen, G. Questions concerning the Indo-Europeanization of Scandinavia: The Germanic languages. in *Oxford Handbook of Scandinavian Archaeology* (eds. Hedeager, L., Prescott, C. & Kristiansen, K.) (Oxford University Press, Oxford, in press).

22. Mallory, J., Dybo, A. & Balanovsky, O. The Impact of Genetics Research on Archaeology and Linguistics in Eurasia. *Russ. J. Genet.* **55**, 1472–1487 (2019).

23. Malmström, H. *et al.* The genomic ancestry of the Scandinavian Battle Axe Culture people and their relation to the broader Corded Ware horizon. *Proc. R. Soc. B Biol. Sci.* **286**, 20191528 (2019).

24. Tornberg, A. & Vandkilde, H. Modelling age at death reveals Nordic Corded Ware paleodemography. Preprint at https://doi.org/10.21203/rs.3.rs-4125486/v1 (2024).

25. Müller, J. & Vandkilde, H. The Nordic Bronze Age Rose from Copper Age Diversity: Contrasts in the Cimbrian Peninsula. in *Contrasts of the Nordic Bronze Age* (eds. Austvoll, K. I. et al.) vol. 1 29–48 (Brepols Publishers, Turnhout, 2020).

26. Patterson, N. *et al.* Large-scale migration into Britain during the Middle to Late Bronze Age. *Nature* **601**, 588–594 (2022).

27. Lang, V. Formation of Proto-Finnic – an archaeological scenario from the Bronze Age / Early Iron Age. in *Congressus duodecimus Internationalis Fenno-Ugristarum, Oulu 2015: plenary papers* (eds. Mantila, H., Leinonen, K., Brunni, S., Palviainen, S. & Sivonen, J.) 63–84 (University of Oulu, Oulu, 2015).

28. Allentoft, M. E. *et al.* Population genomics of post-glacial western Eurasia. *Nature* **625**, 301–311 (2024).

796   29. Allentoft, M. E. *et al.* 100 ancient genomes show repeated population turnovers in

797   Neolithic Denmark. *Nature* **625**, 329–337 (2024).

798   30. Heggarty, P. *et al.* Language trees with sampled ancestors support a hybrid model for

799   the origin of Indo-European languages. *Science* **381**, eabg0818 (2023).

800   31. Schiffels, S. *et al.* Iron Age and Anglo-Saxon genomes from East England reveal British

801   migration history. *Nat. Commun.* **7**, 10408 (2016).

802   32. Amorim, C. E. G. *et al.* Understanding 6th-century barbarian social organization and

803   migration through paleogenomics. *Nat. Commun.* **9**, 3547 (2018).

804   33. Gretzinger, J. *et al.* The Anglo-Saxon migration and the formation of the early English

805   gene pool. *Nature* **610**, 112–119 (2022).

806   34. Stolarek, I. *et al.* Genetic history of East-Central Europe in the first millennium CE.

807   *Genome Biol.* **24**, 173 (2023).

808   35. Vyas, D. N. *et al.* Fine-scale sampling uncovers the complexity of migrations in 5th–6th

809   century Pannonia. *Curr. Biol.* **33**, 3951-3961.e11 (2023).

810   36. Ringbauer, H. *et al.* Accurate detection of identity-by-descent segments in human

811   ancient DNA. *Nat. Genet.* **56**, 143–151 (2024).

812   37. Sousa da Mota, B. *et al.* Imputation of ancient human genomes. *Nat. Commun.* **14**, 3660

813   (2023).

814   38. Lawson, D. J., Hellenthal, G., Myers, S. & Falush, D. Inference of Population Structure

815   using Dense Haplotype Data. *PLOS Genet.* **8**, e1002453 (2012).

816   39. Racimo, F. *et al.* The spatiotemporal spread of human migrations during the European

817   Holocene. *Proc. Natl. Acad. Sci.* **117**, 8989–9000 (2020).

818   40. Furholt, M. Social Worlds and Communities of Practice: a polythetic culture model for

819   3rd millennium BC Europe in the light of current migration debates. *Préhistoires*

820   *Méditerranéennes* (2020) doi:10.4000/pm.2383.

821   41. Hachmann, R., Kossack, G. & Kuhn, H. *Völker zwischen Germanen und Kelten:*

822   *Schriftquellen, Bodenfunde und Namengut zur Geschichte des nördlichen*

823   *Westdeutschlands um Christi Geburt.* (Wachholtz, Rhine River Valley, 1962).

824   42. Saag, L. *et al.* The Arrival of Siberian Ancestry Connecting the Eastern Baltic to Uralic

825   Speakers further East. *Curr. Biol.* **29**, 1701-1711.e16 (2019).

826   43. Lazaridis, I. *et al.* The genetic history of the Southern Arc: A bridge between West Asia

827   and Europe. *Science* **377**, eabm4247 (2022).

828   44. Mittnik, A. *et al.* The genetic prehistory of the Baltic Sea region. *Nat. Commun.* **9**, 442

829   (2018).

830   45. Saag, L. *et al.* Genetic ancestry changes in Stone to Bronze Age transition in the East

831   European plain. *Sci. Adv.* **7**, eabd6535 (2021).

832   46. Posth, C. *et al.* Palaeogenomics of Upper Palaeolithic to Neolithic European hunter-

833   gatherers. *Nature* **615**, 117–126 (2023).

834   47. Blank, M. Mobility, Subsistence and Mortuary practices. An interdisciplinary study of

835   Neolithic and Early Bronze Age megalithic populations of southwestern Sweden.

836   (University of Gothenburg, Gothenburg, 2021).

837   48. Blank, M. *et al.* Mobility patterns in inland southwestern Sweden during the Neolithic

838   and Early Bronze Age. *Archaeol. Anthropol. Sci.* **13**, 64 (2021).

839   49. Larsson, Å. M. Breaking and Making Bodies and Pots : Material and Ritual Practices in

840   Sweden in the Third Millennium BC. (Uppsala University, Uppsala, 2009).

841   50. Brorsson, T., Lucenius, J. & Stenbäck, N. Changing perspectives Thin section and ICP

842   analysis of Neolithic pottery from the Åland Islands. in *Helsinki Harvest: Proceedings*

843   *of the 11th Nordic Conference on the Application of Scientific Methods in Archaeology.*

844   *MASF 7* vol. 7 (MASF (Monographs of the Archaeological Society of Finland), 2021).

845   51. Østmo, E. *Krigergraver: en dokumentarisk studie av senneolittiske hellekister i Norge.*

846   (Kulturhistorisk Museum, Universtitet i Oslo, 2011).

847   52. Austvoll, K. I. *Seaways to Complexity: A Study of Sociopolitical Organisation Along the*

848   *Coast of Northwestern Scandinavia in the Late Neolithic and Early Bronze Age.*

849   (Equinox Publishing Limited, 2021).

850   53. Johannsen, J. W. The Late Neolithic Expansion: Ancient and new traditions 2350-1700

851   BC. *Dan. J. Archaeol.* **12**, 1–22 (2023).

852   54. Larsson, M. N. *et al.* Ancient Sheep Genomes reveal four Millennia of North European

853   Short-Tailed Sheep in the Baltic Sea region. 2023.06.26.544912 Preprint at

854   https://doi.org/10.1101/2023.06.26.544912 (2023).

855   55. Vandkilde, H. *The Metal Hoard from Pile in Scania, Sweden: Place, Things, Time,*

856   *Metals, and Worlds around 2000 BCE.* (Aarhus Universitetsforlag, 2017).

857   56. Kristiansen, K. & Larsson, T. B. 5:3 The iconography of ruling elites: The Kivik burial

858   and the origins of the Nordic Bronze Age. in *The rise of Bronze Age society: travels,*

859   *transmissions and transformations* (Cambridge University Press, Cambridge ; New

860   York, 2005).

861   57. Kristiansen, K. *Europe Before History. Cambridge.* (Cambridge University Press,

862   Cambridge, 1998).

863   58. Narasimhan, V. M. *et al.* The formation of human populations in South and Central

864   Asia. *Science* **365**, eaat7487 (2019).

865   59. Stang, C. S. *Lexikalische Sonderübereinstimmungen zwischen dem Slavischen,*

866   *Baltischen und Germanischen.* (Univ.-Forl., 1972).

867   60. Hedeager, L. *Iron-Age Societies: From Tribe to State in Northern Europe, 500 BC to*

868   *AD 700.* (Blackwell, 1992).

869 61. Fulk, R. D. *A Comparative Grammar of the Early Germanic Languages*. (John
870   Benjamins, 2018). doi:10.1075/sigl.3.

871 62. Kümmel, M. J. Reconstructing Proto-Germanic. in *Oxford Research Encyclopedia of
872   Linguistics* (2023). doi:10.1093/acrefore/9780199384655.013.931.

873 63. Polomé, E. C. Methodological Aspects of Glotto-and Ethnogenesis of the Germanic
874   Peoples. *Entsteh. Von Sprachen Volkern Giotto- Ethnogenetische Aspekte Eur.*
875   *Sprachen Akten Des* **6**, 45–70 (1985).

876 64. Hirt, H. Die Urheimat der Germanen: Nachwort. in *Neue Jahrbücher für das klassische*
877   *Altertum, Geschichte und deutsche Litteratur* vol. 3 570–572 (1899).

878 65. Ringe, D. A. *From Proto-Indo-European to Proto-Germanic*. (Oxford University Press,
879   2017).

880 66. Nielsen, H. F. *The Germanic Languages: Origins and Early Dialectal Interrelations*.
881   (University of Alabama Press, 1989).

882 67. Schwantes, G. Die Jastorf-Zivilisation. in *Reinecke Festschrift* (eds. Behrens, G. &
883   Werner, J.) (Mainz, 1950).

884 68. Hedeager, L. *Iron Age Myth and Materiality: An Archaeology of Scandinavia AD 400-*
885   *1000*. (Routledge, 2011).

886 69. Hines, J. *The Scandinavian Character of Anglian England in the Pre-Viking Period*.
887   (University of Oxford, 1983).

888 70. Knol, E. & IJssennagger, N. Palaeogeography and People. in *Frisians and Their North*
889   *Sea Neighbours: From the Fifth Century to the Viking Age* (eds. Hines, J. &
890   IJssennagger, N.) 5–24 (Boydell & Brewer, 2017).

891 71. Hutterer, C. J. *Die germanischen Sprachen: ihre Geschichte in Grundzügen*.
892   (Akadémiai Kiadó, 1975).

893 72. Diaconus, P. *History of the Lombards*. (University of Pennsylvania Press, 1974).

894   73. Wolfram, H. *History of the Goths*. (University of California Press, 1990).

895   74. Eiddon, I., Edwards, S. & Heather, P. Goths & Huns. in *The Cambridge Ancient History*

896   (eds. Garnsey, P. & Cameron, A.) vol. 13: The Late Empire (Cambridge University

897   Press, 1998).

898   75. Birkmann, T. 79. From Ancient Nordic to Old Nordic: Definition and delimitation of the

899   period. in *79 . From Ancient Nordic to Old Nordic: Definition and delimitation of the*

900   *period* 691–698 (De Gruyter Mouton, 2008). doi:10.1515/9783110197051-079.

