Title:

Ancient genomes from Iceland reveal the making of a human population

One Sentence Summary:

Ancient genomes from Iceland provide unique insights into the making of a human population shaped by extensive genetic drift.

Authors:

S. Sunna Ebenesersdóttir[1, 2], Marcela Sandoval-Velasco[3], Ellen D. Gunnarsdóttir[1, 2], Anuradha Jagadeesan[1, 2], Valdís B. Guðmundsdóttir[1, 2], Elísabet L. Thordardóttir[1, 2], Margrét S. Einarsdóttir[1, 2], Kristján H. S. Moore[1], Ásgeir Sigurðsson[1], Droplaug N. Magnúsdóttir[1], Hákon Jónsson[1], Steinunn Snorradóttir[1], Eivind Hovig[4,5,6], Pål Møller [4,7,8], Ingrid Kockum[9], Tomas Olsson[9], Lars Alfredsson[10], Thomas F. Hansen[11,12], Thomas Werge[11,13,14], Gianpiero L. Cavalleri[15], Edmund Gilbert[15], Carles Lalueza-Fox[16], Joe W. Walser III[17,18], Steinunn Kristjánsdóttir[17,18], Shyam Gopalakrishnan[3], Lilja Árnadóttir[17], Ólafur Þ. Magnússon[1], M. Thomas P. Gilbert[3], Kári Stefánsson[1,19] and Agnar Helgason[1, 2]

Affiliations:
[1]deCODE Genetics/AMGEN, Inc., Reykjavik Iceland.
[2]Department of Anthropology, University of Iceland, Reykjavik, Iceland.
[3]Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5–7, 1350 Copenhagen K, Denmark.
[4]Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway.
[5]Institute for Cancer Genetics and Informatics, Oslo University Hospital, Oslo, Norway.
[6]Department of Informatics, University of Oslo, Oslo, Norway.
[7]Department of Human Medicine, Universität Witten/Herdecke, Witten, Germany.
[8]Research Group Inherited Cancer, Department of Medical Genetics, Oslo University Hospital, Oslo, Norway.
[9]Center for Molecular Medicine, Department of Clinical Neuroscience, Neuroimmunology Unit, Karolinska Institutet, Stockholm, Sweden.
[10]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.
[11]Institute of Biological Psychiatry, Copenhagen Mental Health Services, Copenhagen, Denmark.
[12]Danish Headache Center, Dept. of Neurology, Copenhagen University hospital, DK-2600 Glostrup, Denmark.
[13]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark.
[14]The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Copenhagen, Denmark.
[15]Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, 123 St. Stephen's Green, Dublin, Ireland.
[16]Institut de Biologia Evolutiva, (UPF-CSIC), Barcelona, Spain.
[17]National Museum of Iceland, Reykjavik, Iceland.
[18]Department of Archaeology, University of Iceland, Reykjavik, Iceland.
[19] Norwegian University of Science and Techonology, University Museum, 7491 Trondheim, Norway
[20]Faculty of Medicine, University of Iceland, Reykjavik, Iceland.

Abstract:

Opportunities to directly study the founding of a human population and its subsequent evolutionary history are rare. Using ancient genomes from 27 skeletal remains, we demonstrate that the earliest Icelanders were a combination of Norse, Celtic and admixed individuals. We further show that ancient Icelanders are much more similar to their source populations in Scandinavia and the British Isles and Ireland than to contemporary Icelanders, who have been shaped by 1100 years of extensive genetic drift. Finally, we report evidence for unequal contributions from the ancient founders to the contemporary Icelandic gene pool, which was in part influenced by ancestry. These results provide unique insights into the making of a human population that has proven extraordinarily useful for the discovery of genotype-phenotype associations.

Main text:

Historical sources (*1*) indicate that Iceland was settled by people from Norway and the British Isles (defined here as the United Kingdom and Ireland) between 870-930 A.D. (*2, 3*). Throughout the preceding century, the Norse had raided, traded and settled in the islands and coastal regions of Ireland, Scotland and northern England (*4*). The number of settlers in Iceland is thought to have been in the range of 8,000-16,000, with the population rarely exceeded 50,000 until 1850 (*3*), after which there was a rapid expansion to its current size of 330,000.

Previous studies of mitochondrial DNA (mtDNA) and Y chromosomes from contemporary Icelanders indicate that while 62% of their matrilineal ancestry stems from Scotland and Ireland, 75% of their patrilineal ancestry is Scandinavian (*5-7*). Moreover, mtDNA control region sequences from 68 Icelandic ancient skeletal remains (~1,000 years old) were reported to be more similar to those observed in the contemporary source populations in Scandinavia

and the British Isles than those in contemporary Icelanders (*8*). This was explained by loss of mtDNA haplotypes from the Icelandic gene pool due to extensive genetic drift (*8*), an interpretation that is supported by founder events revealed through whole genome sequencing (WGS), genome-wide association studies (*9*) and population genetics analyses (*10*). Here, we use WGS data from the earliest generations of Icelanders to directly assess the formation and subsequent evolution of a human population.

We selected 35 skeletal remains from 27 geographically dispersed sites within Iceland for sequencing (Fig. 1). Based on archaeological context (typically grave artifacts) (*11, 12*) or radiocarbon dating (*13, 14*), 31 of the remains are deemed to have been from pre-Christian burials (<1000 A.D.), one from an early Christian burial before 1104 A.D. (ÞSK-A26), two with unknown dates (TGS-A1 and FOV-A1), and one (KOV-A2) born in 1678 (Table S1). Following established ancient DNA (aDNA) protocols (SM S2), 61 double-stranded DNA libraries were produced from 50 different DNA extracts, and Whole Genome Enrichment (WGE) (*15*) was used to enrich for human DNA fragments before sequencing. The libraries were assessed for pre-capture DNA concentration, the proportion of post-capture human mapped reads (endogenous content), contamination, and genome coverage, which led to the exclusion of eight individuals from pre-Christian burials from subsequent analyses (Fig. 1 and Table S2).

For the 27 individuals who passed quality control criteria, the average number of reads per position in the human autosomal reference genome obtained ranged from 0.18 to 30.7, with a median of 0.71 (Table 1). The authenticity of the mapped reads is supported by several different analyses (SM S2). First, all libraries yielded short read lengths (Table S3) and patterns of cytosine deamination that were consistent with aDNA (*16*) (Figs. S1 and S2). Second, contamination levels were negligible (<1.9%) (Table S4). Third, mtDNA consensus sequences for all 27 ancient individuals matched those obtained from a previous study (*8*) (Table S5).

Using R_y, the proportion of sex chromosome reads that mapped to the Y chromosome (*17*), it was possible to confidently assign sex to 26 of the individuals (Fig. S3A and Table S6). The result for one individual, YGS-B2, was ambiguous, as $R_y=0.055$ was much greater than the upper limit for females (0.016) and well below the lower limit for males (0.077). An examination of read depth on the sex chromosomes relative to the autosomes indicates that YGS-B2, who died before adulthood, carried two X chromosomes and one Y chromosome (Fig. S3B). To our knowledge, this is the first report of an individual with Klinefelter syndrome or any kind of aneuploidy based on aDNA. Of the 23 individuals from pre-Christian burials, 18 (78%) were male (X^2 test against expectation of 50%, P=0.012). As the individuals were not selected on the basis of morphologically predicted sex, it is likely that there were sex differences in burial practices during this period (*18*), such that female burials were either rarer or less likely to be identified and reported to the National Museum of Iceland as pre-Christian.

We next sought to establish the relationship between the ancient Icelanders and contemporary population samples genotyped on micro-array SNP chips (Table S7). As the read depth for most of the ancient Icelanders was too low to call diploid genotypes, haploid genotypes were used in all analyses (obtained by randomly sampling one allele from high quality reads). Based on previous genetic studies (*5-8, 10, 19, 20*), and historical, archaeological and linguistic evidence (*2, 12, 21*), the ancient Icelanders are expected to be most closely related to the contemporary populations from Iceland, Scandinavia and the British Isles. To validate this expectation, we performed a principal components analysis (PCA) based on 2,139 contemporary individuals from 28 European populations, using genotypes from 227,056 SNPs. When projected onto the first two principal components (PCs), all the ancient Icelanders are placed clearly within the clusters of contemporary individuals from Scandinavia and the British Isles (Fig. S4).

For a more detailed assessment of the relationships between the ancient Icelanders and their source populations, we performed a PCA based on genotypes from contemporary individuals from Scandinavia (n=3,118), the British Isles (n=1,436) and Iceland (n=916), using diploid genotypes from 404,066 autosomal SNPs. Figure 2A shows a clear separation of contemporary individuals from Scandinavia and the British Isles on PC2, reflecting allele frequency differences that have accumulated through several thousand years of drift, added to possible differences in contributions from the ancestral source populations of Europe (22-25). PC1 reveals an even more marked separation of contemporary Icelanders from their source populations. The average genetic distance (F_{ST}) (26) between Icelanders and the other populations in Fig. 2A is 0.00208, whereas comparable distances for the source populations range from a minimum of 0.00092 for England to a maximum of 0.00174 for Wales (Fig. S5).

The most likely cause for the divergence of Icelanders is 1100 years of genetic drift resulting from an initial founder event and subsequent small population size (8-10). This interpretation is supported by the PCA projection of the one early Christian and 23 pre-Christian ancient Icelanders, all of whom are placed on the plane of PC1 occupied by contemporary individuals from Scandinavia and the British Isles in Fig. 2A (see also Figs. S6 and S7). This also implies that the gene pools of these source populations have changed much less during the same period (8). Further support is provided by PCA projections of ancient individuals (n=16) from the British Isles dating from the Neolithic to the Anglo-Saxon period ($22, 27, 28$) (SM S4 and Table S8), who plot among contemporary British and Irish populations (Fig. S8). This justifies the use of contemporary individuals from these source populations as reasonable proxies for ancient inhabitants at the time Iceland was settled.

The PC2 axis in Fig. 2A provides information about the mixed ancestry of the ancient Icelanders (5-8). Thus, while some appear to be fully Scandinavian or British/Irish, others resemble a mixture of these ancestral groups. For a more explicit evaluation of ancestry of the

ancient Icelanders, contemporary genotyped individuals were grouped into Norse (*n*=2,139) (Norway and Sweden), Celtic (*n*=459) (defined here as Ireland and Scotland, excluding Orkney) and Icelandic reference populations. Fig. 2B shows the relationship between the ancient individuals and these two reference populations using D-statistics (*29*) in the form *D(Yoruba, X; Norse, Celtic)* (Table S9). To directly estimate the ancestry of the ancient Icelanders, we used ADMIXTURE in supervised mode with the Norse and Celtic reference populations (K=2) (Fig. 2C). The results were consistent with those obtained from the placement of the ancient Icelanders on the PC2 axis in the PCA in Fig. 2A and D-statistics in Fig. 2B (Pearson's |r|>0.98 between all three assessments of ancestry).

An assessment of Y chromosome variants among the 22 ancient Icelandic males revealed that all belonged to the haplogroups I1, R1a and R1b1 (Table 1). As I1 and R1a are common in Norse and R1b1 in Celts (*7, 30, 31*) (Table S10), we postulated and confirmed that there was a corresponding association between Y chromosome haplogroups and autosomal Norse/Celtic ancestry (Fig. 2C) in the 18 pre-Christian Icelanders (ANOVA p=0.02).

We estimated the mean Norse ancestry of the settlement population (23 pre-Christians and one early Christian) to be 0.554 (95% C.I. 0.308-0.692), with a small non-significant difference between the sexes (0.563 in males and 0.521 in females). When the same ADMIXTURE analysis was applied to the 916 contemporary Icelanders, we obtained a mean Norse ancestry of 0.704 (95% C.I. 0.699-0.709). This suggests that reproductive success among the first generations of Icelanders was stratified by ancestry, as genetic drift alone is unlikely to systematically alter ancestry at thousands of independent loci. This is perhaps not surprising, as many of the individuals of Celtic ancestry are thought to have been brought to Iceland as slaves, whose survival and freedom to reproduce is likely to have been under constraint (*21*).

Of the 23 pre-Christian Icelanders in our study, 20 have been tested for strontium isotopes 86 and 87 in their dental enamel, which can be used to determine whether they spent

the first six years of their lives in Iceland (non-migrants) or elsewhere (migrants) (*14, 32*). Three are certain to have been migrants based on high $^{87}Sr/^{86}Sr$ ratios (>0.710), and most probably, first generation settlers (Table 1). All three were unmixed according to our results; DAV-A8 and DAV-A9 (from the same burial site) were fully Norse, and SSG-A4 was a Celt (Fig. 2C). Two further individuals (SSG-A2 and SSG-A3, both from the same burial site as SSG-A4) have $^{87}Sr/^{86}Sr$ values that are considerably lower, but still too high for a childhood spent only in Iceland. Interestingly, the ancestry of SSG-A3 is estimated to have been a nearly even mix of Norse and Celtic, which indicates that at least some admixture occurred before arrival to Iceland, perhaps in Viking settlements in Scotland or Ireland. Of the 15 pre-Christians with $^{87}Sr/^{86}Sr$ values consistent with non-migrants, four were effectively unmixed (>90%) Norse (SVK-A1, NNM-A1, DKS-A1 and SSJ-A2) and two were unmixed Celts (ORE-A1 and KNS-A1) (Fig. 2C). This suggests that some segregation was maintained between Norse and Celtic individuals in Iceland, at least during the early generations of Icelanders.

An intriguing implication of the extensive drift that has accumulated in the Icelandic gene pool is that skeletal remains from Iceland could be dated based on the proportion of drift shared with the contemporary population. To test this, we sequenced the genome of an Icelander born in 1678 (KOV-A2), projected his genotypes onto the PCA in Fig. 2A and found that he was placed roughly halfway between the clusters of contemporary Icelanders and their source populations. This is consistent with KOV-A2's genome being shaped by only part of the drift that characterizes contemporary Icelanders. Following on from this observation, we projected genotypes from the remains of two Icelanders with uncertain temporal origin (TGS-A1 and FOV-A1). Based on their positions on the PC1 axis, we infer that TGS-A1 dates from the first generations after the settlement of Iceland, as he plots among other pre-Christian individuals. In contrast, FOV-A1 may have been born some centuries later, as he plots outside the range of pre-Christian individuals and closer to KOV-A2 and the contemporary Icelanders.

Further insights into the temporal origin of TGS-A1 and FOV-A1 can be gained through an analysis of drift shared with the Norse, Celtic and Icelandic reference populations. Fig. 3A shows a scatterplot of two D-statistics: *D(YRI, X; Norse, Iceland)* and *D(YRI, X; Celts, Iceland)* for all 27 ancient Icelanders (Table S11), which effectively distinguishes the signature of Iceland-specific drift (an axis that is parallel to the diagonal line) from that of Norse-Celtic ancestry (perpendicular deviation from that line). Consistent with the PCA plot in Fig. 2A, pre-Christian Icelanders plot very close to contemporary Norse and Celts. Furthermore, KOV-A2 is closer to contemporary Icelanders, whereas FOV-A1 plots roughly halfway between them and the contemporary Celtic and Norse source populations. TGS-A1, the second sample of unknown temporal origin, clusters with a group of four pre-Christian Icelanders (VDP-A5, DAV-A9, NNM-A1 and SVK-A1) who fall just outside the space occupied by contemporary Norse (a similar trend is observed in PC1 axis in Fig. 2A).

The greater genetic drift these individuals share with contemporary Icelanders is intriguing and is also apparent in the results of ADMIXTURE, run in supervised mode with three contemporary reference populations (K=3): Norse, Celtic and Icelandic (Fig. 3B, the correlation between the proportion of Icelandic ancestry and PC1 in Fig. A2 is |r|=0.913). However, the evidence that they belonged to the first generations of Icelanders is strong; namely, pre-Christian burials features, early ^{14}C dates (DAV-A9 and SVK-A1), strontium isotopes consistent with being raised outside of Iceland (DAV-A9) (Table 1), and the finding that three have unmixed Norse ancestry (DAV-A9, NNM-A1 and SVK-A1, Fig. 2C). Thus, temporal misclassification is doubtful. Another possible explanation is contamination from contemporary Icelanders, but contamination is estimated to be negligible in all four individuals (Table S4). We postulate that three different factors could account for the greater shared drift of the four pre-Christian Icelanders and TGS-A1 with contemporary Icelanders. First, subtle genetic drift affecting the Norse gene pools during the past 1100 years might have shifted allele

frequencies so as to make contemporary Norse less representative of their ancestors than is the case for the Celts. Second, these ancient Icelanders might have originated from a sub-population within Scandinavia that is not well represented in our reference samples. Third, it is possible that these ancient individuals contributed disproportionately to the gene pool of contemporary Icelanders. We reasoned that if genetic drift or population sub-structure in Scandinavian populations were responsible for the outlying position of VDP-A5, DAV-A9, NNM-A1 and SVK-A1 in Figs. 2A and 4A, then they would also be outliers in a PCA generated based only on the contemporary individuals from Scandinavia and the British Isles. As shown in Fig. 3C, these ancient Icelanders fall well within the cluster of contemporary Scandinavians. Thus, it may be concluded that these individuals contributed more to the contemporary Icelandic gene pool than other pre-Christian Icelanders in our study and that TGS-A1 may also fall into this category. We note that this observation is consistent with our inference that settlers of Norse ancestry had greater reproductive success than those of Celtic ancestry.