901   76. Barnes, M. P. *Runes: A Handbook*. (Boydell Press, 2012).

902   77. Schulte, M. Runology and historical sociolinguistics: On runic writing and its social

903   history in the first millennium. *J. Hist. Socioling.* **1**, 87–110 (2015).

904   78. Margaryan, A. *et al.* Population genomics of the Viking world. *Nature* **585**, 390–396

905   (2020).

906   79. Rodríguez-Varela, R. *et al.* The genetic history of Scandinavia from the Roman Iron

907   Age to the present. *Cell* **186**, 32-46.e19 (2023).

908   80. Knirk, J. E. 73. Runes : Origin, development of the futhark, functions, applications, and

909   methodological considerations. in *The Nordic Languages: Volume 1* 634–648 (De

910   Gruyter Mouton, 2008). doi:10.1515/9783110197051-073.

911   81. Ottosson, K. 90. Old Nordic: A definition and delimitation of the period. in *The Nordic*

912   *languages. An international handbook of the history of the North Germanic languages*

913   vol. 1 787–793 (De Gruyter Mouton, 2002).

914   82. Riisøy, A. I. Danish Legal Procedure and a Common Scandinavian Law. in *Viking*

915   *encounters: Proceedings of the Eighteenth Viking Congress* (eds. Pedersen, A. &

916   Sindbæk, S. M.) 220–231 (Aarhus Universitetsforlag, 2020).

917   83. Haugen, E. & Markey, T. L. INTRODUCTION. in *The Scandinavian Languages* vol.

918   154 11–13 (De Gruyter Mouton, Berlin, 1973).

919   84. Arthur, F. *et al.* The impact of volcanism on Scandinavian climate and human societies during the Holocene: Insights into the Fimbulwinter eruptions (536/540 AD). *The Holocene* 09596836231225718 (2024) doi:10.1177/09596836231225718.

922   85. Price, N. Climate change and the Fimbulwinter. in *Oxford Handbook of Scandinavian Archaeology* (Oxford University Press, Oxford, in press).

924   86. Jordanes. Chapter 1:23. in *The Gothic History of Jordanes in English Version* (Princeton University Press, 1915).

926   87. Procopius. *De Bellis [History of the Wars Book I-VIII]*. (Harvard University Press, 1918).

928   88. Skre, D. 3. Rulership and Ruler's Sites in 1st–10th-century Scandinavia. in *3. Rulership and Ruler's Sites in 1st–10th-century Scandinavia* (ed. Skre, D.) 193–244 (De Gruyter, 2020). doi:10.1515/9783110421101-003.

931   89. Meyer, M. & Kircher, M. Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. *Cold Spring Harb. Protoc.* **2010**, pdb.prot5448 (2010).

934   90. Kapp, J. D., Green, R. E. & Shapiro, B. A Fast and Efficient Single-stranded Genomic Library Preparation Method Optimized for Ancient DNA. *J. Hered.* **112**, 241–249 (2021).

937   91. Gansauge, M. T. *et al.* Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic Acids Research* vol. 45 e79–e79 (2017).

939   92. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).

942   93. Briggs, A. W. *et al.* Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Res.* **38**, e87–e87 (2010).

944   94. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler

945   transform. *Bioinformatics* vol. 25 1754–60 (2009).

946   95. Fu, Q. *et al.* Genome sequence of a 45,000-year-old modern human from western

947   Siberia. *Nature* **514**, 445–449 (2014).

948   96. Rubinacci, S., Ribeiro, D. M., Hofmeister, R. J. & Delaneau, O. Efficient phasing and

949   imputation of low-coverage sequencing data using large reference panels. *Nat. Genet.*

950   **53**, 120–126 (2021).

951   97. Hanghøj, K., Moltke, I., Andersen, P. A., Manica, A. & Korneliussen, T. S. Fast and

952   accurate relatedness estimation from high-throughput sequencing data in the presence of

953   inbreeding. *GigaScience* **8**, giz034 (2019).

954   98. Browning, B. L. & Browning, S. R. Detecting Identity by Descent and Estimating

955   Genotype Error Rates in Sequence Data. *Am. J. Hum. Genet.* **93**, 840–851 (2013).

956   99. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-

957   connected communities. *Sci. Rep.* **9**, 5233 (2019).

958   100. Weissensteiner, H. *et al.* HaploGrep 2: mitochondrial haplogroup classification in the

959   era of high-throughput sequencing. *Nucleic Acids Res* vol. 44 W58-63 (2016).

960

961

962

963   **Methods**

964   **Data Generation**

965   All steps of generating the ancient genomes were undertaken using well established protocols

966   within ancient DNA, as described in Supplementary Note S1. In brief, samples were

967   inspected visually to identify suitable candidates. Teeth and petrous bone have previously

968   been identified as the optimal skeletal material for recovery of ancient DNA and were hence

969   selected where possible. All drilling, extractions and library builds were carried out in

970   dedicated ancient DNA clean lab facilities at the University of Copenhagen, both manually

971   and with automatisation. Double and single stranded libraries[89–91] were built and sequenced

972   on the Illumina Hiseq 4000 and the Novaseq 6000. For the purpose of authenticating ancient

973   reads as human, non-USER treated libraries were built to allow for the detection of post-

974   mortem damage[92]. Where possible, USER treated libraries were built from the authenticated

975   extracts to minimise the effects of post-mortem damage on downstream analysis[93]. All reads

976   were mapped to the human reference genomes (build hs37d5) using bwa aln (0.7.17)[94], and

977   duplicates were removed using picard MarkDuplicates (2.25.0). Contaminated libraries were

978   identified using contamix[95] and excluded. The final 712 samples included in this study ranged

979   from 0.01X to 32.83X in autosomal coverage.

980

981   **Imputation**

982   A set of 4,009 published ancient individuals with coverages suitable for imputation[28,37] was

983   curated, including whole genomes and SNP capture data. The published data were merged

984   with the 712 newly generated genomes and imputed with GLIMPSE v.1.1.1[96]

985   (Supplementary Note S4). As off-target capture sites have been shown to impute poorly, only

986   on-target sites were considered. In addition, sites that imputed poorly (imputation info score

987   < 0.5) or did not pass the 1000 Genomes strict mappability mask were excluded. Individuals

988   with coverage lower than recommended (0.1–0.5X for whole genome, 1.0X for captured

989   samples[28,37]) and average genotype probability resulting from the imputation (<0.9) were also

990   excluded. First and second degree relatives were identified using ngsRelate[97] (Supplementary

991   Note S5.6), and were then excluded from later analyses. This resulted in a final set of 4517

992   individuals (including 578 new) covering 697,179 SNPs.

993

IBD based analyses

Segments of identity-by-descent (IBD) shared between all individuals in the final imputed set of 4517 individuals were detected using IBDseq[98] (Supplementary Note S4). Potentially spurious IBD segments with low LOD scores or in hotspot regions were removed as described in Allentoft et al.[28].

To cluster ancient samples, a network-based hierarchical clustering[99] based on total shared IBD lengths was performed, as described in Allentoft et al.[28]. As short segments are less informative on recent ancestry, single segments less than 2 cM in length or total shared segment lengths of less than 5cM were excluded. Clusters were then curated as described in Allentoft et al.[28]. For the deeper clusters, individuals were plotted on a map of western Eurasia, showing clear correspondences through time and space. To better visualise IBD clusters and finer-scale patterns in the data, we carried out PCA analysis using GCTA (v1.94.1) showing the out-of-Africa and western Eurasian diversity and highlighted positions of individuals from the IBD clusters within PCA space. A second clustering run was performed containing only samples from the Bronze Age and earlier (2800+ BP), as later admixed individuals can bias clustering of the Bronze Age structure (Supplementary Note S5.2.2).

To understand the relationship between and diversity within the clustering, we performed IBD mixture modelling as described in Allentoft et al.[28], allowing us to model target individuals as mixtures of source clusters. For every individual in the dataset, a palette is created, formed by the normalised proportion of total IBD lengths that each individual shares with all 386 clusters in the dataset. As such, the different ways in which the target and source individuals are related to non-source or target individuals are leveraged when modelling the target palette with the source palettes. Here, the single segment length cutoff was 1 cM and no total shared segment length cutoff was applied, as deeper relationships between samples

1019   were of more relevance than to clustering. We first replicated previous results for the

1020   formation of European populations before systematically including more proximal sources in

1021   a series of sets to reveal insight into the genetic structure of the ancient individuals. We then

1022   plotted the mixture modelling proportions on the PCA showing western Eurasian diversity to

1023   identify a series of novel admixture clines within the well established PCA structure

1024   (Supplementary Note S5.3.3). For some specific questions, auxiliary source sets were used

1025   (Supplementary Note S5.3.2.). To understand how the results of the IBD mixture modelling

1026   were reflected on a regional level, we applied spatio-temporal ordinary kriging [39], as detailed

1027   in Supplementary Note S5.3.8. Spatiotemporal variograms were fit via a metric covariance

1028   model using the R package gstat. We used two sets of parameters, one similar to the original

1029   publication for ancestries relevant to Europe more broadly (IBD mixture modelling sets 1-4),

1030   and a second allowing for the closer proximity between populations of relevance during the

1031   Iron Age (Set 5-7) .

1032

Admixture dating

1034   To assess the time of formation of admixed populations detected through the mixture

1035   modelling, DATES[58] was run using default settings (Supplementary Note S5.5). A mean

1036   generation time of 25 years, and the midpoint of the mean radiocarbon ages of all dated

1037   individuals in the target population were used.

1038

Y-chromosome and mitochondrial genomes analyses

1040   Genetic sex and the presence of karyotypes 47,XXY and 47,XYY were determined from the

1041   ratios of reads mapping to the autosomes, X, and Y chromosomes (Supplementary Note

1042   S5.4.3). For individuals with an XY karyotype, we used bcftools v1.17 mpileup to call

1043   genotypes for positions located in the single-copy short-read callable 10Mb region of the Y-

1044   chromosome, excluding indels, triallelic locus and variants with less than 95% frequency in

1045   the population of non-clonal reads in a locus. Haplogroup paths were called conservatively by

1046   matching ancestral and derived calls from ISOGG 2019-2020 in a root-to-tip direction

1047   (Supplementary Note S5.4.2). For reads that mapped to rCRS mitochondrial reference

1048   genome, a consensus was called and the haplogroups classified using Haplogrep 2.0[100] for all

1049   samples with over 5X mitochondrial coverage. Sequences were then aligned with mafft, and

1050   a phylogenetic tree was generated using raxML (Supplementary Note S5.4.1).

1051

1052   **Dendrochronology**

1053   We used dendrochronological data based on tree-ring measurements from 654 wood samples

1054   of *Quercus* sp. from 42 sample sites in Denmark covering the years between AD 300 and AD

1055   800 (Supplementary Note S6.2). From the dendrochronological data we constructed an

1056   average growth-curve for Denmark within this time interval by use of TSAP-Win^{TM} (version

1057   4.82b2).

1058   **Pollen data in southern Sweden**

1059   We used all available pollen data from the province of Scania to investigate the

1060   anthropogenic vegetation changes before the Viking age (Supplementary Note S6.2). The

1061   pollen records were grouped into four characteristic biogeographical regions, standardised to

1062   include the 87 most common pollen/plant taxa and the period 4000–900 BP. We applied a

1063   simplified indicator-species approach using eight pollen taxa from which changes through

1064   time are expected to be mainly due to land-use change. Pollen proportions for these pollen

1065   indicators were then summarised in 100-year intervals.

1066   **Data availability**

1067   Sequence data were deposited in the ENA under accession: xxxxxxxx

1068

Methods References

1070   References 89 - 100

1071

Acknowledgements

1086

Author contributions

1088   EW initiated the study. HMC, EW led the study. HMC, GK, TW, RN, LH, KKr, MS, EW

1089   conceptualised the study. LV, TW, LO, LH, KKr, MS, EW supervised the research. GS,

1090   PBD, SS, AS, CPEZ, TW, RN, EW acquired funding for research. HMC, GS, PE, FD, MLSJ,

1091   SB, PBD, MEA, KGS, AF, TA, IAk, SLA, BA, PB, NC, PC, EC, TTC, ACp, SDD, SD, JD,

1092   KMF, OG, UBG, ZTG, JH, SH, MHe, VH, MHø, MHo, RJ, MDJ, JWJ, TJ, OTK, AK, RK,

1093   KKj, LK, KK, ACL, TL, ML, NL, YM, VYM, KM, VMa, ALM, IM, VMo, EMo, NM, EMu,

1094   BHN, DPK, GP, MSPdL, HR, WR, AS, BS, VSl, VSa, LS, OCU, HV, JV, DV, AVo, SW,

1095   HWe, AW, HWi, KW, JZ, BC, JPD, LO, RN were involved in sample collection. HMC,

1096   MLSJ, PBD, AF, EVB, PC, EC, RJ, TJ, RK, LK, KK, NL, VMa, VMo, EMo, EMu, BHN,

1097   DPK, GP, MSPdL, WR, VSa, AVo, SW, HWi, KW, JZ curated bioarchaeological data.

1098   HMC, FVS, GS, PBD, MEA, LV, CG, JC, JaS, JeS, FEY, ML, SBM, BSP, MSPdL, SW,

1099   TSK, LO were involved in data generation. HMC, GK, JVMM, FVS, TP, TV, SMS, PBD,

1100   ARa, IAl, RAH, EKIP, WB, AI, ARo, AVa, ML, CPEZ, TSK, BC, JPD, RN, MS were

1101   involved in data analysis. HMC, GK, PE, LH, KKr, EW drafted the main text. HMC, JVMM,

1102   FVS, TP, TV, SMS, PE, RF, MJG, HMEL, MFM, MLSJ, SB, ARa, KGS, SS, AF, IAk, SLA,

1103   EVB, EC, MDC, UBG, JZG, ZTG, SH, MHe, MHø, RJ, MDJ, JWJ, TJ, OTK, KKj, VK, LK,

1104   KK, ACL, ML, VYM, KM, VMa, ALM, VMo, EMo, NM, EMu, DPK, HR, WR, AS, BS,

1105   VSa, LS, GT, HV, JV, AVo, SW, HWe, AW, HWi, KW, JZ, TSK, LH drafted suppl notes

1106   and materials. HMC, GK, JVMM, FVS, GS, TP, TV, SMS, PE, RF, MJG, HMEL, MFM,

1107   FD, MLSJ, SB, MEA, LV, CG, ARa, IAl, RAH, EKIP, KGS, SS, AF, WB, AI, ARo, JC,

1108   FEY, SLA, MGB, JD, JZG, ZTG, VH, MHø, MHo, RJ, HJ, MDJ, JWJ, OTK, KKj, VK, LK,

1109   ML, NL, KM, ALM, NM, EMu, BHN, DPK, HR, AS, VSa, LS, GT, HV, SW, AW, KW, JZ,

1110   CPEZ, TSK, TW, LO, RN, LH, KKr, MS, EW involved in reviewing and editing drafts.

1111

Ethics declarations

1113   The authors declare no competing interests

1114

Additional Information

1116   Supplementary Information is available for this paper.

1117

1118   Correspondence and requests for materials should be addressed to: ewillerslev@sund.ku.dk,

1119   hugh.mccoll@sund.ku.dk, guus@hum.ku.dk, kristian.kristiansen@archaeology.gu.se

1120

1121   Reprints and permissions information is available at www.nature.com/reprints.

1122

1123

1124

1125   **Extended Data Figures**

1126

1127
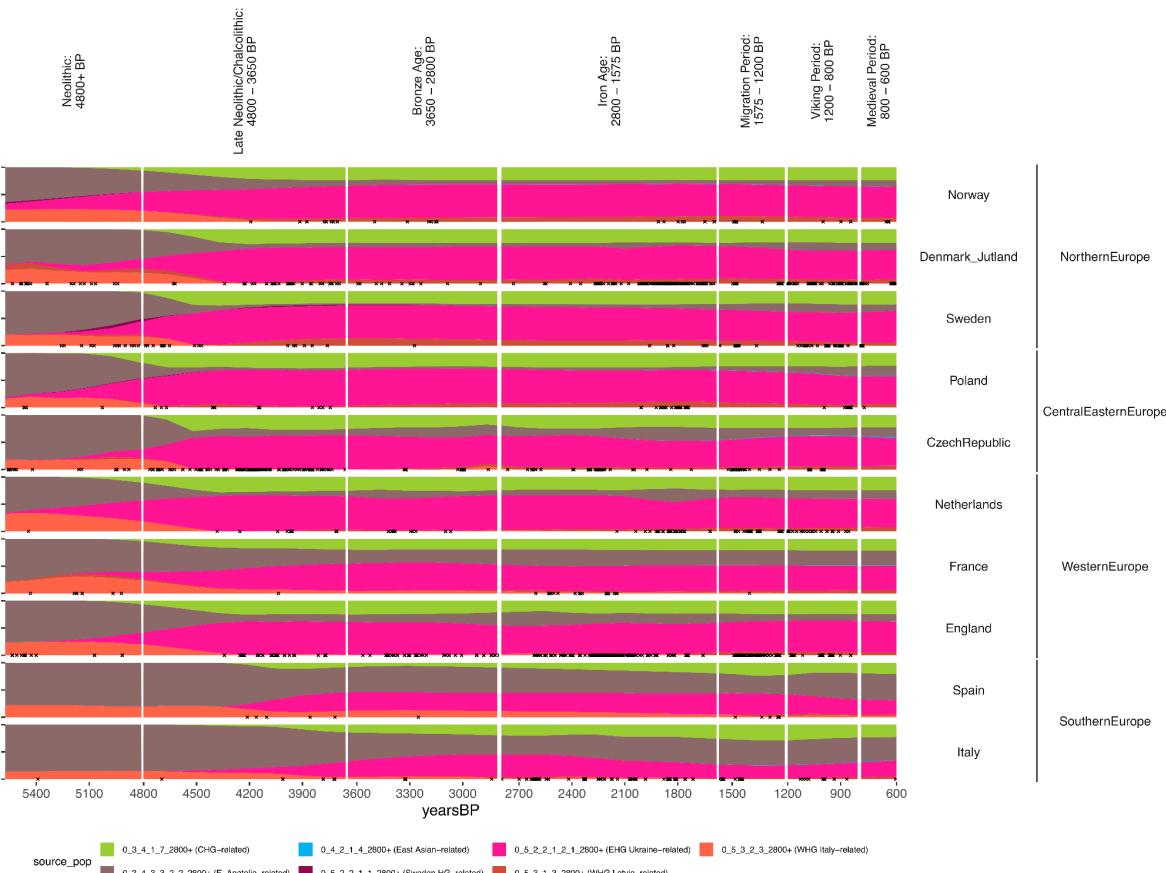
1128

1129   **Extended Data Fig. 1. Spatiotemporal Kriging results showing the general trends of the**

1130   **proportion of ancestries using Hunter Gatherer and Early Neolithic Farming Sources**

1131   **(Set 1).** The arrival of Steppe Ancestry, modelled as EHG (pink) and CHG (light green), onto

1132   a backdrop of Farming related ancestry, modelled as Early Anatolian Farming (brown) and

1133   Western Hunter-Gatherer (orange) is apparent. The 'Europe-wide' kriging parameters are

1134   applied here. Individuals within 250 km of the point used for each country (Supplementary

1135   Fig. S5.134) are indicated with an 'x'. While these results represent the general trends, they

1136   do not capture the heterogeneity present in many regions, and hence should be interpreted

1137   together with the full mixture modelling results (Supplementary Fig. S5.18).

1138

1139
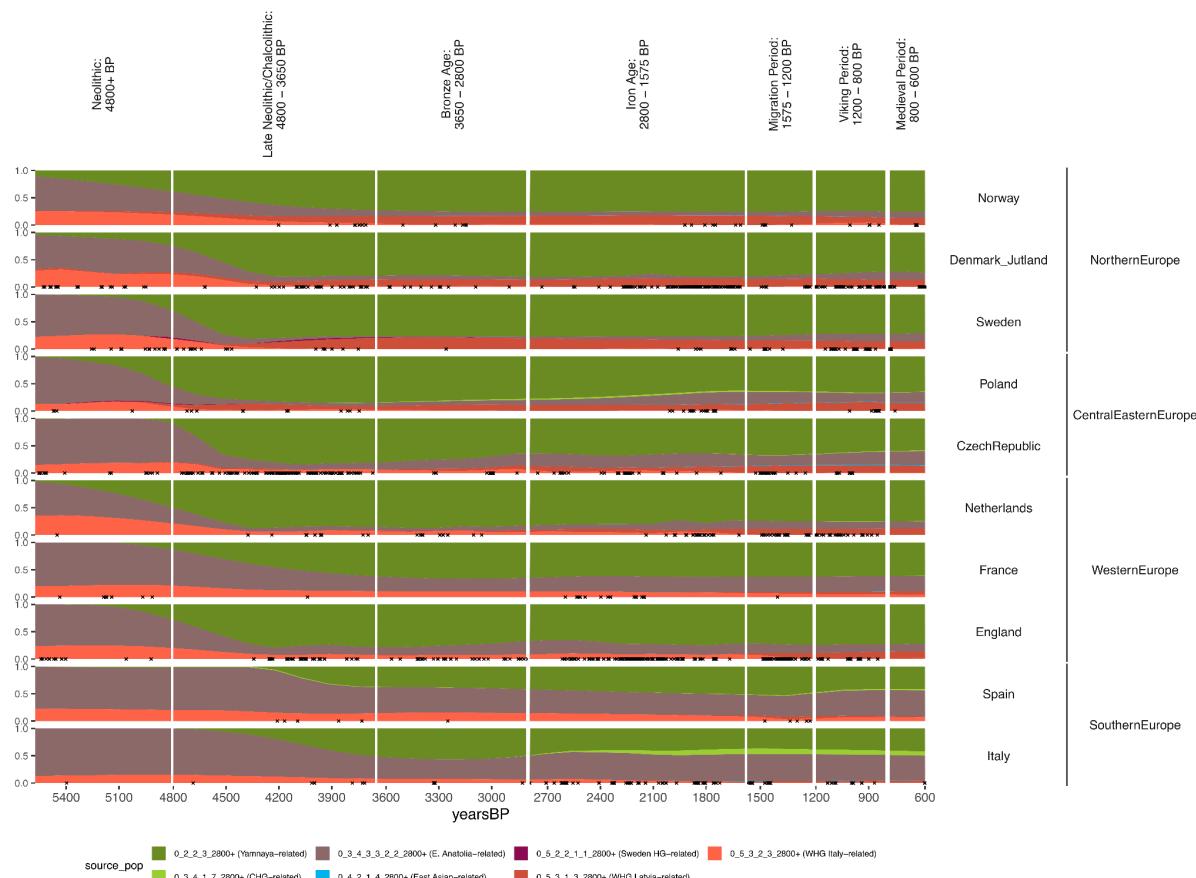
1140   **Extended Data Fig. 2. Spatiotemporal Kriging results showing the general trends of the**