Ancient genomes are key to answering many important questions about the formation and evolution of human populations during recent millennia. The settlement of Iceland occurred around 1100 years ago at the height of the Viking age. Our study of ancient Icelanders reveals a highly admixed Norse and Celtic gene pool of this founding population. In addition to finding evidence for differential reproductive success by ancestry, we show that the Icelandic gene pool has been shaped by substantial genetic drift. The resultant founder events are one reason why this population has proven useful for the discovery of genotype-phenotype associations for rare sequence variants (9).

Table 1. **Summary of genomic sequence data from ancient Icelanders.**

Time period	Classification based on 87Sr/86Sr values	Sample	Archaeo-logical date	C14 date	Auto. genome depth of coverage [X]	Auto. genome covered [%]	SNP overlap w. European ref. [%]†	SNP overlap w. North European ref. [%]‡	Sex chr. karyo-type	mtDNA haplo-group	chrY haplogroup	87Sr/86Sr ratio
Pre-Christian	**Migrant**	DAV-A8	< 1000	<1050	3.31	89.12	90.14	93.92	XX	H1	NA	0.7121
		DAV-A9	< 1000	980-1020	0.43	26.66	21.79	22.66	XY	H1	I1	0.7118
		SSG-A2*	< 1000	NA	10.56	94.96	95.58	99.8	XY	J1c3g	R1b1a1a2a1a2c1	0.7095
		SSG-A3*	< 1000	NA	0.26	16.37	8.16	8.56	XY	T2b2b	I1	0.7093
		SSG-A4*	< 1000	NA	7.26	94.08	94.92	99.08	XX	J1b1a1a	NA	0.7117
	Non-migrant	DKS-A1	< 1000	NA	0.56	36.43	28.38	29.59	XY	U5a1h	R1a1a1b1a3	0.7088
		GRS-A1	< 1000	<1050	0.55	34.77	27.86	29	XY	K1a1b1b	R1a1a1b1a3b	0.7071
		GTE-A1	< 1000	NA	0.25	13.54	12.73	13.18	XY	H4a1a4b	R1a1a1b1a3a1	0.7061
		HSJ-A1	< 1000	NA	30.74	96.29	95.73	99.99	XY	H3g1	I1a1b3b	0.7074
		KNS-A1	< 1000	NA	0.71	43.95	36.2	37.8	XY	H5	R1b1a1a2a1a2c	0.706
		MKR-A1	< 1000	<1050	0.18	11.33	11.25	11.68	XY	K1c1b	R1a1a1b	0.7065
		NNM-A1	< 1000	NA	0.48	32.83	22.23	23.3	XY	H2a2b5a	R1a1a1b1a3a	0.7062
		ORE-A1	< 1000	NA	0.44	26.26	22.88	23.79	XY	K1a3a	R1b1a1a2a1a2	0.7087
		SBT-A1	< 1000	NA	6.01	93.72	94.36	98.46	XY	H3g1a	I1a2a1a2	0.7084
		SSJ-A2	< 1000	NA	0.36	22.41	14.04	14.77	XY	U5a1a1	I1a1b3	0.7077
		STT-A2	< 1000	975-1015	12.92	95.62	95.67	99.9	XY	U4b1b1	R1b1a1a2a1a2c1	0.7069
		SVK-A1	< 1000	<1050	1.07	67.81	60.95	63.47	XY	I2	I1	0.7078
		VDP-A5	< 1000	NA	1.32	81.88	83.64	87.05	XX	H3	NA	0.7085
		VDP-A6	< 1000	NA	1.86	85.92	63.06	65.83	XY	H1c3a	R1a1a1b1a3a	0.7089
		VDP-A7	< 1000	NA	0.83	70.07	83.15	86.68	XY	H4a1a1	R1b1a1a2a1a1b	0.7085
	Not available	FSS-A1	< 1000	NA	0.94	58.92	56.28	58.54	XX	U4a2	NA	NA
		NÞR-A2	< 1000	NA	0.49	31.88	31.79	32.93	XX	K1a2a	NA	NA
		YGS-B2	< 1000	NA	0.27	18.37	14.99	15.67	XXY	J1c1a	R1b1a1a2a1a	NA
Early Christian	**Non-migrant**	ÞSK-A26	1000-1104	1120	0.77	48.41	44.69	46.45	XY	J1b1a1a	R1a1	0.7061

17th century	**Not available**	KOV-A2	b. 1678	NA	0.57	36.05	34.03	35.42	XY	H1	R1b1a1a2a1a	NA
Unknown		FOV-A1	NA	NA	0.68	43.26	37.24	38.91	XY	HV17a	R1b1a1a2a1a2c1a1	NA
		TGS-A1	NA	NA	1.01	59.32	54.86	57.14	XY	T2e1	R1b1a1a2a1a2d	NA

† N=227,056 SNPs. ‡ N=404,066 SNPs. Samples marked with an asterisk symbol (*) were excavated from the same site as a sample that have been subjected to carbon dating, yielding the date estimate 980-1020 A.D
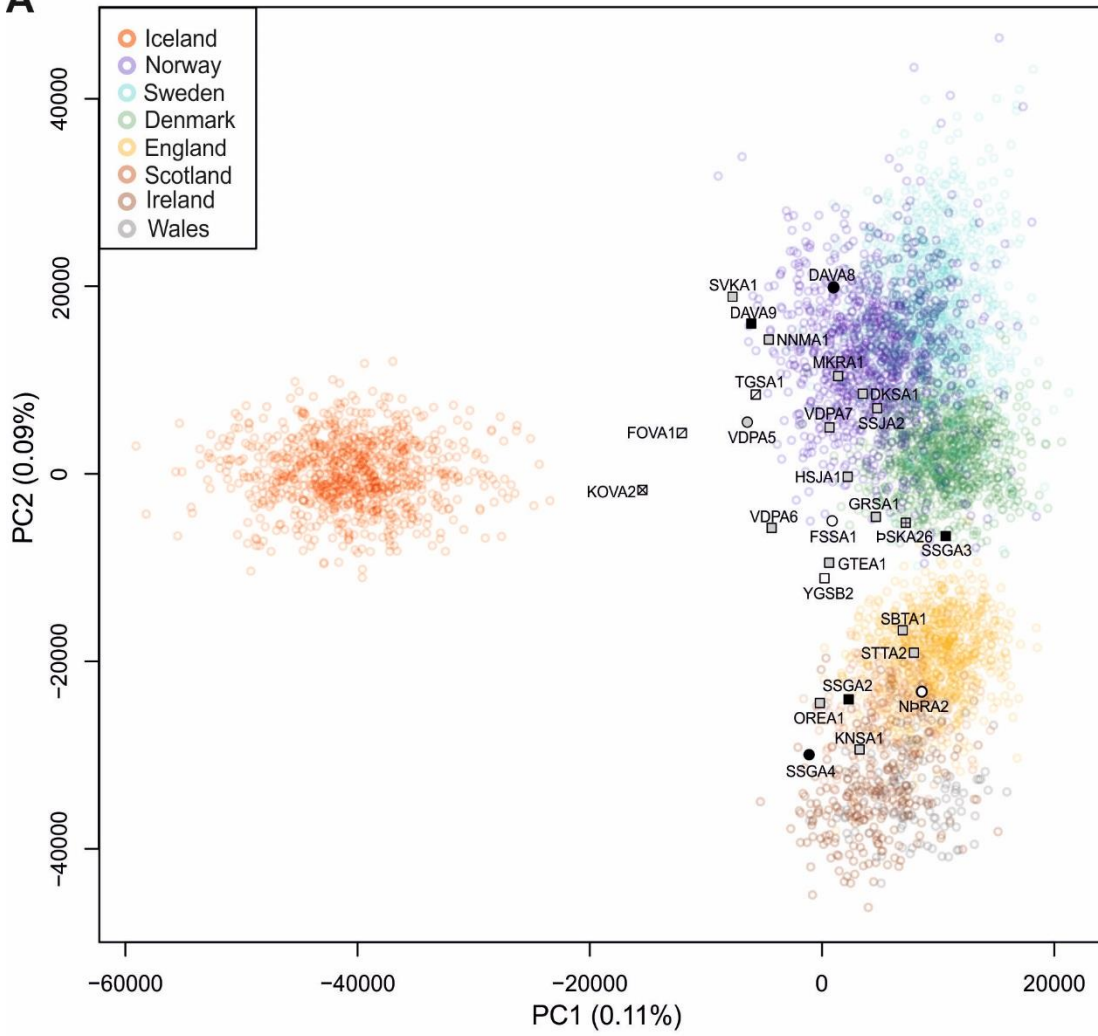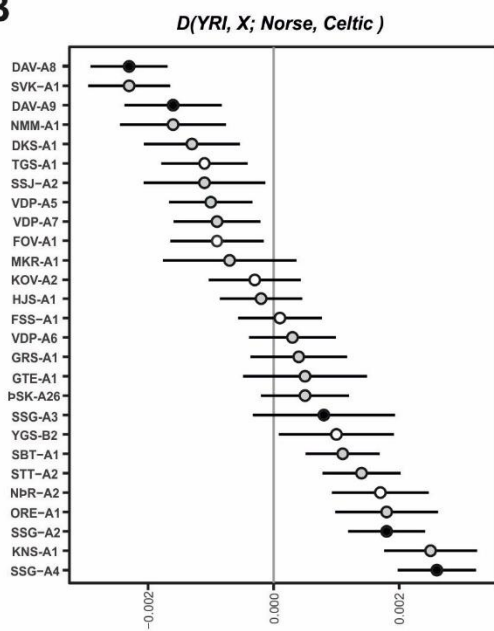
Fig. 1.
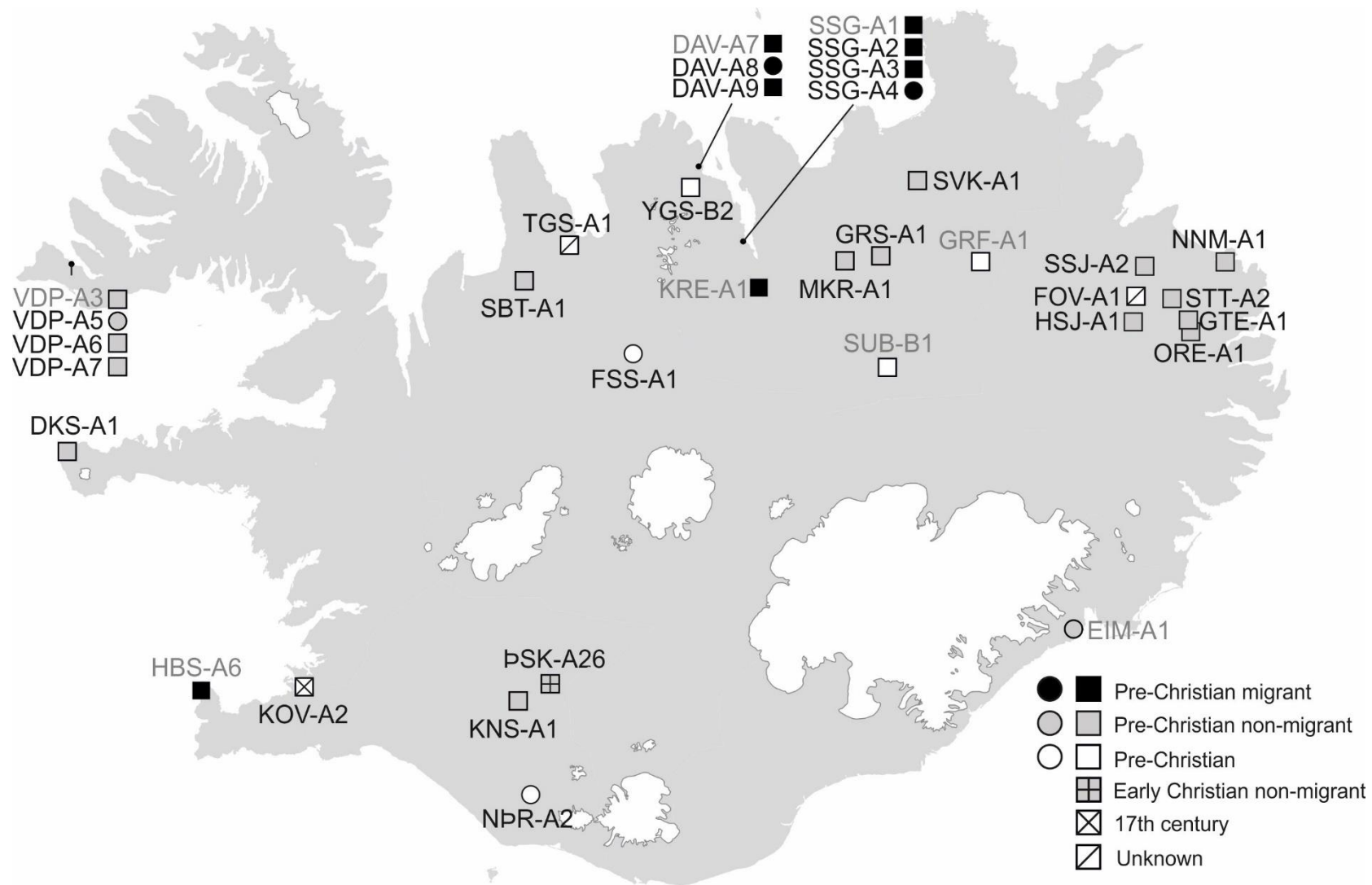
Fig. 1. **A map of Iceland showing the location of excavation sites of skeletal remains sampled for this study.** The first three letters in the sample ID are an abbreviation of the site (SM S1). Circles indicate females and squares males. Eight samples (light grey labels) were ultimately excluded from further analysis.
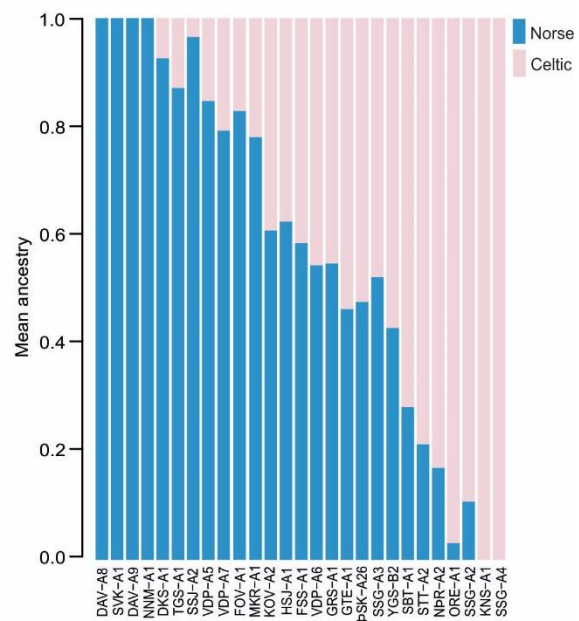
Fig. 2.

Fig. 2. **Ancestry of ancient Icelanders.** (**A**) The ancient Icelanders projected onto the first two eigenvectors of a PCA based on contemporary individuals from Scandinavia, the British Isles and Iceland. Variance explained by the first two components are shown in brackets. Symbols for ancient individuals are as specified in Fig. 1. (**B**) Plot of D-statistics reflecting the differential affinity of ancient Icelanders *(X)* to Norse and Celtic reference populations, using the Yoruba from Nigeria (YRI) as an outgroup (*n*=91). Negative values indicate more Norse ancestry and positive values indicate more Celtic ancestry, whereas values close to zero indicate mixed ancestry (as the contemporary Norse and Celts are the two best candidate source populations for Icelanders). Symbols are black for migrants, grey for non-migrants, and white where no $^{87}Sr/^{86}Sr$ values were available. (**C**) Estimated Norse and Celtic admixture proportions for the ancient Icelanders using ADMIXTURE in supervised mode with K=2.
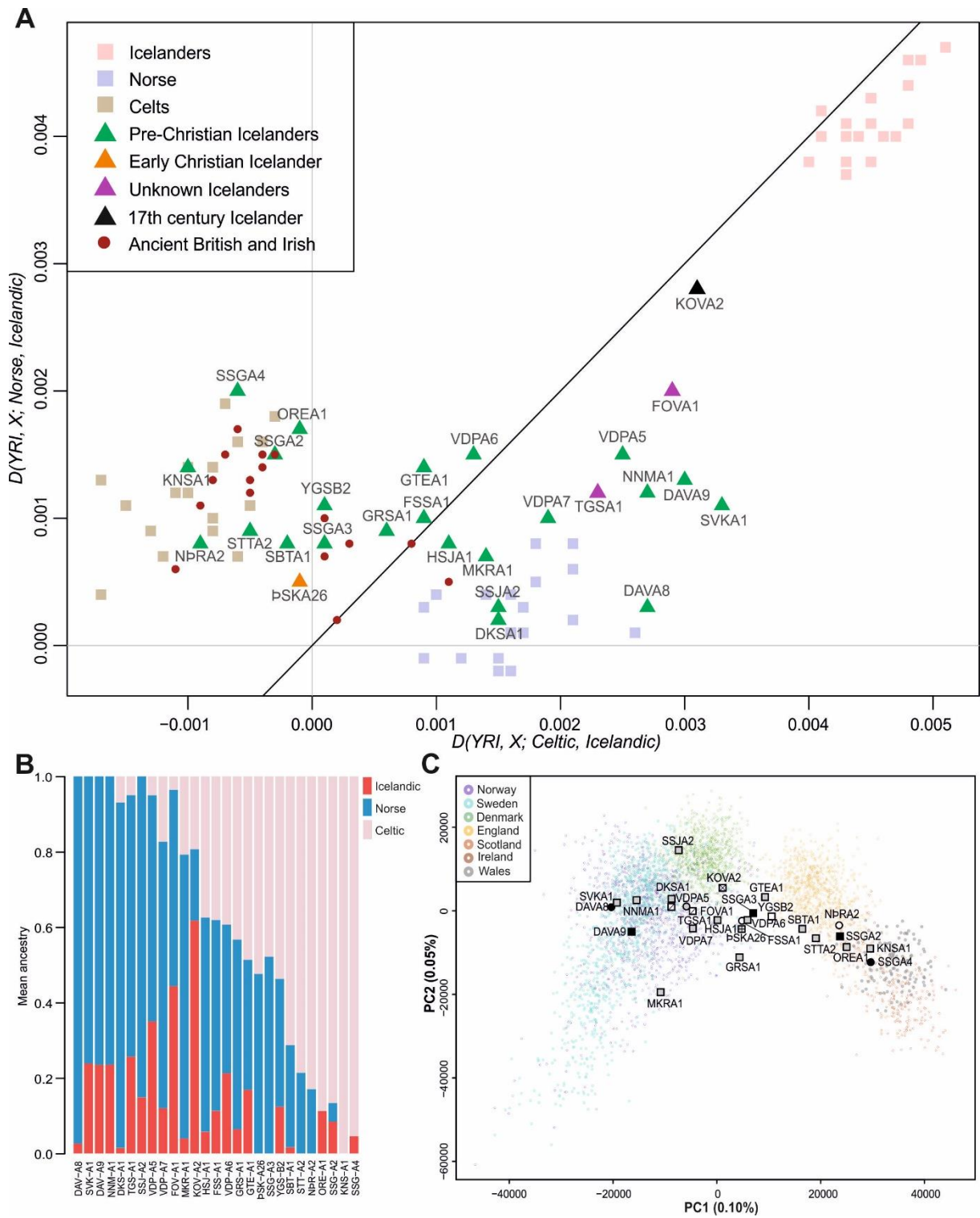
Fig. 3.