1141   **proportion of ancestries when adding a 'Yamnaya-related' source (Set 2).** Here, Steppe

1142   Ancestry that was previously modelled as EHG (pink) and CHG (light green), is now

1143   modelled directly (darker green). The 'Europe-wide' kriging parameters are applied here.

1144   Individuals within 250 km of the point used for each country (Supplementary Fig. S5.134)

1145   are indicated with an 'x'. While these results represent the general trends, they do not capture

1146   the heterogeneity present in many regions, and hence should be interpreted together with the

1147   full mixture modelling results (Supplementary Fig. S5.18).

1148
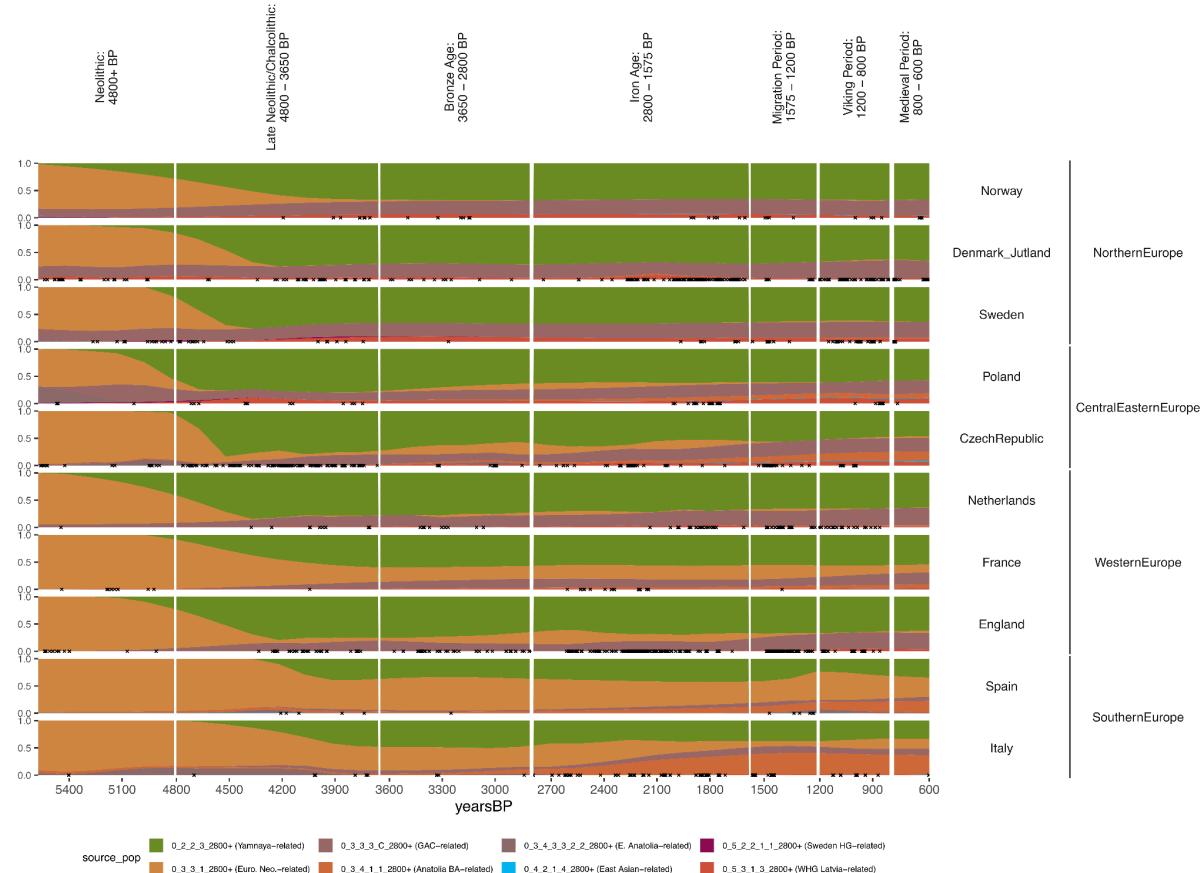
1149

1150   **Extended Data Fig. 3. Spatiotemporal Kriging results showing the general trends of the**

1151   **proportion of ancestries of when including additional Farming-related sources (Set 3).**

1152   To the north, Bronze Age populations are modelled only as Yamnaya and Globular Amphora,

1153   whereas in the south, some European Farming ancestry persists. The 'Europe-wide' kriging

1154   parameters are applied here. Individuals within 250 km of the point used for each country

1155   (Supplementary Fig. S5.134) are indicated with an 'x'. While these results represent the

1156   general trends, they do not capture the heterogeneity present in many regions, and hence

1157   should be interpreted together with the full mixture modelling results (Supplementary Fig.

1158   S5.18).

1159

1160

1161
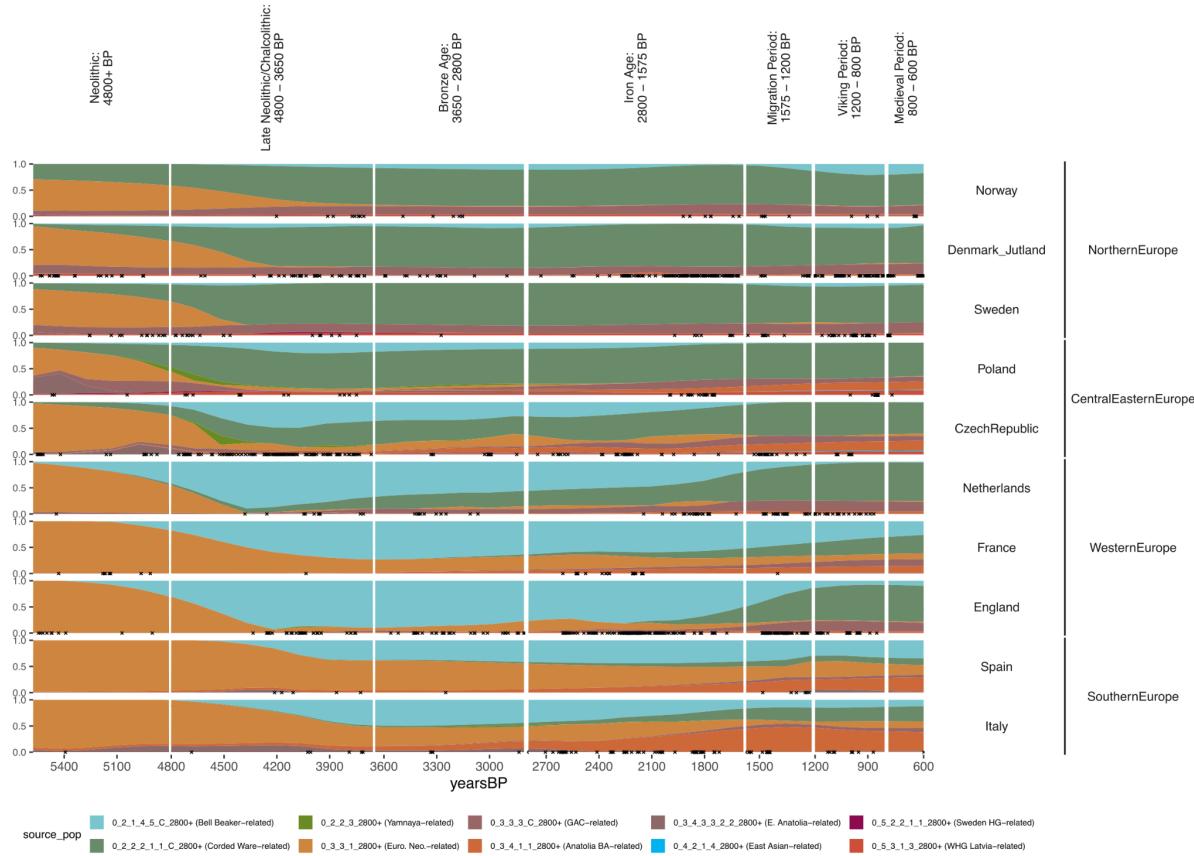
1162

Extended Data Fig. 4. Spatiotemporal Kriging results showing the general trends of the proportion of ancestries when including additional Corded Ware and Bell Beaker-related sources (Set 4). Steppe Ancestry from the Bronze Age populations are modelled only as Corded Ware-related to the North, and Bell Beaker-related in the south. The expansion of Bronze Age Anatolian-related ancestry in Italy during the Iron Age, and to lesser degrees Czech Republic, France and England is apparent. The 'Europe-wide' kriging parameters are applied here. Individuals within 250 km of the point used for each country (Supplementary Fig. S5.134) are indicated with an 'x'. While these results represent the general trends, they do not capture the heterogeneity present in many regions, and hence should be interpreted together with the full mixture modelling results (Supplementary Fig. S5.18).

1173

1174
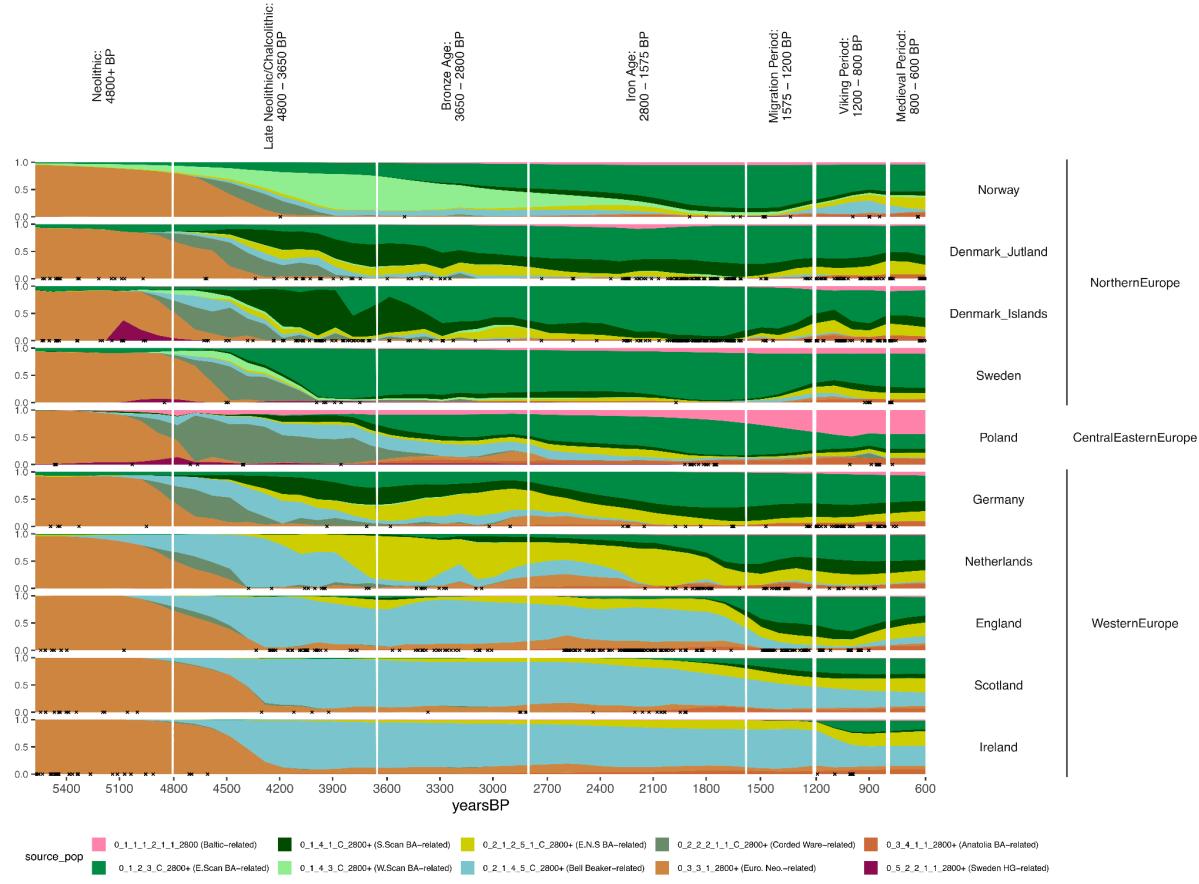
1175

Extended Data Fig. 5. Spatiotemporal Kriging results showing the general trends of the proportion of ancestries when including additional Bronze Age Scandinavian and Baltic-related sources (Set 5). In the Late Neolithic / Chalcolithic, the regions of Northern Europe are distinct and modelled by various sources. Throughout the Bronze Age and Iron Age, Eastern Scandinavian ancestry becomes widespread, despite the persistence of some local Bronze Age ancestry. During the Migration Period, individuals from England and the Netherland experience an increase in Scandinavian ancestry reflecting Iron Age Jutland. Individuals from Denmark and Sweden experience and increase in East North Sea ancestry. Norway experiences an increase in Bell Beaker-related ancestry during the Viking Period. The transition in Poland from Iron Age (East Germanic) to Medieval (Slavic) coincides with the arrival of 'Baltic BA'-related ancestry. The 'North / West Europe' kriging parameters are

1187   applied here. Individuals within 200 km of the point used for each country (Supplementary

1188   Fig. S5.134) are indicated with an 'x'. While these results represent the general trends, they

1189   do not capture the heterogeneity present in many regions, and hence should be interpreted

1190   together with the full mixture modelling results (Supplementary Fig. S5.18).

1191
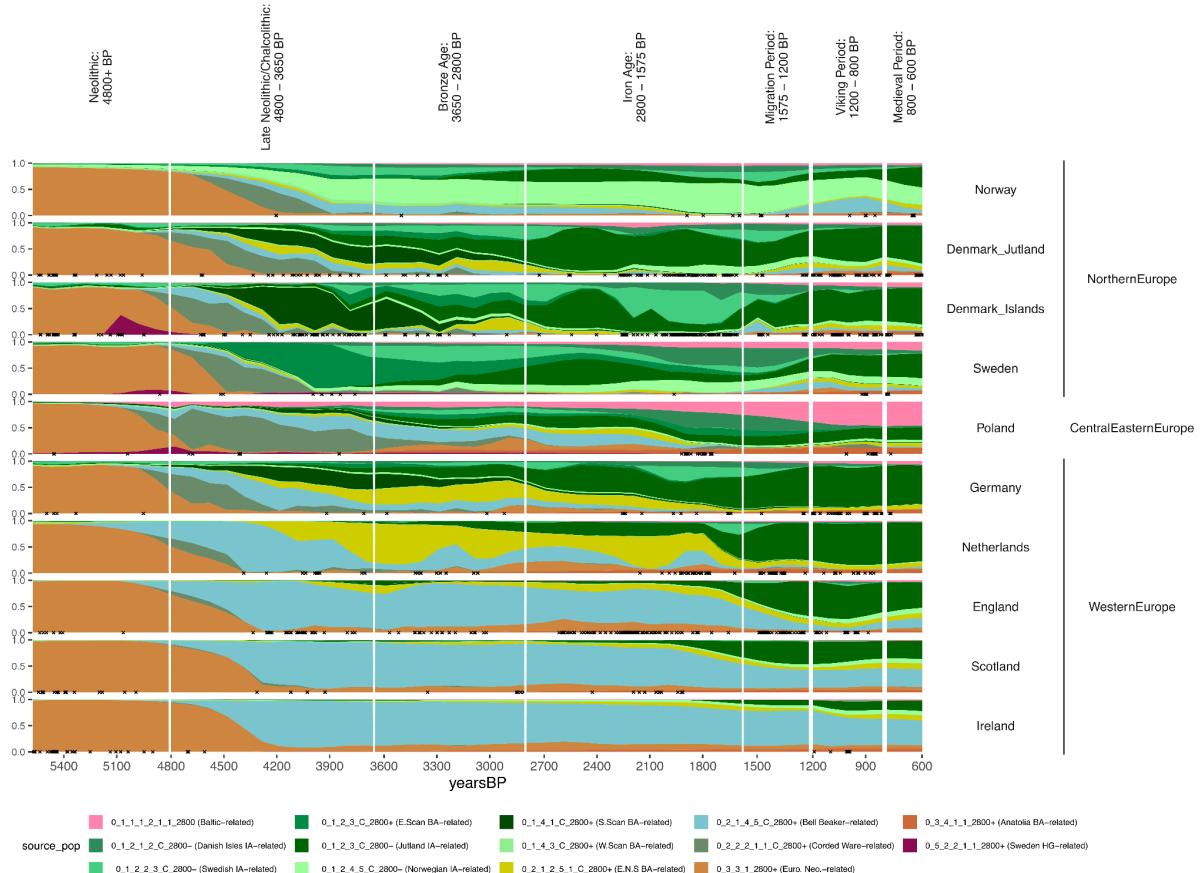
1192

Extended Data Fig. 6. Spatiotemporal Kriging results showing the general trends of the

proportion of ancestries when including Iron Age sources from Norway, Sweden, the

Danish Isles and Jutland (Set 6). Here, the migration period shows an increase of 'Jutland

IA'-related ancestry in England, the Netherlands, the Danish Isles. The 'North / West Europe'

kriging parameters are applied here. Individuals within 200 km of the point used for each

country (Supplementary Fig. S5.134) are indicated with an 'x'. While these results represent

the general trends, they do not capture the heterogeneity present in many regions, and hence

1200   should be interpreted together with the full mixture modelling results (Supplementary Fig.

1201   S5.18).

1202
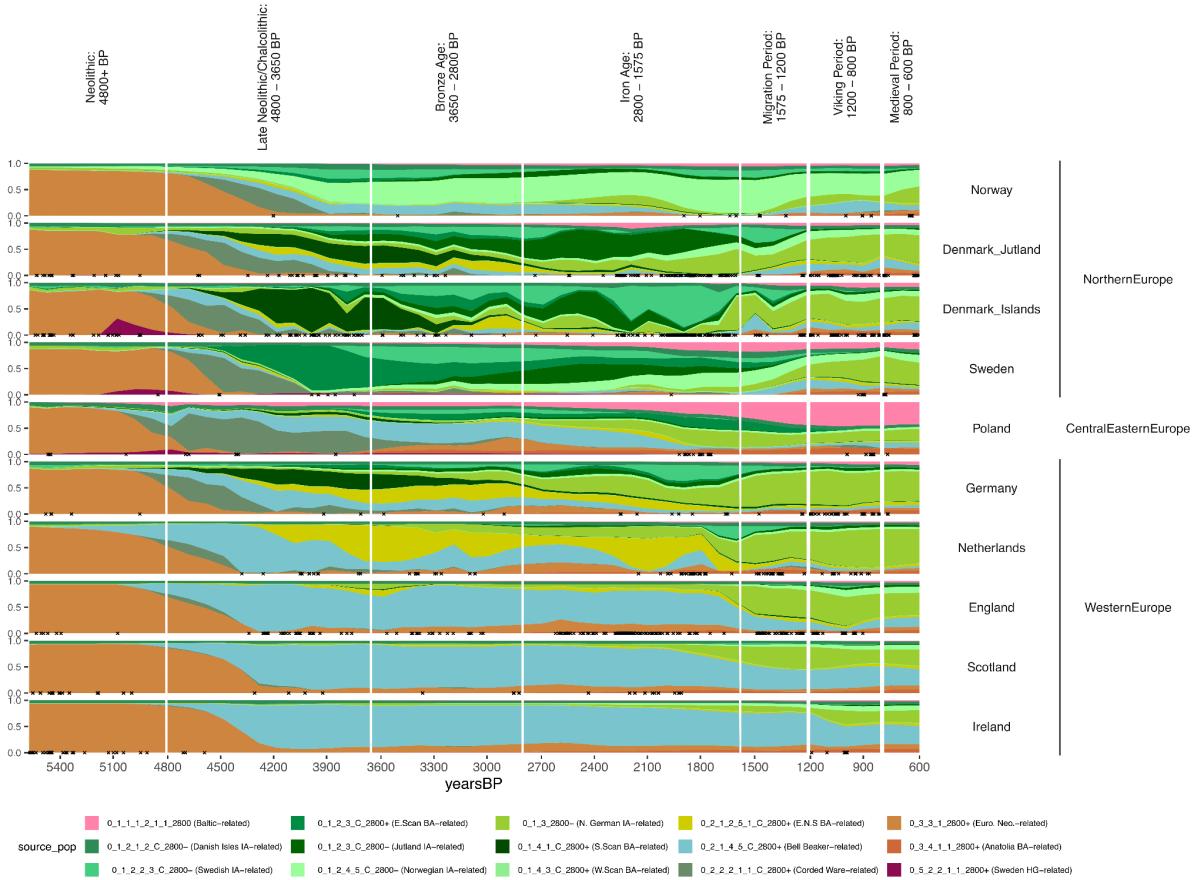
1203

Extended Data Fig. 7. Spatiotemporal Kriging results showing the general trends of the proportion of ancestries when including an additional Iron Age source from Northern Jutland (Set 7). Here, we show migration period migrations transitions mentioned in Set 6 are better modelled by the Northern/Germany Southern Jutland source. In addition, Jutland sees a transition from local to Northern Germany/Southern Jutland ancestry. Individuals within 200 km of the point used for each country (Supplementary Fig. S5.134) are indicated with an 'x'. While these results represent the general trends, they do not capture the heterogeneity present in many regions, and hence should be interpreted together with the full mixture modelling results (Supplementary Fig. S5.18).