Fig. 3. **Shared drift of ancient and contemporary Icelanders** (**A**) Scatterplot of $D(YRI, X;$ *Norse, Iceland*) and $D(YRI, X; Celts, Iceland)$ for all 27 ancient Icelanders (X). To help interpret these results, we included values for ancient individuals from the British Isles and a subset of contemporary individuals (who were removed from the reference population for this analysis). (**B**) Estimated Norse, Celtic and Icelandic admixture proportions for the ancient Icelanders using ADMIXTURE in supervised mode with K=3. (**C**) Ancient Icelanders projected onto a PCA of contemporary individuals from Scandinavia ($n=3,118$) and the British Isles ($n=1,436$). Variance explained by the first two components are shown in brackets. Symbols for ancient individuals are as specified in Fig. 1.

Reference and notes:

1. *The book of settlements: Landnámabók* (University of Manitoba, Winnipeg, 1972).

2. K. Smith, Landnám: the settlement of Iceland in archaeological and historical perspective. *World Archaeol* **26**, 319-347 (1995).

3. J. Steffensen, *Menning og meinsemdir: ritgerðarsafn um mótunarsögu íslenskrar þjóðar og baráttu hennar við hungur og sóttir*. (Ísafoldarprentsmiðja, Reykjavík, 1975).

4. G. Jones, *A History of the Vikings*. (Oxford Univeristy Press, Oxford, 1984).

5. S. Goodacre *et al.*, Genetic evidence for a family-based Scandinavian settlement of Shetland and Orkney during the Viking periods. *Heredity* **95**, 129-135 (2005).

6. A. Helgason, S. Sigurdardottir, J. R. Gulcher, R. Ward, K. Stefansson, mtDNA and the origin of the Icelanders: Deciphering signals of recent population history. *Am J Hum Genet* **66**, 999-1016 (2000).

7. A. Helgason *et al.*, Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland. *Am J Hum Genet* **67**, 697-717 (2000).

8. A. Helgason *et al.*, Sequences From First Settlers Reveal Rapid Evolution in Icelandic mtDNA Pool. *PLoS genetics* **5**, (2009).

9. D. F. Gudbjartsson *et al.*, Large-scale whole-genome sequencing of the Icelandic population. *Nat Genet* **47**, 435-444 (2015).

10. A. Helgason, G. Nicholson, K. Stefansson, P. Donnelly, A reassessment of genetic diversity in Icelanders: Strong evidence from multiple loci for relative homogeneity caused by genetic drift. *Ann Hum Genet* **67**, 281-297 (2003).

11. E. Ó. Hreiðarsdóttir, "Íslenskar perlur frá víkingaöld: með viðauka um perlur frá síðari öldum," M.A. thesis. (2005).

12. K. Eldjárn, *Kuml og haugfé*. A. Friðriksson, Ed., (Mál og menning, Reykjavík, 2000).

13. P. L. Ascough *et al.*, Radiocarbon reservoir effects in human bone collagen from northern Iceland. *J Archaeol Sci* **39**, 2261-2271 (2012).

14.   H. G. a. T. D. Price, "The settlement of Iceland; analysis of strontium isotopes on human teeth. A preliminary discussion of results". (Fornleifastofnun Íslands, Reykjavik, Iceland, 2006).

15.   J. M. Enk *et al.*, Ancient whole genome enrichment using baits built from modern DNA. *Mol Biol Evol* **31**, 1292-1294 (2014).

16.   S. Sawyer, J. Krause, K. Guschanski, V. Savolainen, S. Paabo, Temporal Patterns of Nucleotide Misincorporations and DNA Fragmentation in Ancient DNA. *PloS one* **7**, e34131 (2012).

17.   P. Skoglund, J. Stora, A. Gotherstrom, M. Jakobsson, Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J Archaeol Sci* **40**, 4477-4482 (2013).

18.   C. Spatacean, "Women in the Viking Age. Death, life after death and burial customs", M.A. Thesis. (2006).

19.   A. Helgason, B. Hrafnkelsson, J. R. Gulcher, R. Ward, K. Stefansson, A populationwide coalescent analysis of Icelandic matrilineal and patrilineal genealogies: Evidence for a faster evolutionary rate of mtDNA lineages than Y chromosomes. *Am J Hum Genet* **72**, 1370-1388 (2003).

20.   A. L. Price *et al.*, The Impact of Divergence Time on the Nature of Population Structure: An Example from Iceland. *PLoS genetics* **5**, e1000505 (2009).

21.   G. Karlsson, *Iceland's 1100 Years: The History of a Marginal Society*. (C. Hurst, London, 2000).

22.   L. M. Cassidy *et al.*, Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 368-373 (2016).

23.   Q. M. Fu *et al.*, The genetic history of Ice Age Europe. *Nature* **534**, (2016).

24.   P. Skoglund *et al.*, Genomic Diversity and Admixture Differs for Stone-Age Scandinavian Foragers and Farmers. *Science* **344**, 747-750 (2014).

25.   P. Skoglund *et al.*, Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336**, 466-469 (2012).

26. D. Reich, K. Thangaraj, N. Patterson, A. L. Price, L. Singh, Reconstructing Indian population history. *Nature* **461**, 489-494 (2009).

27. R. Martiniano *et al.*, Genomic signals of migration and continuity in Britain before the Anglo-Saxons. *Nature Communications* **7**, (2016).

28. S. Schiffels *et al.*, Iron Age and Anglo-Saxon genomes from East England reveal British migration history. *Nat Commun* **7**, 10408 (2016).

29. N. Patterson *et al.*, Ancient Admixture in Human History. *Genetics* **192**, 1065-1093 (2012).

30. E. W. Hill, M. A. Jobling, D. G. Bradley, Y-chromosome variation and Irish origins. *Nature* **404**, 351-352 (2000).

31. A. O. Karlsson, T. Wallerstrom, A. Gotherstrom, G. Holmlund, Y-chromosome diversity in Sweden - a long-time perspective. *Eur J Hum Genet* **14**, 963-970 (2006).

32. T. D. Price, H. Gestsdottir, The first settlers of Iceland: an isotopic approach to colonisation. *Antiquity* **80**, 130-144 (2006).

Acknowledgments:

Author contributions:

SSE, KS and AH designed and directed the research. SSE and AH analysed the data, with EDG, AJ, VBG, ELT, MSE, HJ, SG and MTPG providing assistance with particular tasks. SSE and MS-V generated the ancient genomic data with laboratory guidance and support

from ÁS, DNM, SS, CL-F, ÓÞS and MTPG. Modern reference datasets were provided by EH, PM, IK, TO, KA, TFH, TW, GLC and EG. Archaeological and osteological context was provided by JWW, SK and LÁ, who also provided access to samples. SSE, KS and AH wrote the manuscript and supplements with input from MS-V, MTPG, HJ, KHSM and SS.

Competing financial interests:

Authors affiliated with deCODE Genetics are employed by the company, which is owned by Amgen, Inc.: SSE, EDG, AJ, VBG, ELT, MSE, KHSM, ÁS, DNM, HJ, SS, ÓÞS, KS and AH.

Supplementary Materials

Materials and Methods
Supplementary Text
Figs. S1 to S7
Tables S1 to S11
References (*33-69*)

Supplementary Materials for

Ancient genomes from Iceland reveal the making of a human population

S. Sunna Ebenesersdóttir, Marcela Sandoval-Velasco, Ellen D. Gunnarsdóttir, Anuradha Jagadeesan, Valdís B. Guðmundsdóttir, Elísabet L. Þórðardóttir, Margrét S. Einarsdóttir, Kristján H. S. Moore, Ásgeir Sigurðsson, Droplaug N. Magnúsdóttir, Hákon Jónsson, Steinunn Snorradóttir, Eivind Hovig, Pål Møller, Ingrid Kockum, Tomas Olsson, Lars Alfredsson, Thomas F. Hansen, Thomas Werge, Gianpiero Cavalleri, Edmund Gilbert, Carles Lalueza-Fox, Joe W. Walser III, Steinunn Kristjánsdóttir, Shyam Gopalakrishnan, Lilja Árnadóttir, Ólafur Þ. Magnússon, M. Thomas P. Gilbert, Kári Stefánsson and Agnar Helgason

correspondence to: sunna@decode.is (SSE); agnar.helgason@decode.is (AH); kstefans@decode.is (KS)

This PDF file includes:

SUPPLEMENTARY TEXT: TABLE OF CONTENTS

S1. Archaeological information

Selection of skeletal remains

For this study, 35 ancient Icelandic skeletal remains were screened for DNA preservation. The remains were unearthed during the period 1900-1996 A.D. and are stored at the National Museum of Iceland. Details of the sampled individuals are given in Table S1. Several factors that contributed to the initial sample selection process are discussed below.

(1) Dating evidence: We focused on human remains dating to or shortly after the settlement of Iceland (870-930 A.D). The replacement of the Old Norse religion by Christianity around 1000 A.D. entailed major changes in burial customs. In Christian burials, the dead were typically laid in coffins in relatively large cemeteries, with graves in an east-west orientation. However, pre-Christians were buried according to Old Norse customs in peripheral sites called *kuml*. They are usually found as single inhumations, with the body lying on its side in a north-south orientation. A key factor in identifying a pre-Christian burial site is the presence of grave goods, which were typically buried with the deceased to be used by them in the afterlife. Typical grave goods include animals, boats, weapons, jewelry and household items (*12*). The archaeological classification of a burial site can effectively be used to date the death of an individual to either the pre-Christian period (~870-1000 A.D.) or the Christian period (>1000 A.D). This dating method is well supported by results from both radiocarbon dating and datable grave goods (*11, 13, 14, 32*).

(2) Place of development: We preferentially selected remains that had been subjected to enamel strontium isotope analysis. The enamel of adult teeth is formed in early childhood, and because the ratio of isotopes 87 and 86 ($^{87}Sr/^{86}Sr$) absorbed by organisms varies geographically, they can be used to infer the provenance of human remains (*33*). In particular, the relatively large geological difference between $^{87}Sr/^{86}Sr$ in Iceland and the rest of Europe allows for inferences about whether remains found in Iceland were from an individual who spent their first years there or elsewhere (*14, 32*). It is very likely that that an individual who, according to burial features, died before 1000 A.D., and whose enamel displays a $^{87}Sr/^{86}Sr$ value indicative of being born and bred outside of Iceland, was a first generation settler.

(3) Geographical location and archaeological relevance: We selected samples unearthed from widely geographically distributed sites to best represent the population of Iceland during the settlement period (Main text Fig. 1). We sampled multiple remains from three sites which contained more than one burial. (DAV: 3, SSG: 4, and VDP: 4).

Other practical issues contributed to our selection of samples, such as the availability of remains and their quality. All the remains analyzed in this paper were sampled for a previous study of Helgason et al. (*34*), in which one tooth (usually a molar or premolar) was sampled from each of the remains. Frozen DNA extracts were available for many of the samples; in some cases, frozen ground tooth samples were also available. For sequencing, we prioritized samples that had a sufficient volume of DNA extract and/or ground tooth for fresh extraction. Prioritization was also influenced by the quality of each sample, assessed based on success in PCR amplification and clonal sequencing of mtDNA sequences in the previous study (*34*).

Archaeological context summaries of ancient Icelandic skeletal remains

Here, we provide a summary of archaeological context of the sites and remains that were sampled for this study. More detailed information about the pre-Christian sites can be found in the reference publication Kuml and Haugfé (*12*).

Dalvík (Brimnes) [DAV]: A pre-Christian burial site was found in 1908 and excavated the following year. It is one of the largest and most studied pre-Christian burial sites in Iceland. Thirteen human skeletal remains, six horse skeletons, and the remains of three dogs were found at the site. In one of the graves, the deceased had been placed in a sitting position at the rear of a boat. We selected three samples for sequencing from this site: DAV-A7, DAV-A8 and DAV-A9.

Öndverðarnes [DKS]: In 1962, a human skeleton (DKS-A1) was discovered during road construction. Grave goods included a sword, a spearhead, a knife, a shield-boss, a bone-pin, and fragments of iron. According to a morphological analysis, the skeletal remains show evidence of developmental delay that could be explained by hypogonadism caused by Klinefelter syndrome, testicular disorder or castration (*35*).

Einholt [EIM]: In 1979 a pre-Christian burial was found near a farmhouse. The remains of a female skeleton (EIM-A1), along with grave goods, were found at the site.

Fossvellir [FOV]: A single human skeleton (FOV-A1) was found by bridge builders near the bridge over Jökulsá á Brú in 1929. The remains seem to have been placed at the site after the individual was deceased. The bones had been carefully arranged on top of each other and were surrounded by stone slabs and turf. The remains are of unknown temporal origin. However, there have been some speculations about who the deceased might have been. It has been claimed that these are the remains of a silver salesman who was thought to have disappeared (possibly murdered) in this area in 1830 (*36, 37*). However, it is also known that there was a parliament site (ca. 1624-1800) in close proximity, where individuals convicted of serious crimes were executed (*37*).

Fremri-Svartárdalur [FSS]: In 1960, an eroded burial was found around 150 meters from an old farmstead, which contained a human skull and a few other bones (FSS-A1). No grave goods were found at the site, which had clearly been disturbed previously. The site is classified as pre-Christian.

Gilsárteigur [GTE]: In 1949, field-leveling exposed a pre-Christian burial site near an old farm site. The remains of two skeletons were excavated in 1957. Both burials contained grave goods. One of the remains (GTE-A1) found at the site was screened for DNA preservation.

Grímsstaðir [GRS]: Three pre-Christian burials were found in close proximity to each other near the site of a farmstead. We analysed one of the skeletal remains (GRS-A1), which were excavated in 1937. No grave goods were found at the site.

Grímsstaðir 'the old' [GRF]: In 1962 an eroded pre-Christian burial site revealed the skeletal remains of three individuals along with a spear-head. We sequenced one of the remains (GRF-A1) excavated from this site.

Hafurbjarnarstaðir [HBS]: An eroded pre-Christian burial site was discovered in 1828 and excavated in 1947. The remains of seven to eight individuals (one child, two females and four to five individuals whose sex could not be identified by morphological analysis) were found at the site. Grave goods were found at the site, along with dog and horse bones. We analysed a sample from skeletal remains found in one of the burials (HBS-A1).

Hrólfsstaðir [HSJ]: In 1996, human skeletal remains (HSJ-A1) were unearthed by a bulldozer. This pre-Christian site had clearly been disturbed before. A comb, knife, and pieces of charcoal were found in the grave.

Karlsnes [KNS]: Human skeletal remains (KNS-A1) were found in an eroded pre-Christian burial site in 1932. There are no records of a farm having existed nearby. The remains, which were in a supine position, appeared to have been undisturbed since burial. Grave goods included a spearhead, a knife, two lead weights, three beads, and a small stone.

Kópavogur [KOV]: In 1988, two skeletal remains were excavated during road constructions in Kópavogur. Based on archaeological evidence, the remains were identified as a female, born 1664, and a male, born 1678. According to historical records (*38*), they were executed in 1704 for the murder of the female's husband. The male was beheaded, and his impaled head publicly exhibited, whereas the female was drowned. Their remains were buried buried in unconsecrated ground at a site called Hjónadysjar. The remains of the male (KOV-A2) were selected for sequencing in this study.

Kroppur [KRE]: During road work in 1900, two pre-Christian skeletal remains were revealed west of a farm. Grave goods were found in both graves. We sequenced one of the remains (KRE-A1) found at the site.

Neðri-Þverá [NÞR]: In 1949, two human skulls were found near a farm site. No grave goods were found, but archaeologists concluded that it was likely a badly eroded pre-Christian burial site. The remains of NÞR-A2 were selected for sequencing.