1213

1214
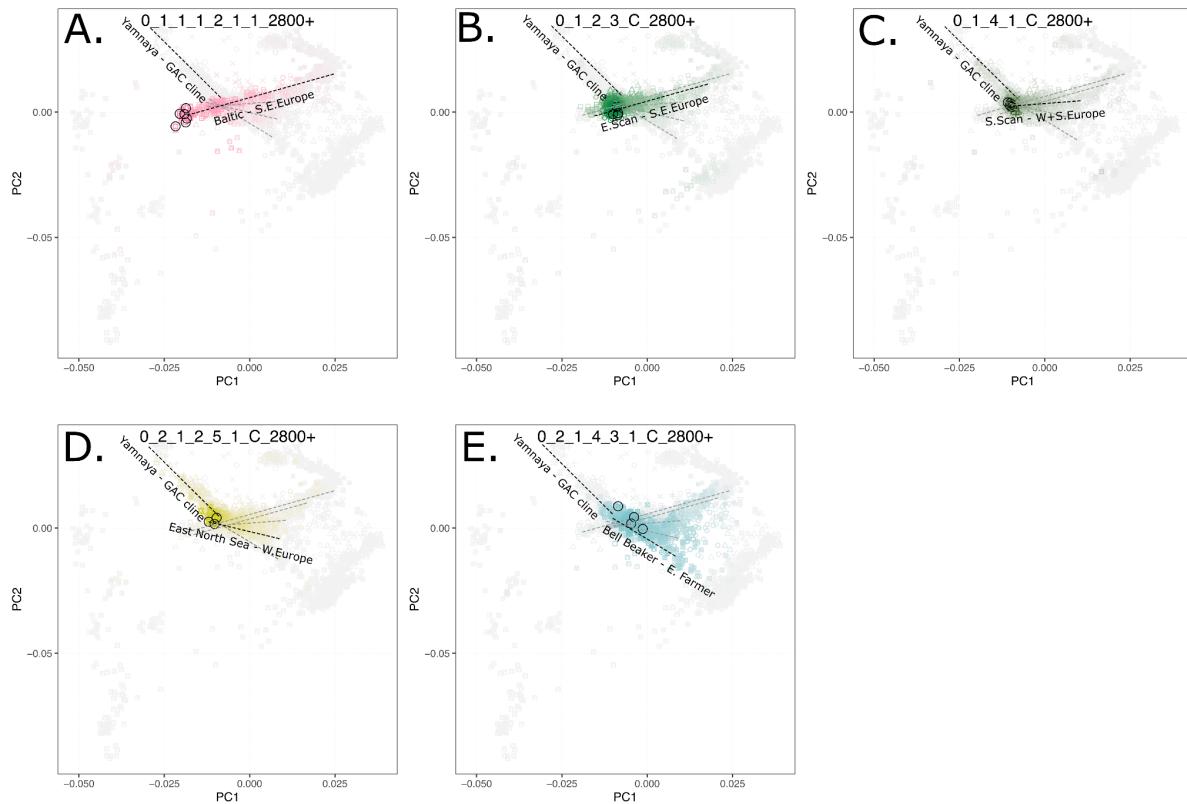
1215

Extended Data Fig.-8. A subset of mixture modelling results from Set 5 displayed on the

western Eurasian PCA, (Supplementary Note S5.3.1, Supplementary Fig. S5.20),

showing mixture modelling proportions for a series of subclusters from the Corded

Ware (North) (A, B, C) and Bell Beaker (D, E) clusters, revealing clines admixing with

groups of varying European Farmer- and Bronze Age western Mediterranean-related

ancestry. Source individuals are circled, and admixture proportions follow a cline from full

colour (100%) to grey (0%).

1223
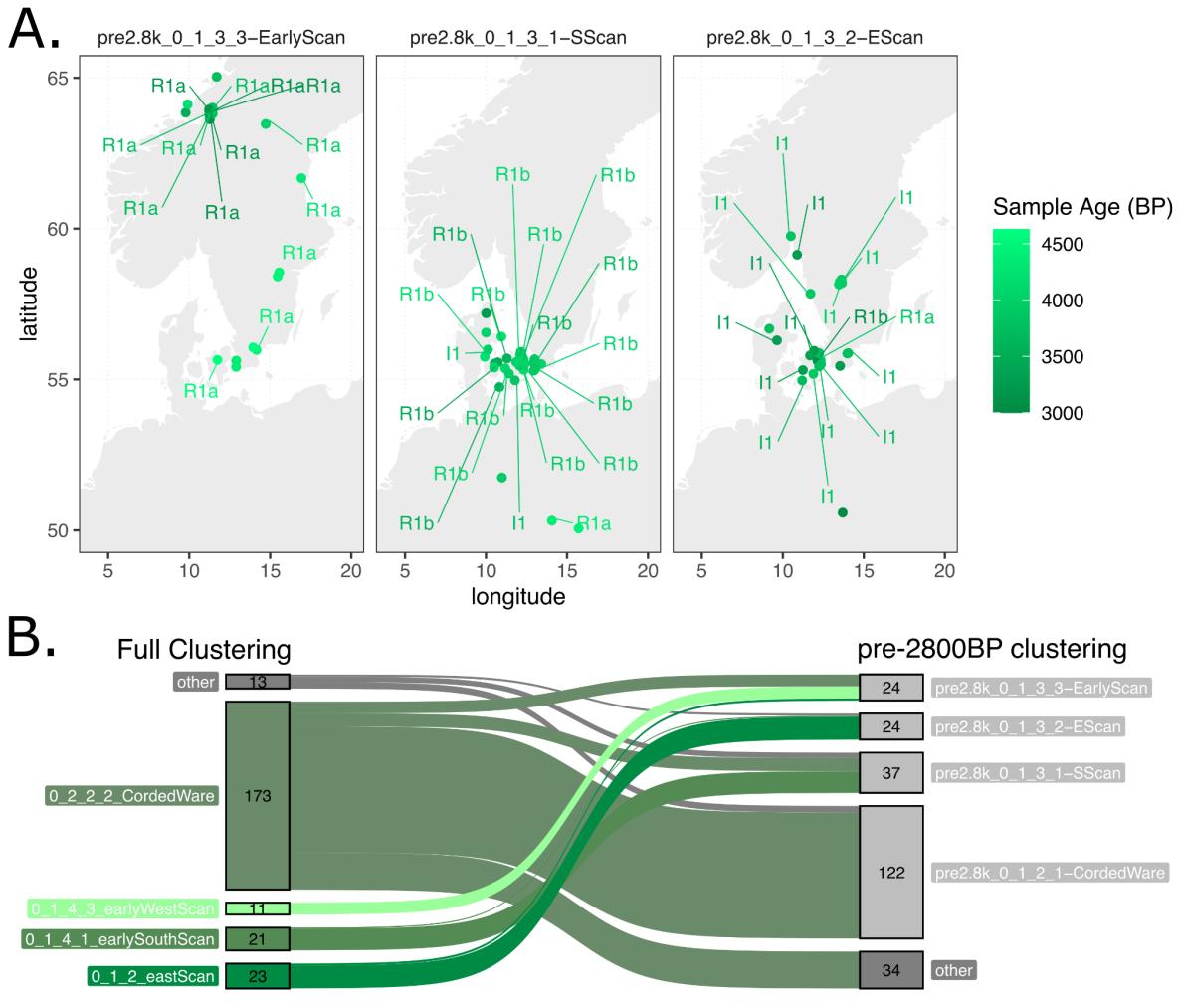
Extended Data Fig.-9. Scandinavian IBD clusters from pre-2800BP clustering. (A) Geographical distribution of individuals within the Scandinavian sub-clusters from the pre-2800 BP re-clustering. For males with sufficient coverage, major Y-haplogroups are noted. (B) Sankey diagram showing the correspondence between the three main Scandinavian clusters and the Eastern Corded Ware clusters in the Full and pre-2800 BP clustering.
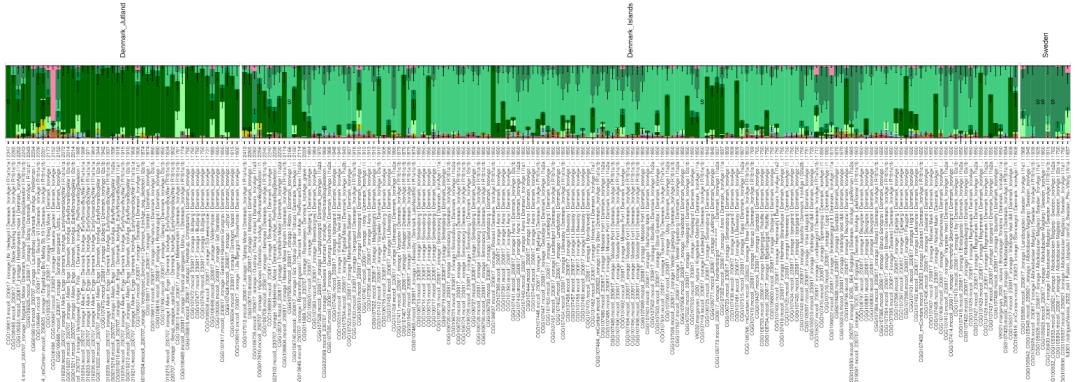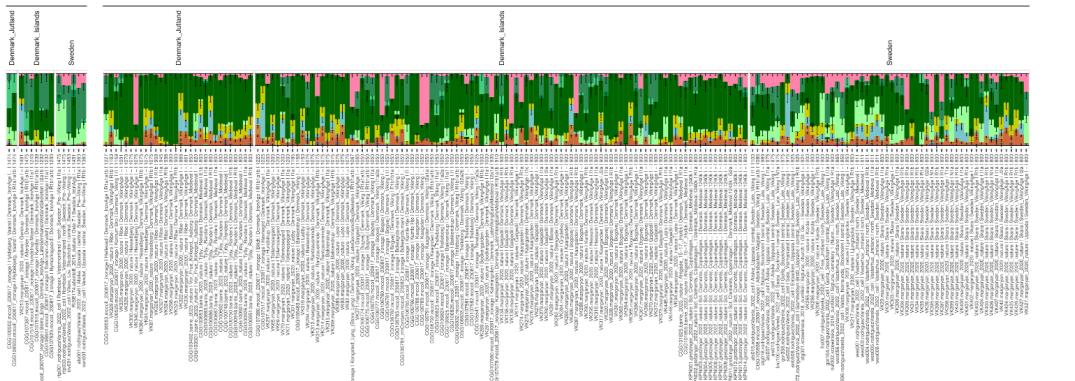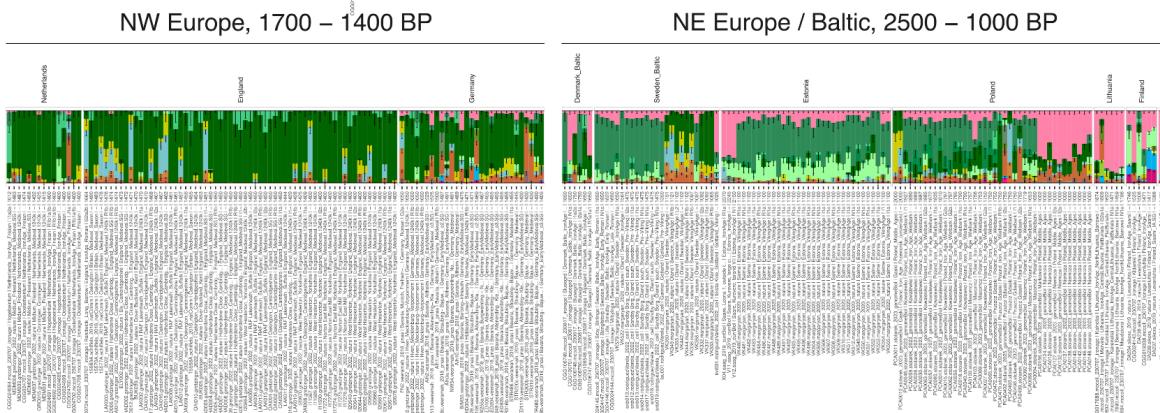
Extended Data Fig.-10. A northern European subset of IBD mixture modelling results for Iron Age sources (Set 6). Row 1 shows variation from Denmark_Jutland to the Islands of Denmark, to southern Sweden. Row 2 shows Denmark and Sweden during the Migration Period (1575–1200 BP, left) to the Viking Age (1200–800 BP, right). Row 3 shows the