Njarðvík [NNM]: In 1981, a human skull (NNM-A1) was found at a site considered to be a badly damaged pre-Christian burial.

Ormsstaðir [ORE]: In 1966, a pre-Christian site near an old farmstead was excavated after being exposed during field leveling. One human skeleton (ORE-A) was found, along with an axe, a knife, and three lead weights. A single human bone from another individual was found nearby.

Skeljastaðir [ÞSK]: The remains of ÞSK-A26 are from the Christian cemetery at Skeljastaðir in Þjórsárdalur, excavated in 1939. The remains are dated to before 1104 A.D., as the site was abandoned in the wake of a volcanic eruption of Mount Hekla in that year (*39*).

Sílastaðir [SSG]: A cluster of four pre-Christian graves was discovered in 1947. Based on morphological analysis, three of the skeletons were deemed male, and one female. We selected all four remains for sequencing: SSG-A1, SSG-A2, SSG-A3 and SSG-A4.

Smyrlaberg [SBT]: In 1952, a pre-Christian burial site was found in an old gravel quarry. Two years later its excavation revealed a male skeleton (SBT-A1) and an iron knife. Another grave, badly damaged, was found nearby, but only fragments of bone could be identified within

Straumur [STT]: In 1952, a pre-Christian burial site was excavated, which included the remains of four individuals (one child, one male, one female, and another adult whose sex could not be determined by morphological analysis). Grave goods included a horse bone, a small axe, thirty boat rivets, a lead weight, two pebbles and a knife. We selected STT-A2 for sequencing.

Suðurárbotnar [SUB]: An eroded grave containing two skulls and a few bones, thought to be pre-Christian, was discovered in 1947. We selected SUB-A1 for sequencing.

Surtsstaðir [SSJ]: In 1945, an eroded and previously disturbed pre-Christian burial was found and excavated four years later. The remains of two individuals were found at the site, along with grave goods. We selected SSJ-A2 for sequencing.

Svínadalur [SVK]: In 1961, human skeletal remains (SVK-A1) were brought to the National Museum. They had been exposed for many years near an old farmhouse. There were no grave goods found at the site, but the remains are thought to be pre-Christian.

Tunga í Gönguskörðum [TGS]: Human skeletal remains (TGS-A1) of unknown temporal origin were excavated in 1981 by inhabitants at a nearby farm.

Vatnsdalur [VDP]: In 1964, a pre-Christian boat grave was discovered by chance at this site. In it were seven previously disturbed human skeletal remains (three females and four male), along with a dog skeleton. Grave goods included a knife, thirty beads, a silver Thor's hammer, a fragmented Cufic coin (ca. 870-930 AD), and jewelry. We sequenced four of the remains found at the site (VDP-A3, VDP-A5, VDP-A6 and VDP-A7).

Viðar (Másvatn) [MKR]: In 1989, human skeletal remains (MKR-A1) were revealed during road work. The remains are thought to be from a pre-Christian burial site, dated to pre-1477 based on volcanic ash chronology.

Ytra-Garðshorn [YGS]: In 1952, a pre-Christian burial site was discovered during road construction east of a farm and excavated two years later. The site included the disturbed remains of nine human skeletons (four males, two females, one child and two individuals whose sex have not be inferred based on morphological analysis). This site is one of the largest pre-Christian burial sites discovered in Iceland. There were grave goods in all graves. We selected one of the remains (YGS-B2) for sequencing.

S2. Ancient DNA processing

DNA extraction

Extraction of DNA was conducted in dedicated aDNA clean laboratory facilities located several kilometers from the main deCODE site in Reykjavik. Surfaces of the laboratory and equipment were cleaned repeatedly with bleach, and full body suits, gloves, hoods and facemasks were worn at all times. Blank controls were incorporated in extractions, library preparations and each step of PCR to monitor for contamination.

Prior to this study, all sampled teeth (*n*=36) were ground to powder using an analytical mill, followed by DNA purification with phenol-chloroform, as described elsewhere (*8, 34*). All remaining extraction aliquots and/or ground teeth were stored at -20°C. In total, 32 samples had enough DNA extraction aliquot remaining (>20 µL) for at least one library build. In addition, 17 new DNA extracts were made from ground tooth samples using a modified version of the silica-in-solution method based on Rohland and Hofreiter (*40*), with adjustments (*41*) (Table S2). The amount of ground tooth varied, but was generally 200-400 mg. A brief description of the method follows. The powdered samples were incubated for 24 hours at 37°C in 5 ml of an EDTA-based digestion buffer containing 0.25 mg/mL Proteinase K. Following digestion, the DNA was bound to a silica suspension in the presence of chaotropic salts. For this step, we used the Dabney et al. (*41*) binding buffer recipe. The pellets were then washed twice in 1 ml 80% EtOH, and DNA was eluted in 80 µL low-salt EB buffer (Qiagen, Valencia, California). We kept any undissolved bone powder, binding supernatant and the silica pellet, and stored it with the aliquots of DNA extracts at -20°C.

DNA library preparation, enrichment and sequencing

For each sample, between 20-30 µL of extracted DNA was converted to a double-stranded library using NEBNext ® DNA Library Prep Master Mix set for 454 (catalog number E6070L, New England BioLabs, Ipswich, Massachusetts). One to two libraries were built per extract, resulting in 61 libraries (1 to 4 libraries per individual) (Table S2). All libraries were prepared according to the manufacturer's instructions, with a few modifications and adjustments listed below. The extracts were not nebulized as aDNA is already highly fragmented in nature. The protocol consists of three modules. In the first module, DNA was blunt ended in a 50 µL reaction using 5 µL of NEBNext 10x end repair buffer and 2.5 µL of NEBNext End Repair Enzyme Mix (T4 polynucleotide kinase and T4 DNA polymerase). The samples were then incubated in a thermal cycler for 20 minutes at 12°C, followed by 15 minutes at 37°C. Following this, the reactions were purified in MinElute silica spin columns (Qiagen, Valencia, California), using $5\times$ PB binding buffer and 720 µL of PE wash buffer. Samples were eluted in 30 µL of EB buffer, with an incubation step at 37°C for 15 minutes. All centrifugation steps were conducted at $16,000\times g$. In the second module, Illumina-specific adapters, prepared following the protocol described by Meyer and Kircher (*42*), were ligated to the blunt-ended DNA. Ligation of adapters was conducted in 50 µL reactions using 10 µL of NEBNext $5\times$ quick ligation buffer, 5 µL of Quick T4 DNA ligase and 0.25 µM as final adapter concentration. Reactions were incubated in a heat block for 20 minutes at 20°C. After ligation, reactions were purified as above, with the exception of being eluted in 42 µL of EB buffer in this step. In the third module, remaining DNA nicks were repaired, and the fill-in procedure was conducted in a 50 µL reaction using 3 µL of *Bst* DNA polymerase and 5 µL of NEBNext adapter Fill-in reaction buffer, along with the adapter-ligated DNA from above. Reactions were incubated in a thermal cycler for 20 minutes at 65°C, followed by 20 minutes at 80°C to inactivate the *Bst* enzyme. Once molecules had been ligated to suitable adapters and the PCR setup was finished,

28

the libraries were handled outside of the clean room facilities.

The libraries were amplified in 100 µL PCR reactions containing 25 µL of library template, 5 U AmpliTaq Gold Polymerase (Applied Biosystems, Foster City, CA), 1× AmpliTaq Gold Buffer, 2.5 mM MgCl$_2$, 0.8 mg/mL BSA, 200 µM dNTPs, and 200 nM of forward, and reverse indexing primers (Meyer and Kircher, 2010). Thermocycling conditions were 10 min at 95°C, followed by 14 cycles of 20 secs at 95°C, 30 secs at 60°C, and 40 secs at 72°C, and a final 5 min elongation step at 72°C. Samples were eluted in 31 µL of EB buffer, with an incubation step at 37°C for 15 minutes.

Amplified indexed libraries were purified in QiaQuick silica spin columns using 5× PB binding buffer and 720 µL of PE wash buffer. Samples were eluted in 30 µL of EB buffer, with an incubation step at 37°C for 15 min. After purification, 1 µL of each library was measured on an Agilent 2100 Bioanalyzer with High Sensitive DNA Analysis Kit (Santa Clara, California). Six libraries were excluded after first amplification on the basis of low molar concentration (<100 pmol/L) (Table S2). If needed, samples were reamplified for 4-10 cycles to obtain ≥100 ng of library template. Reamplification of libraries were performed in 100 µL PCR reactions, containing 25 µL of DNA template, 1 × Phusion® High-Fidelity PCR Master Mix with GC Buffer (NEB), 2.5% DMSO, 0.2 µM agnostic primers IS5 (5'- AATGATACG GCGACCACCGA) and IS6 primers (5'- CAAGCAGAAGACGGCA TACGA) (42), and H$_2$O. Thermocycling conditions were 3 s at 98°C, followed by 20 s at 98°C, 30 s at 60°C, 30 s at 72°C, with a final 5 min elongation step at 72°C.

With the purpose of screening for endogenous DNA content, the double-stranded DNA libraries were pooled in equimolar concentrations by the values obtained by quantification and sequenced on Illumina MiSeq (2x150 cycles + 7 cycle index run). Based on the initial screening, two libraries (27_P2 and 52_P1) were excluded due to their low proportion of human reads (< 1%) (Table S2). However, several new libraries were constructed from the most efficient samples. Following molecular screening, 51 candidate libraries were selected for enrichment (15) (Table S2) using the MyBait Human Whole Genome Capture Kit (MYbaits-HuWGE) from MYcroarray (Ann Arbor, Mi) (15). Selection was performed following the manufacturer's instructions for version 1.3.8 of the MyBait user manual, along with a few modifications from Version 2.3.1. The libraries were concentrated to the required volume using a SpeedVac (Savant Plus Model SC210A-120). Bait library hybridization was left for 24 hours at 65°C. Instead of releasing the captured DNA target molecules from the RNA baits using a NaOH treatment (as suggested in version 1.3.8), we resuspended the beads with 30 µL Molecular Biology Grade Water, and amplified the enriched DNA directly after recovering the targets.

After enrichment, libraries were amplified for 14-16 cycles in 50 µl PCR reactions, containing 15 µL of enriched library as template, 1× KAPA HotStart ReadyMix, 0.8 mg/mL BSA, 0.3 µM agnostic primers IS5 and IS6 (42) (5'-CAAGCAGAAGACGGCATACGA) and the same PCR set-up condition as above. Following amplification, the enriched libraries were purified and quantified using the same method as before. Next, they were pooled together and sequenced on Illumina MiSeq platform (2x150 bp). We required the endogenous DNA to be >15% to be sequenced further. Seven libraries did not meet this criterion and were excluded (Table S2). In total, 46 libraries, including four from pre-capture, were paired-end sequenced on the Illumina HiSeq2500 (2x100 bp) (≥1 lane) platform. One to five lanes were sequenced per library, resulting in a total of 85 HiSeq lanes (Table S3).

Processing and alignment of reads

Nucleotide base calling and quality score assessment was performed using specific Real Time Analysis (RTA) software provided by Illumina. Due to postmortem degradation, aDNA fragments are usually very short, resulting in sequencing of the adapter, which has been ligated during the library preparation. Thus, AdapterRemoval v.1.5 (*43*) was used to remove adapter sequences, as well as stretches of consecutive low quality bases and N's. We required trimmed reads to be at least 25 bp for alignment. Alignment was performed to NCBI build 38 of the human reference genome using Burrows-Wheeler Algorithm (BWA) v. 0.7.10, (*44*) with the seed disabled (-l 1024) to improve accuracy and the minimum base quality set to 15. Data was merged into library level and duplicated using Picard (PicardTools v.1.79, http://broadinstitute.github.io/picard/). Reads with mapping quality inferior to 30 (SAMtools v.1.5) (*45*) were discarded before remerging to sample level. Read depth and coverage were determined using BEDtools-2.18.2 (*46*) and an in-house python script. Finally, base quality scores were rescaled with mapDamage 2.0 (*47*) to exclude likely-damaged bases (q<30). Ancient genomes from previous published studies (*22, 27, 28*) (Table S8) were processed in the same way after their BAM files were downloaded.

S3. Quality control

Error rate estimates

The error rates were estimated for each of our ancient human samples with ANGSD (*48*) v. 0.911 (*48*) using a high quality 'error free' human genome and a human outgroup. The estimation is based on the idea that any excess of derived alleles for the ancient sample compared to the high quality genome is due to errors, since they should have the same expected number of derived alleles compared to the outgroup (*49*). We used an inferred ancestor of Homo sapiens (*50*) and a randomly sampled high quality (30X) whole genome sequenced at deCODE Genetics. Type-specific error rates and overall error rates are shown in Fig. S1. Both estimation methods show that the most dominant errors in our ancient samples are transitions (C→T, G→A) typical of ancient DNA damage caused by post-mortem deamination of cytosine.

MapDamage

To verify the authenticity of our samples, the existence of nucleotide misincorporation was investigated using the Bayesian approach implemented in mapDamage 2.0 (*47*). An increase in C→T transitions at the at the 5' end was observed in reads from all samples, which is consistent with the presence of authentic aDNA (Fig. S2).

mtDNA motifs and contamination estimates

Although mtDNA is a circular molecule, mapping algorithms treat the ends of the artificially linear mtDNA reference genome as unconnected. There is therefore a bias against the mapping of paired-end reads in these locations. To circumvent this bias, reads mapped to the rCRS were remapped to an extended version of the linear rCRS reference sequence (*51*), obtained by adding the first 500 bp to the end and the last 500 bp to the start (*45*). All mitochondrial consensus sequences were identical to the ones previously reported for the mtDNA control region (sites 16517 to 409 and 16055 to 16410) obtained using PCR and cloning, which were partly based on different DNA extractions (*34*). After calling consensus mtDNA sequences, we estimated the contamination affecting mtDNA by estimating the frequency of minor alleles at positions where the major allele was rare (<5%) in a large sample of contemporary Icelanders (n=30,219). If no alternative allele was found, we allowed for 10% minor allele frequency. A point estimate c of mtDNA contamination was estimated following Skoglund et al. (*17*), assuming independence of the bases: $\hat{c}=N_{alternative}/(N_{consencus}+N_{alternative})$. Estimates are listed in Table S5.

X chromosome contamination

We used the 'Contamination' program in ANGSD (*48*) v.0.911 to estimate X chromosome contamination, as described in (*52*). As the X chromosome is haploid in males, mismatches at polymorphic sites are either due contamination or error in sequencing or mapping. Assuming that sequencing errors are randomly distributed across the genome, the discordance in the rate of heterozygous calls between known polymorphic sites and their adjacent monomorphic sites is evaluated to assess the level of contamination. We used default parameters and the provided mapFile and hapFile (after liftover to NCBI build 38) based on fixed set of SNPs known to be polymorphic in Europeans (HapMap) to define our polymorphic sites, and restricted the analysis to the non-recombining portion of the X chromosome. We only considered bases with

quality ≥20 and reads with mapping quality ≥30. ANGSD applies two different methods to estimate the contamination rate: Method 1 considers all bases, providing greater power, while assuming that errors are independent between reads and sites. However, Method 2 does not have this bias, as it randomly samples a single read at each site, but is also less precise than Method 1. Contamination estimates are shown in Supplementary Table S5. Low X chromosome contamination estimates (<3.6%) were observed in all samples used in this study (median=0.64%).

Genetic sex determination

We used the ratio of reads mapping to the Y and X chromosomes to determine the sex of each sample as described in Skoglund et al. (*17*). To do so, we calculated R_y, the fraction of reads mapping to the Y chromosome compared to the total number of reads mapping to both the Y and the X chromosomes. We used this value to assign each sample XX or XY karyotypes. Results are shown in Fig. S3A and Table S6.

S4. Population genomics

Reference population datasets

We compiled two main autosomal reference datasets of microarray SNP data for the purpose of analyzing the ancestry of the ancient Icelanders. We removed loci and individuals with <95% call rate, and loci that failed Hardy-Weinberg Equilibrium exact tests with p-value <0.0001. We ran ADMIXTURE (*53*) with individuals of Asian and African ancestry and filtered out individuals with <90% European ancestry. Finally, the reference datasets were filtered for related individuals, defined as pairs with IBD values >0.06. The remaining SNPs were subsequently pruned for loci on three previously reported long range LD regions on chromosomes 6, 8 and 11 using PLINK (*54*).

Our European reference dataset contained genotypes for 227,056 loci (the intersection of several different Illumina chip types) from 2,139 contemporary individuals sampled from 29 European populations. Of these, 1166 individuals were from previously published datasets (*55-59*), with the remainder from this study (Table S7). Our second reference dataset contained 404,066 SNPs from 5,470 contemporary individuals from the North Atlantic. Of these, 1,246 were individuals from the United Kingdom and Ireland from the previously published People of the British Isles (POBI) dataset (*60*) (Table S7). We selected individuals from the POBI dataset that had been genotyped on both the Illumina HumanHap 1M chip and Affymetrix 500K to maximize the intersection of SNPs with the remaining samples (n=4,224), which were all genotyped on the Illumina OmniExpress chip. We included 190 Irish individuals from the Irish DNA Atlas dataset(*61*), a collection of individuals who have all eight great-grandparents born within 50 km of each other in different counties of Ireland. Saliva samples were obtained (Oragene OG-250, DNA Genotek, Canada), and extracted according to standard protocol, before genotyping on the Illumina OmniExpress chip at Edinburgh Genomics. The remaining samples (n=4,034) were genotyped on the Illumina OmniExpress chip at deCODE Genetics. The Danish samples (n=980) were recruited in 2011 among blood donors in the area of Greater Copenhagen, Denmark. All participants were healthy with no known illness, for the majority (85%), parents and grandparents were born in Denmark. The Swedish samples (n=1,103) are population-based controls collected as part of the case-control study "epidemiology of multiple sclerosis (EIMS)" (*62*). The Norwegian samples (n=1,035) are unrelated controls from an ongoing study of heritable cancer genes, who were sampled from all the major regions of Norway. The Icelandic samples (n=916) were randomly selected from a larger set of individuals genotyped on the Illumina OmniExpress chip as a part of ongoing research at deCODE into the genetic basis of human phenotypic variation. All sample identifiers were encrypted in accordance with the regulations of the Icelandic Data Protection Authority.

The genotype data from Norway, Sweden and Denmark are from controls used in ongoing GWAS studies, and as such are subject to some access restrictions. Further information can be obtained from Eivind Hovig (ehovig@ifi.uio.no) regarding the Norwegian data, Thomas Werge regarding the Danish data (Thomas.Werge@regionh.dk) and Tomas Olsson (Tomas.Olsson@ki.se) regarding the Swedish data. Icelandic law and the regulations of the Icelandic Data Protection Authority prohibit the release of individual-level and personally identifying data. Access to the genotype data from contemporary Icelanders is contigent on an application to deCODE Genetics and travel to its Icelandic facilities.

Y chromosome analysis

We used the ISOGG 2016 Y chromosome tree (*63*) to assign haplogroups to the 22 ancient Icelanders identified as males, along with 3,018 males from the North European reference data set (main Table 1 and Table S10). The ISOGG 2016 database is based on 13,793 loci, each of which marks a mutation on the human Y chromosome tree. For each SNP in the dataset, we determined whether our ancient sample carried the derived or ancestral allele (according to majority alleles with base quality >30), and assigned each individual to the branch in the tree most consistent with the overall configuration of derived alleles observed.

Principal components analysis

For the two reference data sets of contemporary individuals described above, PCAs were computed using the 'smartpca' (*64*) program from the EIGENSOFT v.6.01 package. The PC weights for each locus were then used to calculate projected eigenvector values based on haploid genotypes from the ancient individuals (*20, 64*). Using the R package "vegan" (http://vegan.r-forge.r-project.org/)., we also used Procrustes transformation (*65*) to combine separate PCA analyses for each ancient individual against a given reference set of contemporary individuals.

D-statistics test

We used D-statistics to explore the patterns of shared genetic drift between ancient and contemporary individuals in our data set. The statistics were calculated with the ADMIXTOOLS package v. 3.0 (*29*), based on allele frequencies using the estimators described previously. Significant deviation from zero can be interpreted as rejection of the tree population typology ((Outgroup, X);(Pop1, Pop2)). Under the assumption that no gene flow occurred between Pop1 and Pop2 and the outgroup, a positive D-statistic suggests affinity between X and Pop2, whilst a negative value indicates affinity between X and Pop1. Standard errors were obtained using a block jackknife 5 cM in size, and D-statistic values were considered significantly different from 0 when absolute Z-scores were >3 (*26, 66, 67*).

Admixture

To assess the different proportions of Norse, Celtic and Icelandic ancestry in our ancient individuals, we used ADMIXTURE v.1.3 (*53*). Unsupervised ADMIXTURE analysis for various values of K did not yield clear clusters corresponding to Celts, Norse and Icelanders. This is most likely due to the fact that the Norse and Celtic populations are very closely related, and that the Icelanders are both admixed and separated from their source populations by genetic drift. We therefore resorted to a supervised approach (*68*), where we estimated admixture proportions for the ancient Icelandic individuals using Norse and Celtic reference populations with K=2, and Norse, Celtic and Icelandic references with K=3. These analyses were run on haploid data. We also estimated admixture proportions for contemporary Icelanders (*n*=916) using ADMIXTURE in supervised mode with Norse and Celtic reference populations (K=2).
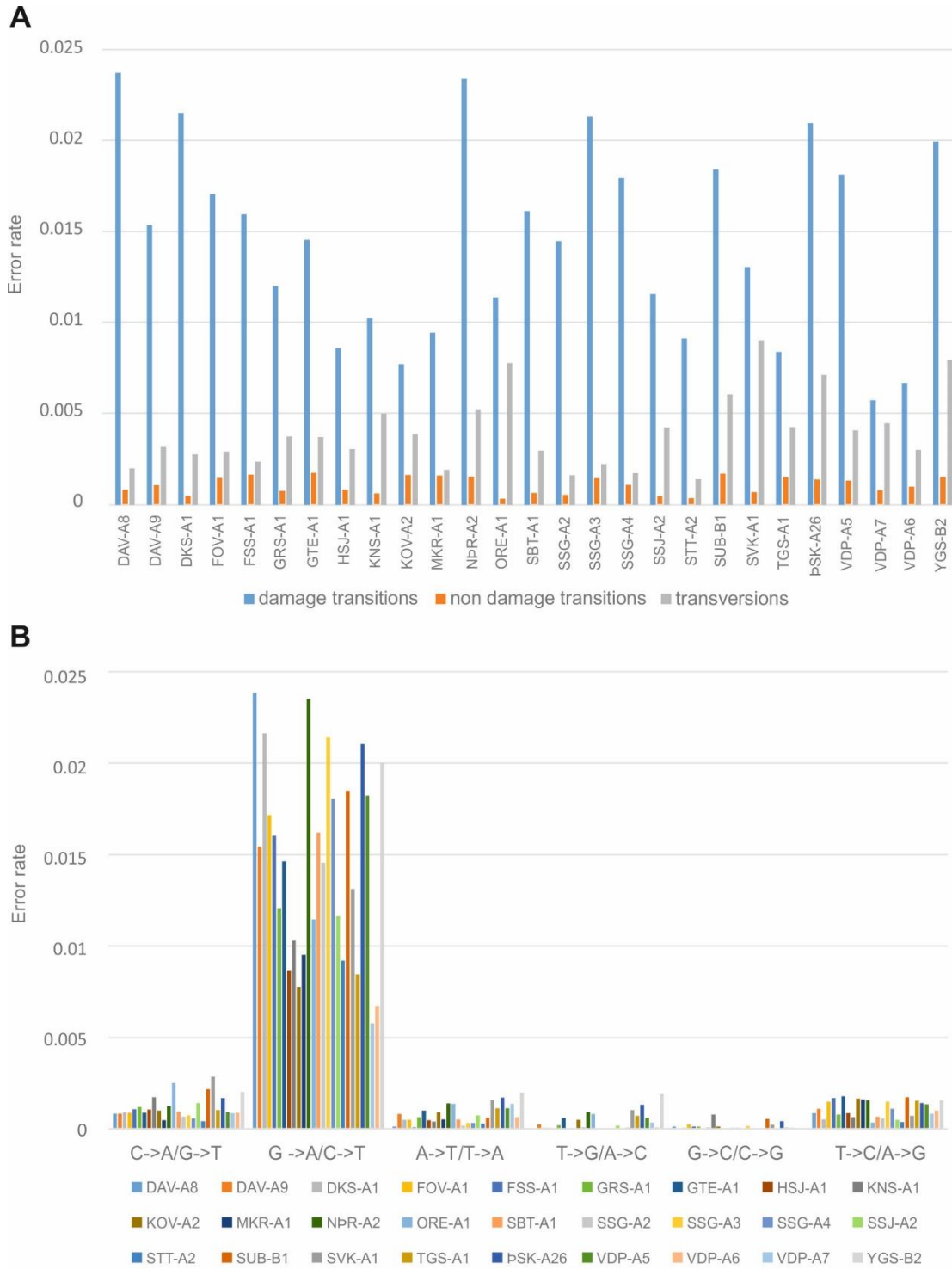
Fig. S1. **Error rate estimates for each ancient sample used in this study.** (**A**) Error rate per sample using an outgroup and high quality genome. (**B**) Type-specific error rate using an outgroup and high quality genome.

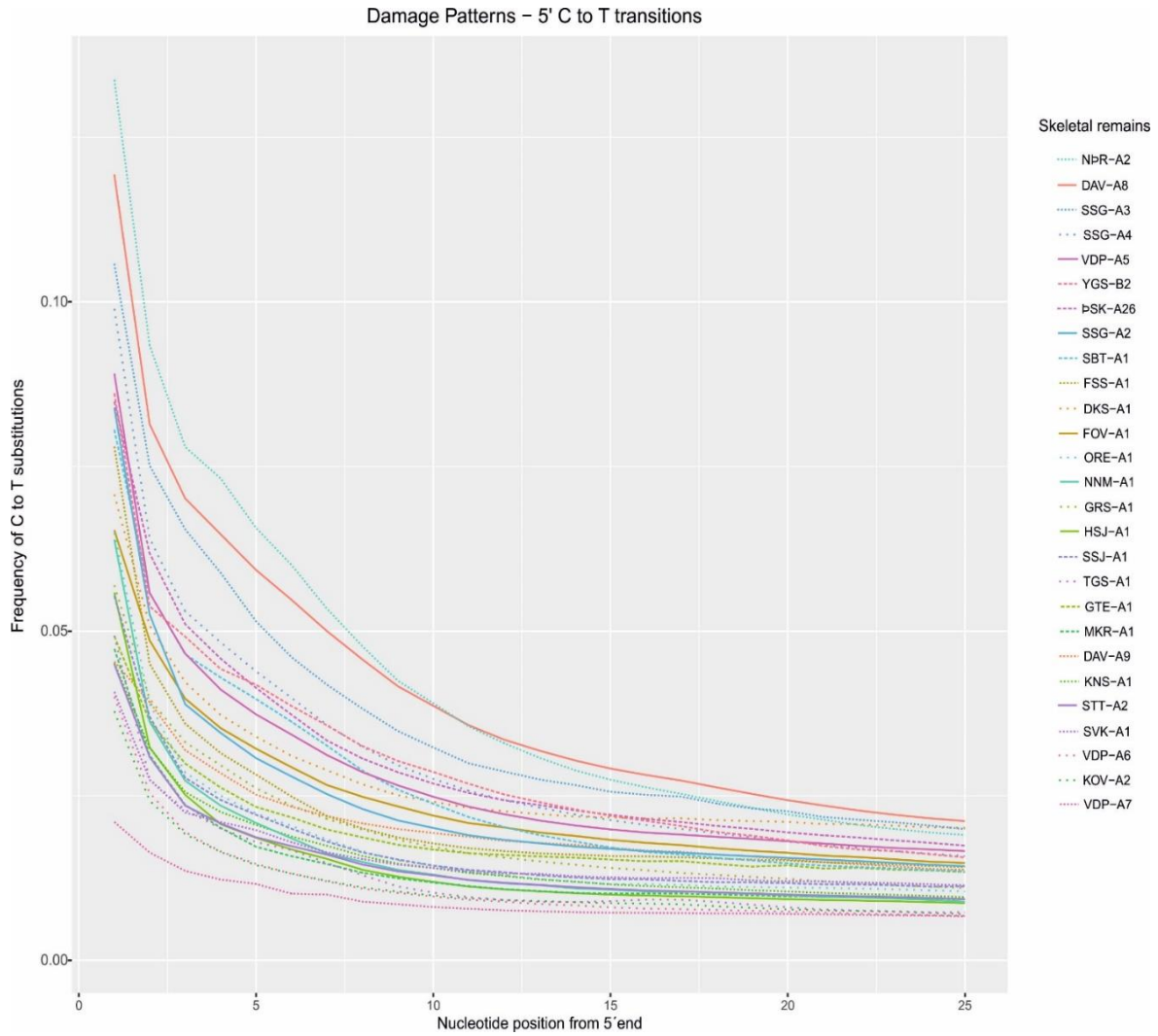**Damage Patterns − 5' C to T transitions**