1241   surrounding regions to the west (left) and east (right). Subset from Supplementary Note S5.3,

1242   Supplementary Fig. S5.18.

1243
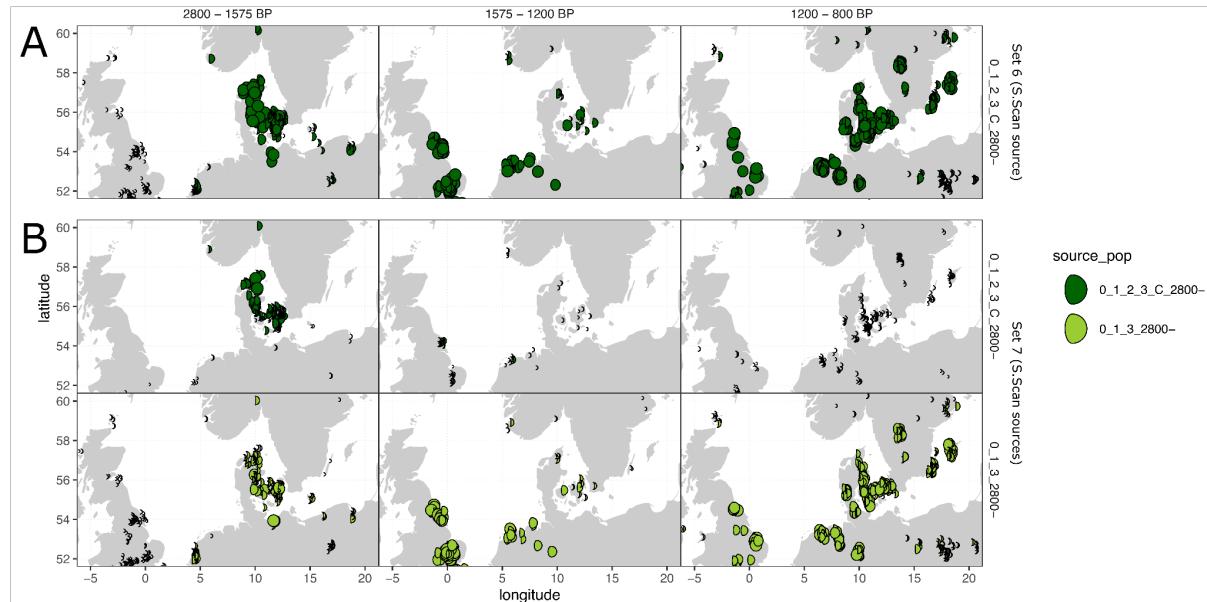
1244

1245   **Extended Data Fig.-11. A subset of IBD mixture modelling results showing the**

1246   **proportion of Southern Scandinavian Iron Age ancestry for northern Europe with**

1247   **varying source sets.** A) Set 6, which contains a single Southern Scandinavian Iron Age

1248   source (0_1_2_3_C_2800-, northern Jutland), in comparison to B) Set 7, with two Southern

1249   Scandinavian Iron Age sources (0_1_2_3_C_2800-, northern Jutland and 0_1_3_2800-

1250   Mecklenburg (northern Germany). The proportion of ancestry modelled is indicated by the

1251   proportion filled and the size of each circle. Full mixture modelling results for northern

1252   Europe are shown in Supplementary Fig. S5.25 and Supplementary Fig. S5.26

1253

1254

1255
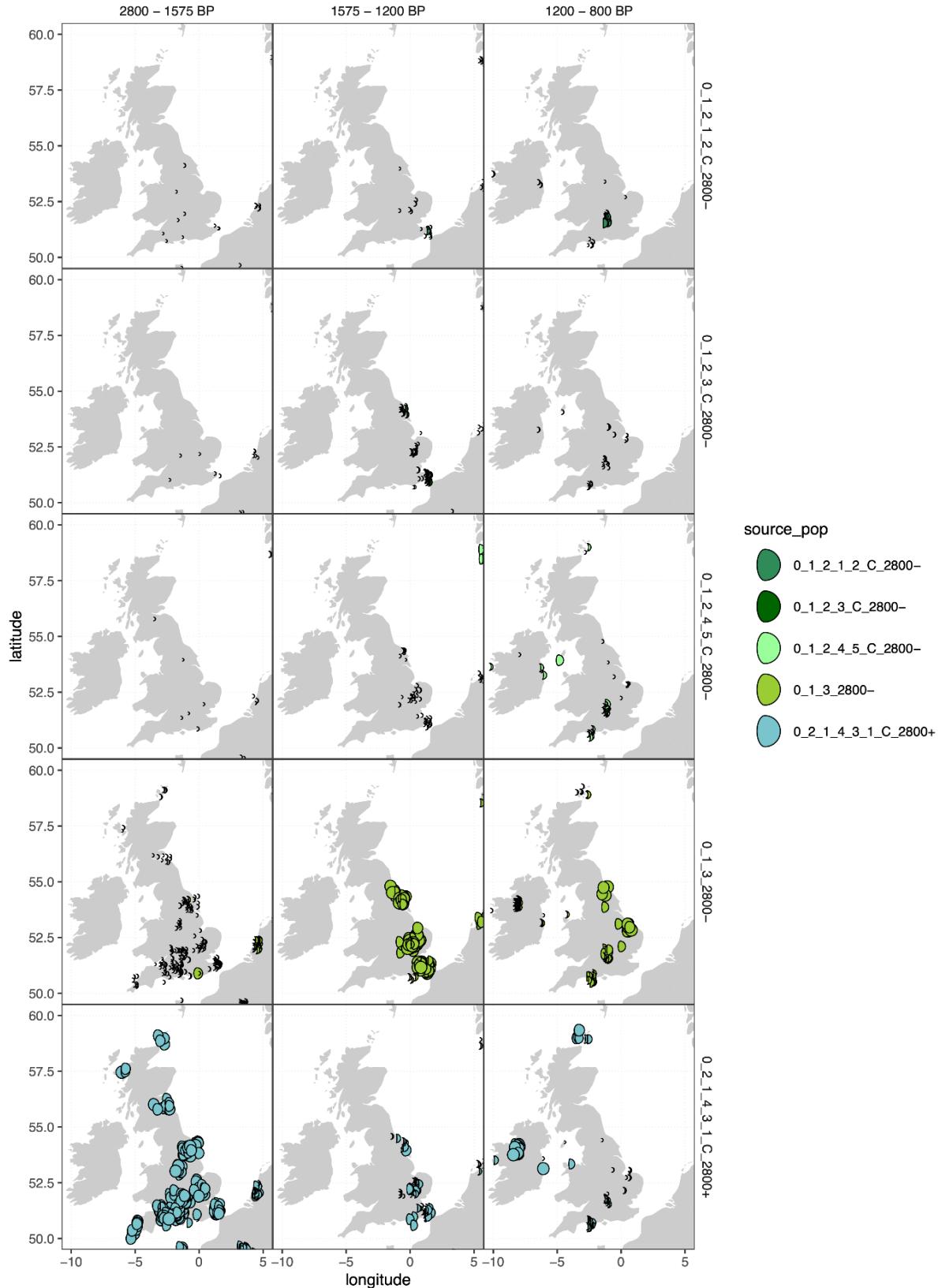
1256 **Extended Data Fig.-12. A subset of IBD mixture modelling results showing the**

1257 **proportion of ancestry in Britain and Ireland (Set 7).** In column 1 (2800–1575 BP), the

1258 dominant ancestry modelled is 0_2_1_1_2 Celtic Bronze Age. In column 2 (1575–1200 BP)

1259 during the Anglo-Saxon period, a transition causing individuals to be modelled primarily as

1260 0_1_3_2800- Southern Scandinavian Iron Age (Mecklenburg, northern Germany) has

1261 occurred, with small proportions of 0_1_2_3_C_2800- Southern Scandinavian Iron Age

1262 (northern Jutland, Denmark). In column 3 (1200–800 BP), the appearance of other

1263 Scandinavian ancestries (cluster 0_1_3_2_2_2 Eastern Scandinavian Iron Age (Sweden) and

1264 cluster 0_1_6_2 Western Scandinavian Iron Age (Norway)) is apparent during the Viking

1265 Age. Full mixture modelling results for northern Europe are shown in Fig. S5.26.

1266
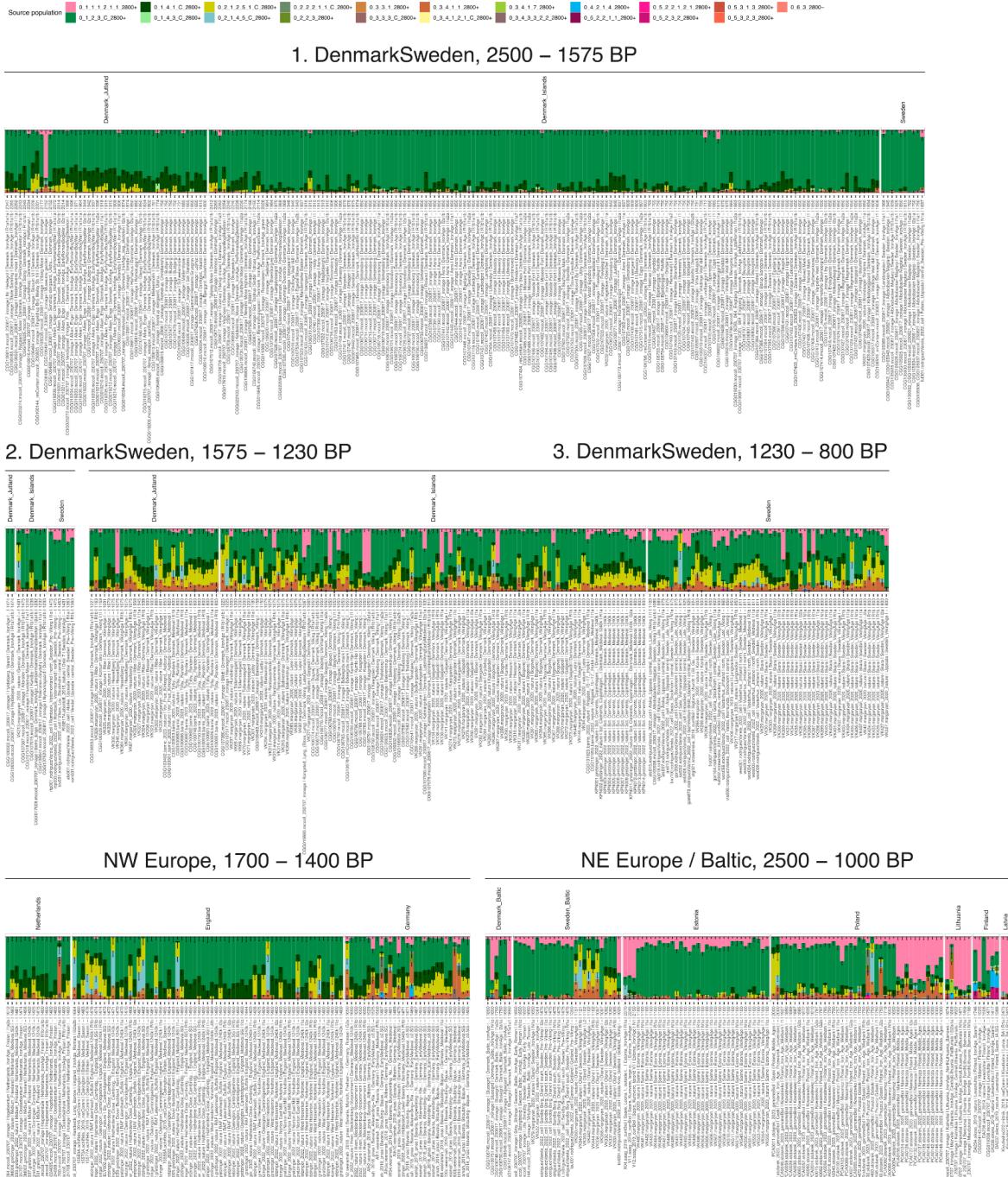
Extended Data Fig.-13. A northern European subset of IBD mixture modelling results for Bronze Age sources (Set 5). Row 1 shows the decreasing proportion of Southern Scandinavian ancestry from Denmark_Jutland to the Islands of Denmark, to southern

1273   Sweden. Row 2 shows Denmark and Sweden from the Migration Period (1575–1200 BP,

1274   left) to the Viking Age (1200–800 BP, right). Row 3 shows the surrounding regions to the

1275   west (left) and east (right). Subset from Supplementary Note S5.3, Fig. S5.18

1276

1277

1278

1279
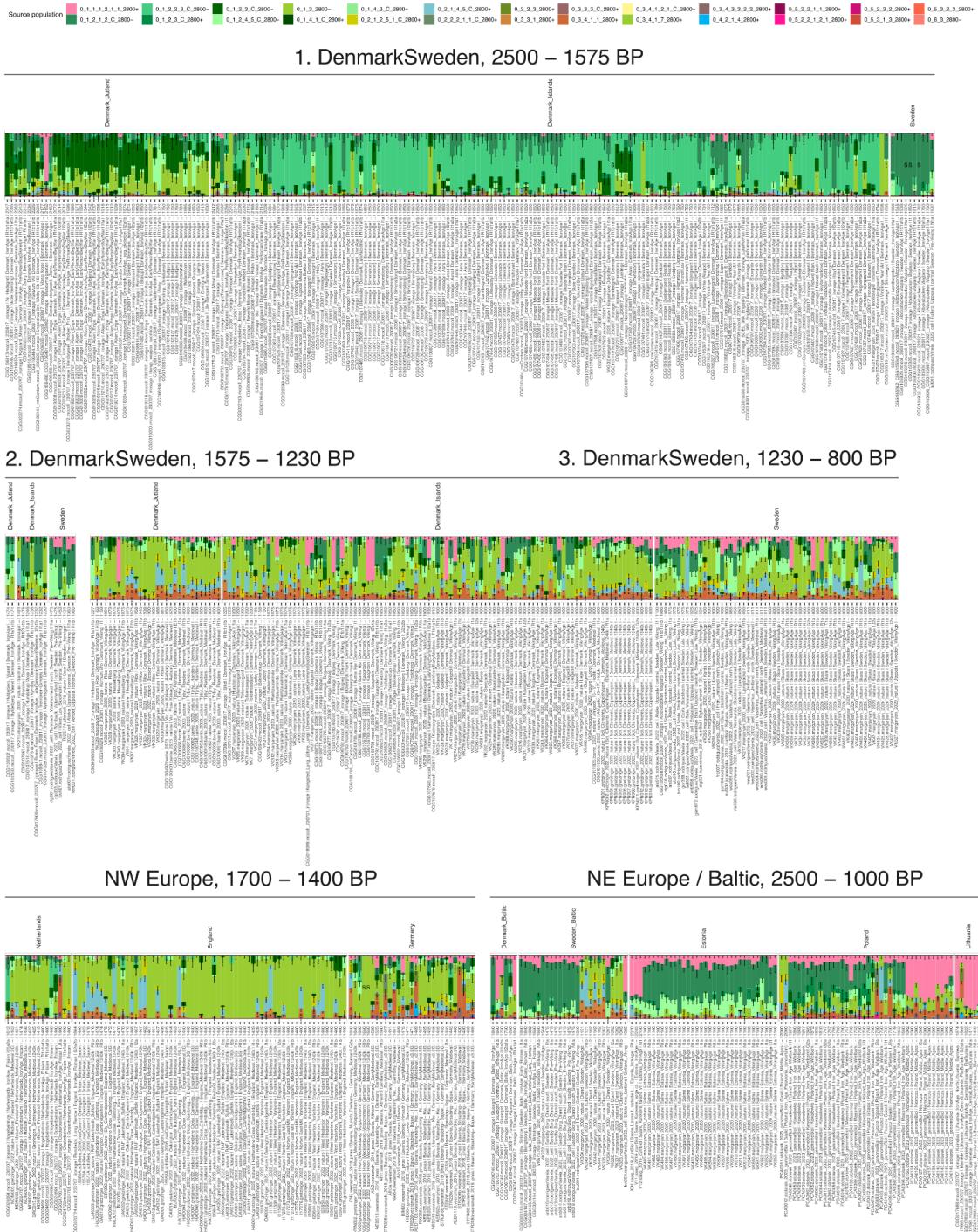
Extended Data Fig.-14. A northern European subset of IBD mixture modelling results for Iron Age sources, when including two Southern Scandinavian Iron Age sources (Set 7). Row 1 shows variation from Denmark_Jutland to the Islands of Denmark, to southern Sweden. Row 2 shows Denmark and Sweden during the Migration Period (1575–1200 BP,

1285   left) to the Viking Age (1200–800 BP, right). Row 3 shows the surrounding regions to the

1286   west (left) and east (right). Subset from Supplementary Note S5.3, Fig. S5.18.

1287

1288
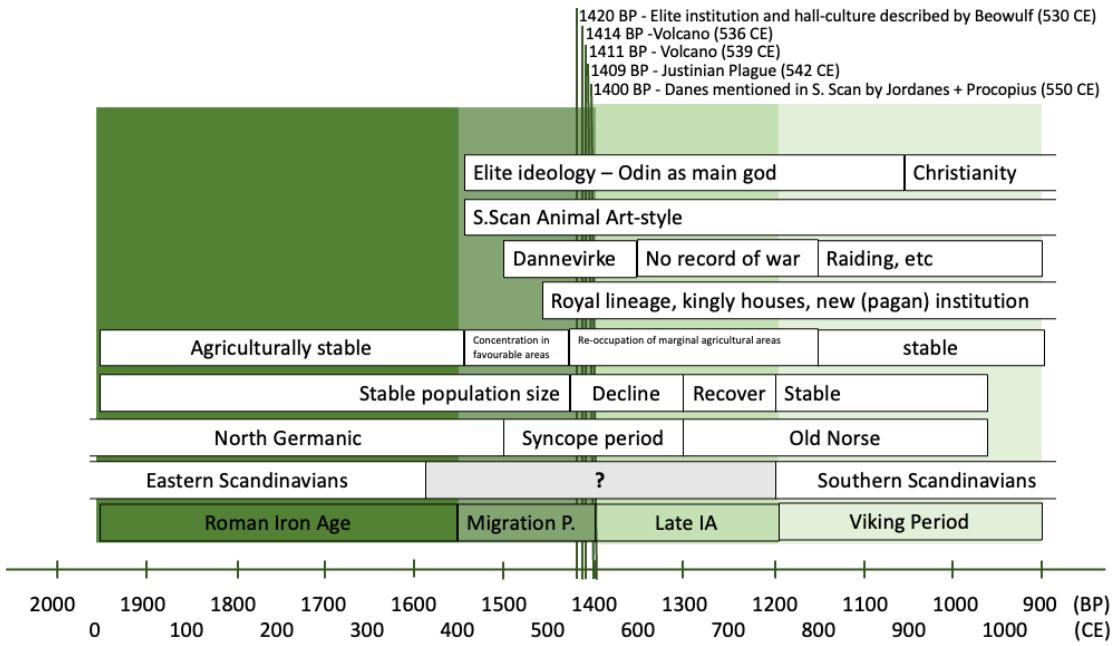
1289

Extended Data Fig.-15. Timeline showing climatic events, and social, demographic, linguistic and genetic shifts in the Danish Isles and southern Sweden from the Migration Period to the Viking Age.

1293

1294

1295

1296

1297