Fig. S2**. Ancient DNA damage**. aDNA damage parameters were estimated using the Bayesian approach implemented in mapDamage 2.0 (*47*). The line chart records the average C to T transitions for the first 25 bases at the 5´end of reads. Samples are arranged in order of damage patterns at first position, from highest to lowest.
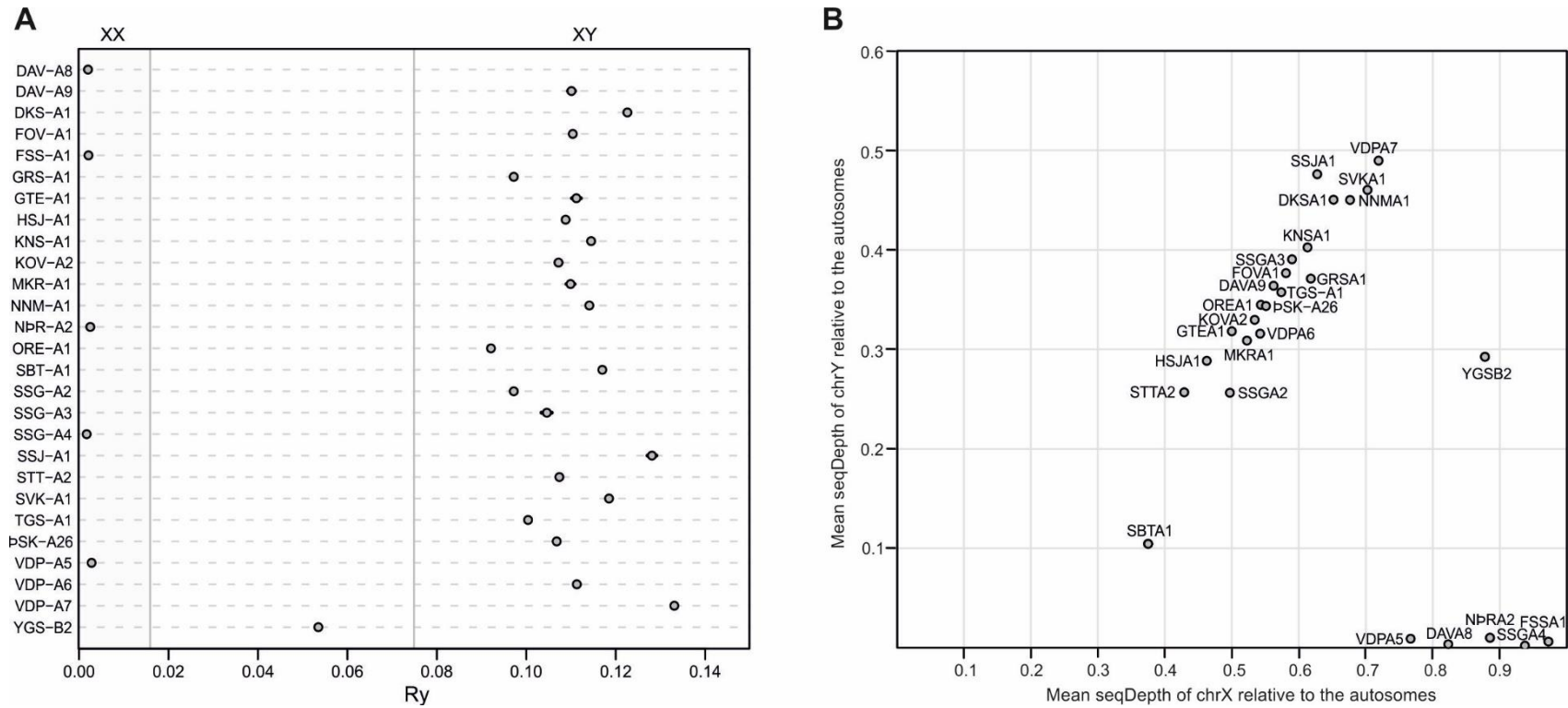
Fig. S3. **Sex determination based on WGS data.** (**A**) R_y - ratio of the number of reads aligning to the Y chromosome compared to the number of reads mapping to both sex chromosomes. Only reads with mapping quality above 30 were considered. Error bars correspond to 95% CI. Samples were classified as female when the upper bound of this CI was <0.016 (left vertical line), and classified as male when the lower bound was >0.075 (right vertical line). (**B**) Scatterplot showing the average sequence depth for X and Y chromosomes, scaled by the average autosomal sequence depth for each sample. For females, X chromosome coverage is expected to be similar to that of autosomes, whereas Y chromosome coverage should be minimal. For males, the coverage of both X and Y chromosomes should be half of the autosomes. Males and females, as defined by R_y, are clearly separated into distinct clusters, but YGS-B2 falls within the range of males for Y chromosome depth and within the range of females for X chromosome depth.
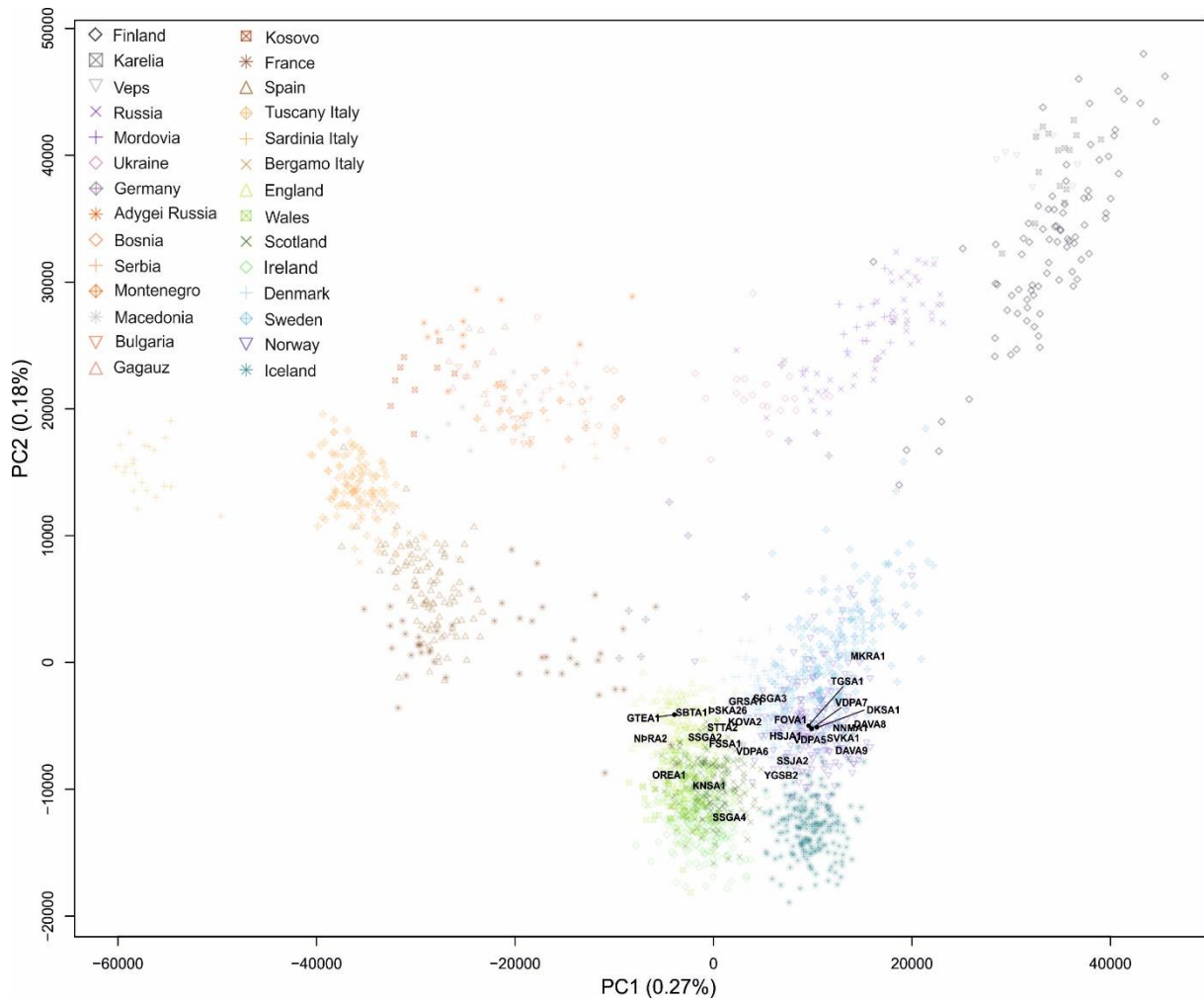
Fig. S4. **PCA projection**. Ancient Icelanders projected onto the first two eigenvectors of a PCA based on contemporary individuals from 28 European populations (*n*=2,139) and 227,056 autosomal SNPs. Variance explained by the first two components is shown in brackets.
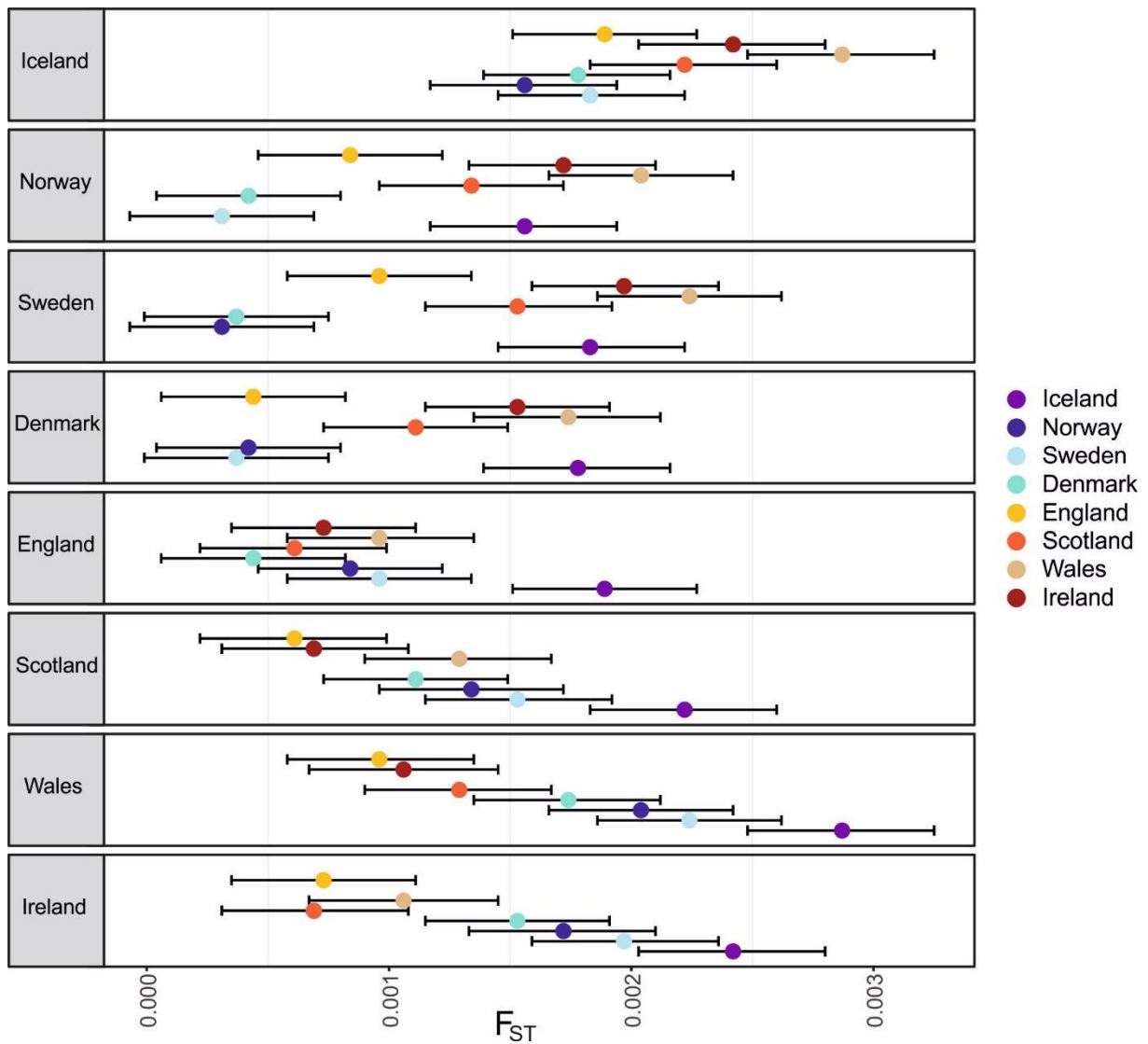
Fig. S5. **The average F$_{ST}$ distances**. Calculated between each pair of contemporary individuals from Iceland, Scandinavia and the British Isles using diploid genotypes from 404,066 autosomal SNPs.
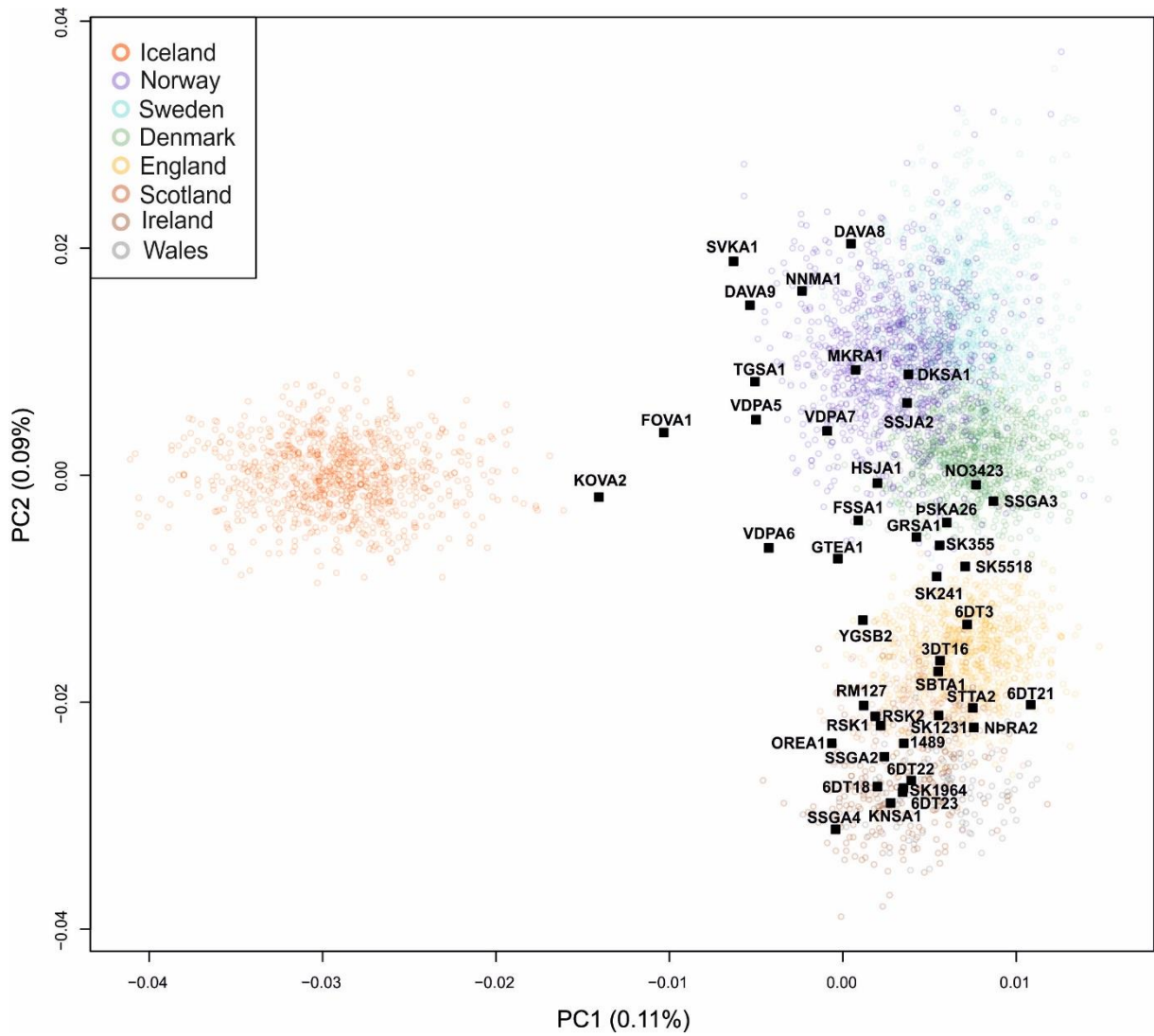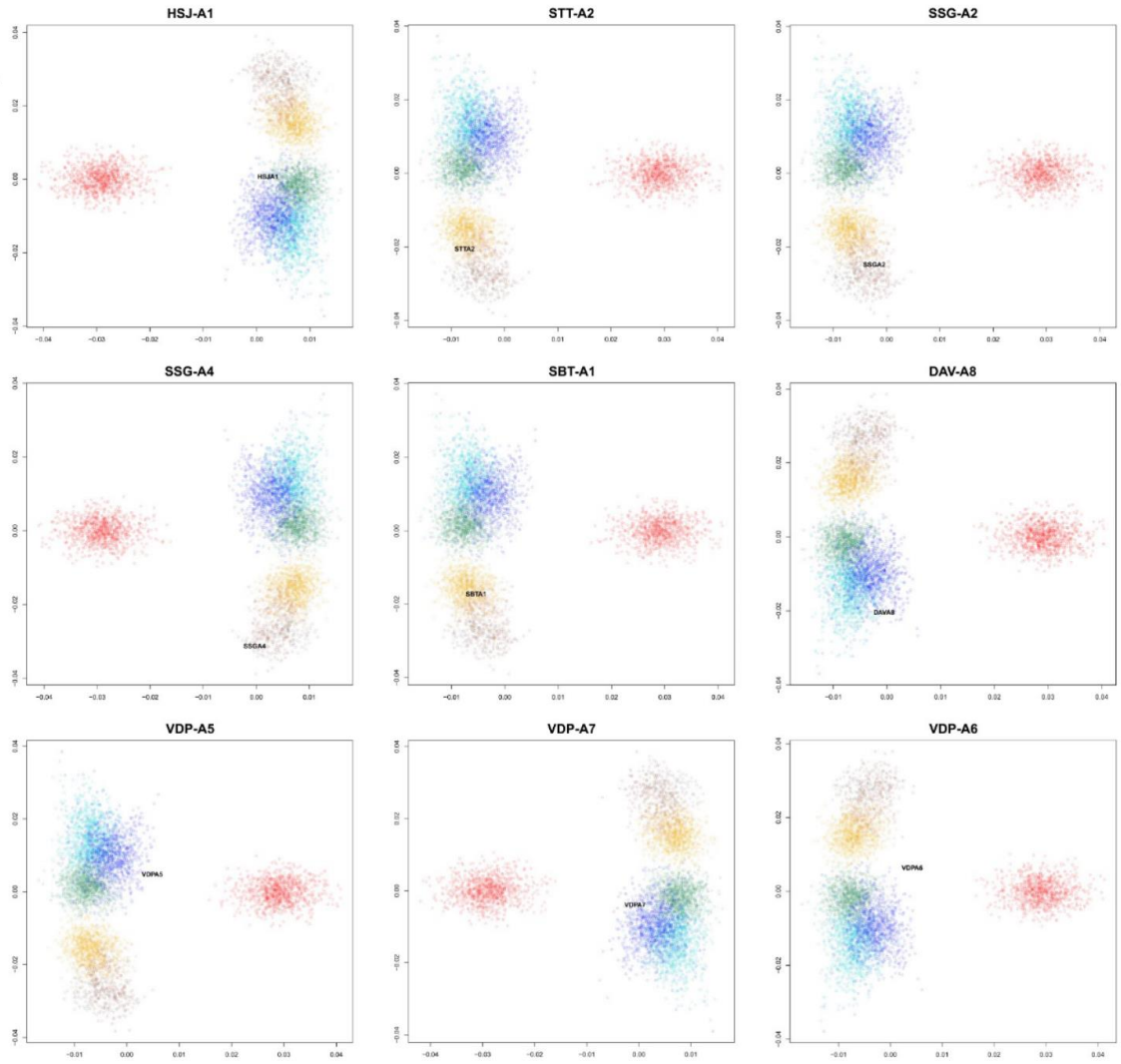
Fig. S6. **Procrustes transformation.** The coordinates for each ancient sample (*n*=43) were transformed to match the first two eigenvectors based on contemporary individuals from Scandinavia (*n*=3,118), the British Isles (*n*=1,436) and Iceland (*n*=916).Variance explained by the first two components is shown in brackets.
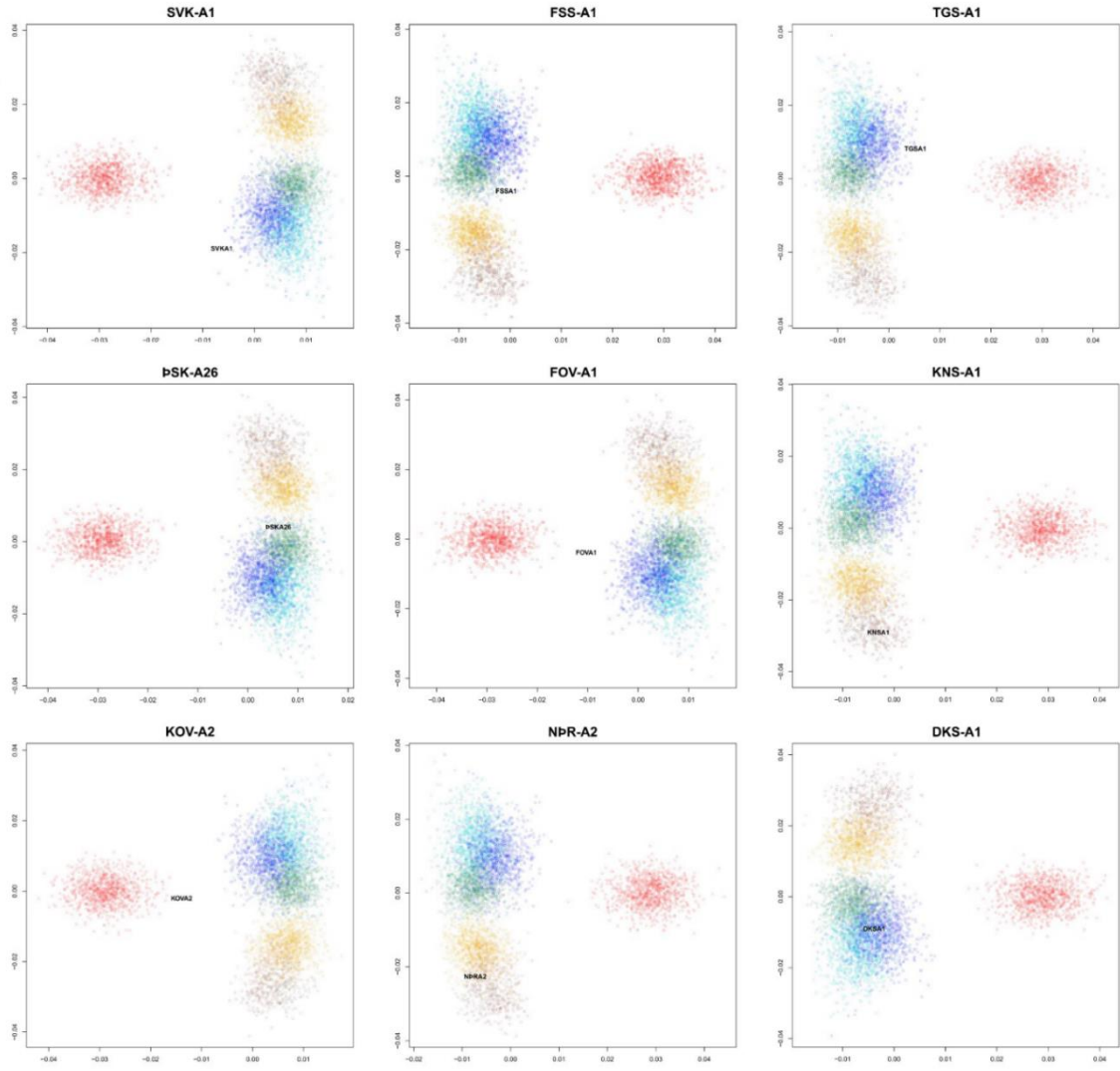
A

Iceland
Norway
Sweden
Denmark
England
Scotland
Ireland
Wales

HSJ-A1   STT-A2   SSG-A2
SSG-A4   SBT-A1   DAV-A8
VDP-A5   VDP-A7   VDP-A6

B

Legend: Iceland, Norway, Sweden, Denmark, England, Scotland, Ireland, Wales

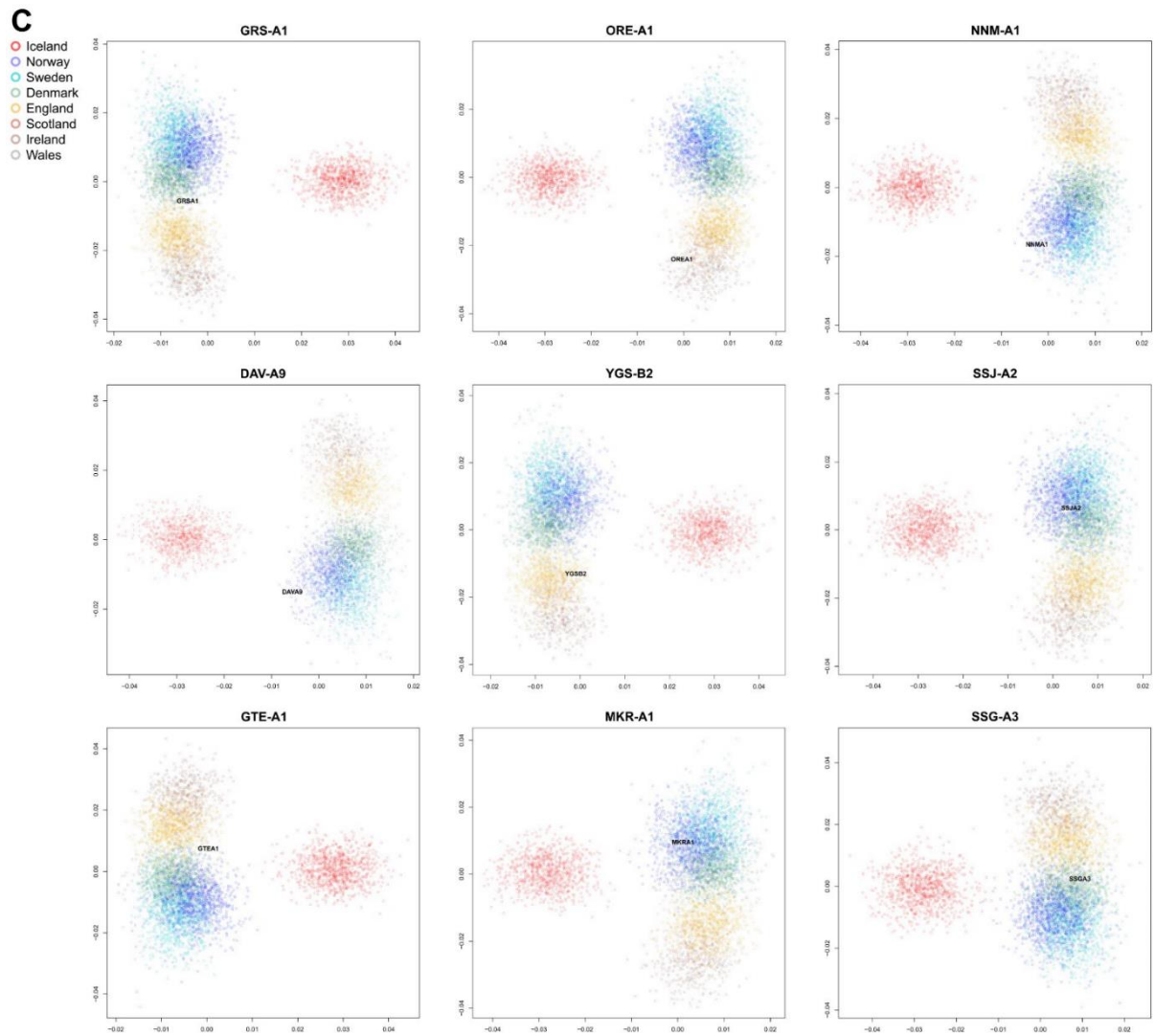SVK-A1, FSS-A1, TGS-A1, ÞSK-A26, FOV-A1, KNS-A1, KOV-A2, NÞR-A2, DKS-A1

Fig. S7. **PCAs of individual samples.** Based on overlapping SNPs after merging each ancient individual with the North European reference data. Figures (**A**), (**B**) and (**C**) are sorted by percentage of overlapping SNPs (see main Table 1), from highest to lowest.
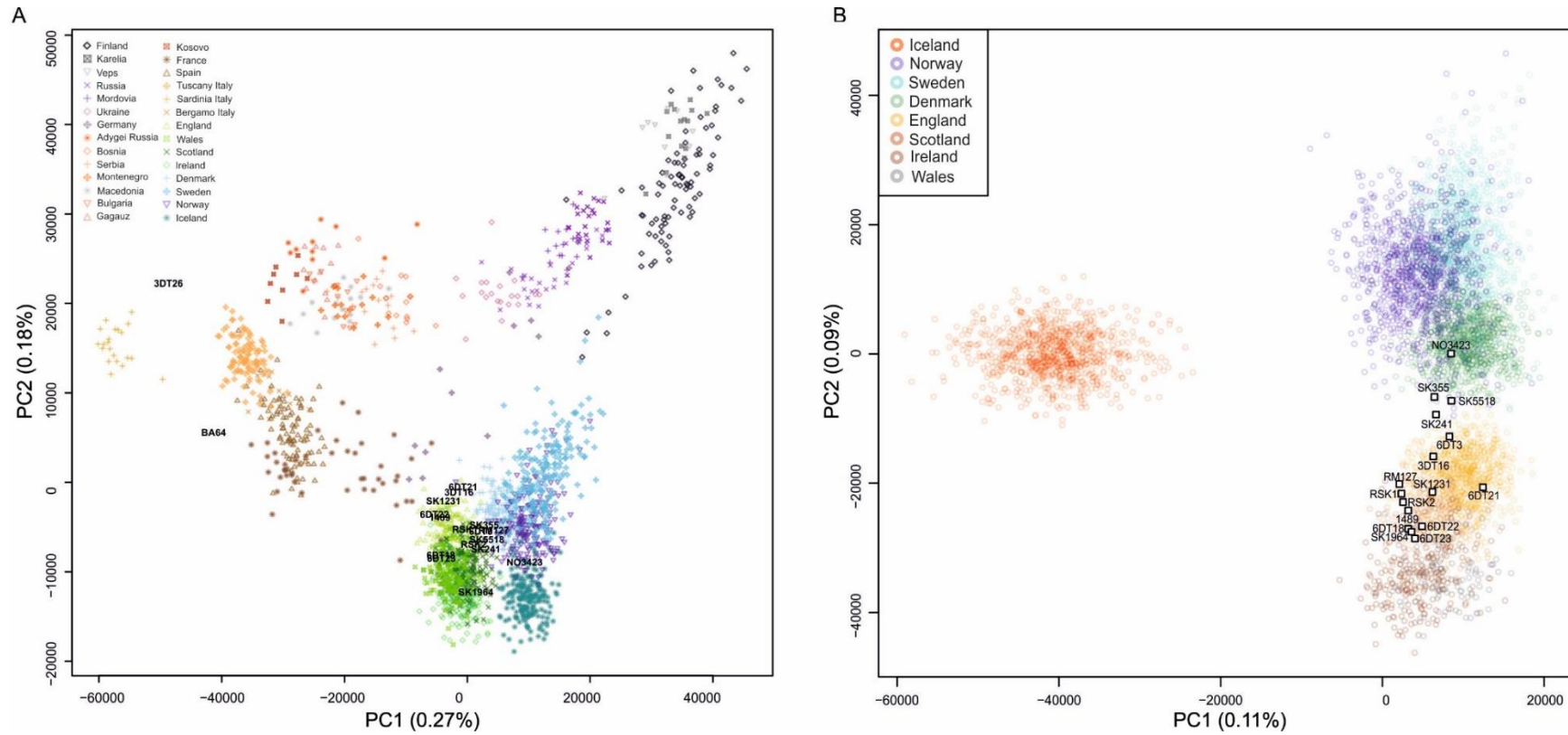
Fig. S8. **PCAs showing the relationship between previously published ancient genomes from the British Isles and contemporary populations.** (**A**) Ancient British and Irish genomes (*n*=18) projected onto the first two eigenvectors of a PCA based on contemporary individuals from 28 European populations (*n*=2,139) using diploid genotypes from 227,056 autosomal SNPs. (**B**) Ancient British and Irish genomes plotting within the Northern Europe cluster in Fig. S8A (*n*=6) projected onto the first two eigenvectors of a PCA based on contemporary individuals from Scandinavia (*n*=3,118), the British Isles (*n*=1,436) and Iceland (*n*=916) using diploid genotypes from 404,066 autosomal SNPs.

Supplementary tables

Table S1. **Overview of all ancient skeletal remains screened for this study.**

Time period	Classification based on ⁸⁷Sr/⁸⁶Sr value (3, 4)	Sample ID	Grave goods	Archaeo-logical date	¹⁴C date	Typology date[1]	Archaeological sex[2]	Age or age class[2]	N†	Y‡	Site	Region
Pre-Christian	Migrant	DAV-A7*	yes	<1000	<1050	900-1000	Male	< 35	13	1909	Dalvík (Brimnes)	North
		DAV-A8	yes	<1000	<1050	900-1000	Female?	25-34	13	1909	Dalvík (Brimnes)	North
		DAV-A9	yes	<1000	980-1020	900-1000	Female?	18-25	13	1909	Dalvík (Brimnes)	North
		HBS-A6*	yes	<1000	NA	900-1000	Male	35-45	7-8	1947	Hafurbjarnarstaðir	South
		KRE-A1*	yes	<1000	NA	900-1000	Male	35-45	2	1900	Kroppur	North
		SSG-A1*	yes	<1000	980-1020	850-1000	Male	45+	4	1947	Sílastaðir	North
		SSG-A2	yes	<1000	NA	850-1000	Male	45+	4	1947	Sílastaðir	North
		SSG-A3	yes	<1000	NA	850-1000	Male	35-45	4	1947	Sílastaðir	North
		SSG-A4	yes	<1000	NA	850-1000	Female?	35-45	4	1947	Sílastaðir	North
	Non-migrant	DKS-A1	yes	<1000	NA	850-1000	Unknown	18-25	1	1962	Öndverðarnes	West
		EIM-A1*	yes	<1000	NA	900-1000	Female	25-35	1	1979	Einholt	East
		GRS-A1	yes	<1000	<1050	NA	Male	35-45	3	1937	Grímsstaðir	North
		GTE-A1	yes	<1000	NA	NA	Male	18-25	1	1957	Gilsárteigur	East
		HSJ-A1	yes	<1000	NA	NA	Male	35-45	4	1996	Hrólfsstaðir	East
		KNS-A1	yes	<1000	NA	950-1050	Male	35-45	1	1932	Karlsnes	South
		MKR-A1	no	<1000	<1050	NA	Male	35-45	1	1989	Víðar (Másvatn)	North
		NNM-A1	no	<1000	NA	NA	Unknown	Adult	1	1981	Njarðvík	East

		Sample					Sex	Age	†	‡	Site	Region
		ORE-A1	yes	<1000	NA	900-1000	Male	45+	1	1966	Ormsstaðir	East
		SBT-A1	yes	<1000	NA	NA	Male	45+	1	1954	Smyrlaberg	North
		SSJ-A2	yes	<1000	NA	850-1000	Male	35-45	2	1949	Surtsstaðir	East
		STT-A2	yes	<1000	975-1015	900-1000	Male	45+	1	1952	Straumur	East
		SVK-A1	yes	<1000	<1050	NA	Male	35-45	1	1961	Svínadalur	North
		VDP-A3*	yes	<1000	NA	850-1050	Male	18-25	7	1964	Vatnsdalur	West
		VDP-A5	yes	<1000	NA	850-1050	Female	35-45	7	1964	Vatnsdalur	West
		VDP-A6	yes	<1000	NA	850-1050	Male	25-35	7	1964	Vatnsdalur	West
		VDP-A7	yes	<1000	NA	850-1050	Male	35-45	7	1964	Vatnsdalur	West
	Not available	FSS-A1	no	<1000	NA	NA	Female	35-45	1	1960	Fremri-Svartárdalur	North
		GRF-A1*	yes	<1000	723-1122	NA	Unknown	7-12	3	1962	Grímsstaðir 'the old'	North
		NÞR-A2	no	<1000	NA	NA	Female?	Adult?	2	1952	Neðri-Þverá	South
		SUB-A1*	no	<1000	683-970	NA	Male	Adult	2	1947	Suðurárbotnar	North
		YGS-B2	yes	<1000	NA	NA	Unknown	7-12	9	1954	Ytra-Garðshorn	North
Early Christian	Non-migrant	ÞSK-A26	NA	1000-1104	1120	NA	Male	Adult	NA	1939	Skeljastaðir	South
17th Century	Not available	KOV-A2	NA	b. 1678	NA	NA	Male	26	2	1988	Kópavogur	West
Unknown		FOV-A1	no	Unknown	NA	NA	Male	Adult	1	1929	Fossvellir	East
		TGS-A1	no	Unknown	NA	NA	Male	Adult	1	1981	Tunga	North

Samples marked with an asterisk symbol (*) were ultimately excluded from further analysis. †Number of remains found at site. ‡Year of exhumation. [1] (*11, 12*) [2](*69*)

Table S2. **Overview of extractions, library builds and sequencing done on ancient genomes analysed in this study**. We constructed 61 libraries from 50 DNA extractions of 35 ancient Icelandic skeletal remains. Library names indicate which DNA extraction method was used: phenol/chloroform purification method (P) or the modified silica extraction method (S).

Sample	Library	Pre WGE (no. of HiSeq lanes)	Post WGE (no. of HiSeq lanes)
			MiSeq - endogenous (endo) %
DAV-A8	25_P1	59.16	60.46 (1)
DAV-A8	25_S1	29.72	51.14 (2)
DAV-A9	26_P1	2.04	30.25 (2)
DAV-A9	26_P2	2.07	21.77 (1)
DKS-A1	27_P1	NA	16.36 (1)
DKS-A1	27_S1	1.64	22.99 (2)
DKS-A1	27_1	0.22 (1)	NA
FOV-A1	32_P1	7.31	23.51 (1)
FOV-A1	32_P2	5.47	33.78 (1)
FSS-A1	34_P1	40.94	58.22 (1)
GRS-A1	38_P1	8.43	31.83 (2)
GRS-A1	38_P2	5.86	31.1 (1)
GTE-A1	44_P1	6.62	32.65 (2)
HSJ-A1	2_P1	29.08	50.6 (4)
HSJ-A1	2_P2	33.04	53.2 (2)
HSJ-A1	39_P1	35.26	47.37 (2)
HSJ-A1	39_S1	60.14	50.26 (4)
KNS-A1	52_S1	6.01	32.95 (2)
KOV-A2	3_P1	8.67	26 (1)
MKR-A1	49_P1	4.44	27.52 (1)
NNM-A1	50_P1	6.72	32.89 (1)
NÞR-A2	63_P1	13.83	29.1 (1)
ORE-A1	56_S1	22.43	28.57 (1)
SBT-A1	66_P1	18.11	62.03 (1)
SBT-A1	66_S1	54.09 (2)	64.15 (2)
SSG-A2	67_P2	65.51 (4)	37.07
SSG-A2	67_S1	29.29	50.45 (4)
SSG-A3	58_P2	2.7	21.32 (1)
SSG-A4	63_P1	44.82	57.3 (2)
SSG-A4	63_P2	38.47	53.54 (2)
SSG-A4	63_S1	31.35	56.92 (2)
SSJ-A2	77_S1	4.41	36.45 (1)
STT-A2	74_P1	23.46	40.44 (1)
STT-A2	74_S1	43.8	55.96 (4)
SVK-A1	70_S1	6.24	32.5 (1)
TGS-A1	8_P1	NA	61.88 (1)
VDP-A5	97_P1	30.21	40.94 (5)
VDP-A5	97_S1	11.24	49.49 (1)
VDP-A6	101_P1	1.99	26.27 (1)
VDP-A6	101_S1	14.37	36.39 (2)
VDP-A7	94_P1	8.81	31.85 (1)
VDP-A7	94_S1	10.07	63.44 (2)
YGS-B2	92_P1	NA	41.02 (1)
YGS-B2	92_1	0.36 (1)	NA
ÞSK-A26	9_P1	15.4	43.49 (2)
Libraries excluded based on Bioanalyzer results			
DAV-A7	24_S1	NA	NA
HBS-A6 *	42_P1	NA	NA
KRE-A1 *	53_P1	NA	NA
SSG-A2	67_P1	NA	NA
SSG-A3	58_P1	NA	NA
VDP-A3 *	88_P1	NA	NA
Libraries excluded based on MiSeq results (endo <15%)			
DAV-A7 *	24_P1	1.26	12.49
DAV-A7 *	24_P2	1.81	WGE failed

DKS-A1	27_P2	0.96	NA
GRF-A1 *	55_S1	6.96	WGE failed
KNS-A1	52_P1	0.85	NA
MKR-A1	49_P2	2.92	WGE failed
SSG-A1 *	68_P1	NA	14.24
Libraries excluded based on contamination results (Table S4)			
GTE-A1	44_S1	6.5	35.42 (2)
SUB-B1 *	100_P1	35.7	47.76 (1)
Libraries excluded based on sequencing depth/coverage (Table S3)			
EIM-A1 *	28_P1	3.29	19.47 (1)

Samples marked with an asterisk symbol (*) were ultimately excluded from further analysis.

Table S3. **Details and statistics for all libraries (*n*=49) sequenced on the Illumina HiSeq platform. Trimmed reads** is the total number of reads passing quality and length filtering. **Mapped reads** is the number of reads mapping to the human genome. **Endo %** is the percentage of reads mapped to the human genome. **Endo % without dup** is the percentage of reads mapped to the human genome after removing duplicate reads. **Dup %** is the percentage of duplicates reads after mapping to the human genome. **Final reads** indicate the number of reads mapped to the human genome without duplicates. **Read length** indicates the average read length of final reads. **DoC (X)** indicates the mean autosomal depth of coverage. **Auto. Cov.** shows the fraction of the autosomal genome that is covered by at least one read.

Sample	Library ID	WGE	Trimmed reads	Mapped reads	Endo %	Trimmed reads	Final reads	Endo % without dup	Dup %	Read length	DoC (X)	Auto. Cov.
DAV-A8	25_P1	Yes	269547770	201283564	74.67	197078840	128814634	65.36	36	80	1.1	0.76
DAV-A8	25_S1	Yes	261186173	159838243	61.2	214636371	113288441	52.78	29.1	77	1.17	0.78
DAV-A8	25_S1	Yes	297520712	185583494	62.38	239432152	127494934	53.25	31.3	77	1.22	0.81
DAV-A9	26_P1	Yes	282633784	133832869	47.35	170246657	21445742	12.6	84	89	0.21	0.14
DAV-A9	26_P2	Yes	305892099	95050113	31.07	233805417	22963431	9.82	75.8	79	0.21	0.15
DAV-A9	26_P1	Yes	281111011	98803057	35.15	209527794	27219840	12.99	72.5	91	0.39	0.24
DKS-A1	27_1	No	363194526	2027247	0.56	362707894	1540615	0.42	24	104	0.03	0.02
DKS-A1	27_P1	Yes	229253823	70640728	30.81	167944567	9331472	5.56	86.8	99	0.1	0.06
DKS-A1	27_S1	Yes	309244488	104089571	33.66	238452900	33297983	13.96	68	121	0.59	0.38
DKS-A1	27_S1	Yes	295343683	98850875	33.47	235686371	39193563	16.63	60	126	0.59	0.38
EIM-A1	28_P1*	Yes	287039389	110406182	38.46	192088689	15455482	8.05	86	92	0.12	0.08
FOV-A1	32_P2	Yes	266981327	115005525	43.08	190930150	38954348	20.4	66.1	80	0.4	0.3
FOV-A1	32_P1	Yes	288740067	95669302	33.13	230146760	37075995	16.11	61.2	87	0.58	0.38
FSS-A1	34_P1	Yes	262581240	187568997	71.43	163251215	88238972	54.05	53	76	0.94	0.66
GRS-A1	38_P1	Yes	168873197	71706374	42.46	118639923	21473100	18.1	70.1	69	0.3	0.19
GRS-A1	38_P1	Yes	262428188	110482391	42.1	180829038	28883241	15.97	73.9	69	0.31	0.2
GRS-A1	38_P2	Yes	296394594	125513294	42.35	214698912	43817612	20.41	65.1	83	0.47	0.35
GTE-A1	44_P1	Yes	203843887	110783054	54.35	111842242	18781409	16.79	83	94	0.31	0.19
GTE-A1	44_P1	Yes	289159138	154556495	53.45	158926635	24323992	15.31	84.3	95	0.32	0.2
GTE-A1	44_S1*	Yes	236774327	127186947	53.72	176271023	66683643	37.83	47.6	114	0.96	0.66
GTE-A1	44_S1*	Yes	283426244	153019255	53.99	201043785	70636796	35.14	53.8	109	1.01	0.69
HSJ-A1	2_P2	Yes	251817854	168561801	66.94	197083791	113827738	57.76	32.5	104	0.95	0.73

49

HSJ-A1	2_P2	Yes	302593056	201657118	66.64	224160436	123224498	54.97	38.9	100	1.01	0.76
HSJ-A1	39_P1	Yes	226083722	143683836	63.55	175506146	93106260	53.05	35.2	102	1.39	0.85
HSJ-A1	39_P1	Yes	226874875	143799561	63.38	176063126	92987812	52.82	35.3	102	1.39	0.85
HSJ-A1	2_P1	Yes	262411059	172156230	65.61	206666814	116411985	56.33	32.4	113	1.5	0.91
HSJ-A1	2_P1	Yes	265893809	173936214	65.42	209391653	117434058	56.08	32.5	113	1.86	0.91
HSJ-A1	2_P1	Yes	287039398	188900584	65.81	212725934	114587120	53.87	39.3	108	2.7	0.92
HSJ-A1	39_S1	Yes	258276576	195101589	75.54	230631102	167456115	72.61	14.2	118	4.39	0.99
HSJ-A1	39_S1	Yes	258865031	195569312	75.55	231004376	167708657	72.6	14.2	118	4.39	0.99
HSJ-A1	39_S1	Yes	275578087	209212979	75.92	244446390	178081282	72.85	14.9	114	4.57	0.99
HSJ-A1	39_S1	Yes	277565387	210155661	75.71	246316729	178907003	72.63	14.9	114	4.58	0.99
HSJ-A-1	2_P1	Yes	288594731	189609855	65.7	214257463	115272587	53.8	39.2	108	2.7	0.92
KNS-A1	52_S1	Yes	276846239	126459508	45.68	198299728	47912997	24.16	62.1	103	0.72	0.48
KNS-A1	52_S1	Yes	313445420	137826293	43.97	223327861	47708734	21.36	65.4	98	0.73	0.48
KOV-A2	3_P1	Yes	127306063	127306063	37.09	267735210	51844301	19.36	59.3	83	0.69	0.45
MKR-A1	49_P1	Yes	338463336	169673733	50.13	194788951	25999348	13.35	84.7	102	0.23	0.16
NNM-A1	50_P1	Yes	212025629	88107388	41.56	169053463	45135222	26.7	48.8	97	0.51	0.38
NÞR-A2	62_P1	Yes	118916533	118916533	36.84	254251853	50381029	19.82	57.6	70	0.64	0.43
ORE-A1	56_S1	Yes	232416772	79568411	63.91	48563478	29196363	60.12	63.3	119	0.48	0.31
SBT-A1	66_S1	Yes	246712167	201207825	81.56	161584163	116079821	71.84	42.3	116	0.96	0.73
SBT-A1	66_S1	Yes	292536776	241197984	82.45	125998969	125998969	71.05	47.8	108	1	0.75
SBT-A1	66_S1	No	208402396	133191914	63.91	197557009	122346527	61.93	8.1	98	3.69	0.94
SBT-A1	66_P1	Yes	232684639	154280574	66.3	219449833	141045768	64.27	8.6	89	3.73	0.95
SBT-A1	66_S1	No	217957154	141317196	64.84	205882343	129242385	62.77	8.5	89	3.9	0.95
SSG-A2	67_P2	No	273323781	51511862	18.85	269841029	48029110	17.8	6.7	75	1.14	0.67
SSG-A2	67_P2	No	282549156	55378994	19.6	278554045	51383883	18.45	7.2	75	1.2	0.7
SSG-A2	67_P2	No	279751689	58656670	20.97	273834435	52739416	19.26	10.1	76	1.26	0.73
SSG-A2	67_P2	No	309035007	65039734	21.05	303628198	59632925	19.64	8.3	75	1.39	0.77
SSG-A2	67_S1	Yes	245090998	156524811	63.86	208851482	120285295	57.59	23.2	95	1.4	0.89
SSG-A2	67_S1	Yes	247752596	158105644	63.82	210835097	121188145	57.48	23.3	95	1.43	0.89
SSG-A2	67_S1	Yes	283443406	188708599	66.58	235622612	140887805	59.79	25.3	93	2.72	0.92

SSG-A2	67_S1	Yes	295423134	191226474	64.73	142496799	142496799	57.76	25.5	93	2.74	0.93
SSG-A3	58_P2	Yes	291970190	81214988	27.82	240288191	29532989	12.29	63.6	83	0.31	0.22
SSG-A4	63_S1	Yes	270354863	211756142	78.33	160144559	101545838	63.41	52	98	0.85	0.65
SSG-A4	63_S1	Yes	269030789	211431179	78.59	159071895	101472285	63.79	52	100	0.85	0.65
SSG-A4	63_P1	Yes	216706160	138219100	63.78	169804761	91317701	53.78	33.9	65	1.19	0.72
SSG-A4	63_P1	Yes	361511422	250439573	69.28	244972090	133900241	54.66	46.5	64	1.38	0.82
SSG-A4	63_P2	Yes	285994792	189371075	66.21	222165362	125541645	56.51	33.7	84	1.3	0.85
SSG-A4	63_P2	Yes	287085611	189400879	65.97	223737064	126052332	56.34	33.4	88	1.3	0.85
SSJ-A2	77_S1	Yes	98953193	51993207	52.54	71042878	24082892	33.9	53.7	125	0.4	0.27
STT-A2	74_P1	Yes	279871003	153982319	55.02	182182703	56294019	30.9	63.4	90	0.82	0.52
STT-A2	74_S1	Yes	227525447	167042213	73.42	200290765	139807531	69.8	16.3	112	2.96	0.95
STT-A2	74_S1	Yes	230430139	168857804	73.28	202932787	141360452	69.66	16.3	112	2.97	0.95
STT-A2	74_S1	Yes	239372866	176721169	73.83	210042923	147391226	70.17	16.6	108	3.01	0.96
STT-A2	74_S1	Yes	237989922	175832545	73.88	209759497	147602120	70.37	16.1	109	3.02	0.96
SUB-B1*	100_P1*	Yes	218976406	136064613	62.14	169074983	86163190	50.96	36.7	69	0.82	0.58
SVK-A1	70_S1	Yes	274607770	112221111	40.87	238704721	76318062	31.97	32	126	1.11	0.73
TGS-A1	8_P1	Yes	257789798	183690530	71.26	86362934	86362934	53.82	53	78	1.01	0.64
ÞSK-A26	9_P1	Yes	282421266	176969238	62.66	129361464	23909436	18.48	86.5	79	0.14	0.11
ÞSK-A26	9_P1	Yes	243301319	103301941	42.46	200101054	60101676	30.04	41.8	82	0.91	0.58
VDP-A5	97_S1	Yes	308029269	224948773	73.03	164359090	81278594	49.45	63.9	108	0.64	0.47
VDP-A5	97_P1	Yes	201938806	118188301	58.53	152583724	68833219	45.11	41.8	85	1.06	0.66
VDP-A5	97_P1	Yes	210235653	122381387	58.21	158714718	70860452	44.65	42.1	85	1.07	0.66
VDP-A5	97_P1	Yes	211926473	123422332	58.24	159638187	71134046	44.56	42.4	85	1.07	0.67
VDP-A5	97_P1	Yes	223494645	129585631	57.98	167880062	73971048	44.06	42.9	85	1.09	0.67
VDP-A5	97_P1	Yes	311609897	169950822	54.54	233797876	92138801	39.41	45.8	86	1.14	0.72
VDP-A6	94_P1	Yes	291553062	125784181	43.14	195912238	30143357	15.39	76	104	0.31	0.21
VDP-A6	94_S1	Yes	335276781	285170998	85.06	171141994	121036211	70.72	57.6	130	0.93	0.69
VDP-A6	94_S1	Yes	338329991	288018954	85.13	172178102	121867065	70.78	57.7	132	0.93	0.69
VDP-A7	101_P1	Yes	235798021	101426904	43.01	178215657	43844540	24.6	56.8	90	0.77	0.5
VDP-A7	101_P2	Yes	240990132	118807912	49.3	208859871	86677651	41.5	27	117	0.93	0.7

VDP-A7	101_P2	Yes	302961387	147897392	48.82	254560930	99496935	39.09	32.7	112	1.02	0.75
YGS-B2	92_1	No	315058614	4893032	1.55	314071208	3905626	1.24	20.2	79	0.07	0.06
YGS-B2	92_P1	Yes	243688076	127549142	52.34	155241741	39102807	25.19	69.3	63	0.37	0.26

Table S4. **DNA contamination estimates based on the mtDNA and the X chromosome for males**. The table records the contamination point estimates and p-values, the average depth of sequence coverage (DoC) for the mitochondrial genome and the X chromosome, and the predicted sex cell karyotype. Contamination estimates based on the X chromosome were conducted on all males. Contamination >5% (red) and P-values >0.05 (bold) are highlighted. Two libraries (bold red) with point estimates >5% were excluded from downstream analysis.

		CHR M						CHR X	METHOD 1			METHOD 2			
Sample	Library ID	DoC (q20) [X]	Minor allele frequency in the Icelandic cohort - less than [%]	Mean mismatch at non-MD rare derived alleles (positions)	95% CI	Contamination at possible damaged rare derived alleles (positions)	95% CI	DoC (q30) [X]	Contamination estimate [%]	SE	p-value	Contamination estimate [%]	SE	p-value	Sex chr. karyo-type
DAV-A8	25_P1 25_S1	448	5	0	NA	6.03 (1)	4.08-8.76	2.725	NA	NA	NA	NA	NA	NA	XX
DAV-A9	26_P1 26_P2	2153	5	0	NA	7.16 (1)	6.15-8.32	0.244	0.2	0.13	**0.3357**	0.47	0.29	**0.07174**	XY
DKS-A1	27_P1 27_S1	19967	5	0.28 (7)	0.25-0.31	0.4 (3)	3.75-4.1	0.363	1.11	0.22	<2.20E-16	0.71	0.34	0.00012	XY
FOV-A1	32_P1 32_P2	1485	5	1.08 (2)	0.64-1.56	2.94 (2)	2.45-3.81	0.396	1.04	0.2	3.97E-09	0.64	0.28	0.01448	XY
FSS-A1	34_P1	445	5	0.09 (3)	0.004-0.54	10.52 (6)	8.96-11.49	0.918	NA	NA	NA	NA	NA	NA	XX
GRS-A1	38_P1 38_P2	1279	5	4.28 (5)	4.24-5	4.3 (6)	3.8-4.73	0.338	2.99	0.41	<2.20E-16	3.59	0.7	9.41E-16	XY
GTE-A1	44_P1 44_S1*	19990	5	11.77 (5)	11.52-11.94	13.91 (4)	13.67-14.15	0.782	14.5	0	<2.20E-16	12.09	1.05	<2.20E-16	XY
	44_P1	6007	5	0.17 (5)	0.13-0.23	3.41 (4)	3.3-3.78	0.125	0.57	0.27	0.000929	0.94	0.43	0.002988	XY
	44_S1*	13844	5	14.17 (5)	13.89-14.38	16.18 (4)	15.9-16.46	0.242	15.41	0	<2.20E-16	12.26	0.00001	<2.20E-16	XY
HSJ-A1	2_P1 2_P2 39_P1 39_S1	18484	5	1.3 (3)	1.18-1.38	6.7 (4)	4.97-5.32	14.214	0.52	0	<2.20E-16	0.53	0	7.91E-08	XY

KNS-A1	52_S1	1550	10	0.81 (1)	0.4-1.59	3.87 (1)	2.79-5.33	0.437	0.37	0.11	0.00052	0.41	0.19	0.02273	XY
KOV-A2	3_P1	3925	5	0.34 (3)	0.24-0.47	0	NA	0.306	0.46	0.16	0.0003	0.88	0.31	0.00021	XY
MKR-A1	49_P1	1496	5	0.26 (1)	0.07-0.83	13 (3)	11.57-13.51	0.093	0.08	0.09	**0.2877**	0.08	0.14	**0.7294**	XY
NNM-A1	50_P1	420	5	0.27 (4)	0.09-0.73	5.34 (2)	3.82-7.31	0.324	0.49	0.27	0.02491	0.35	0.4	**0.2588**	XY
ORE-A1	56_S1	739	5	0.67 (3)	0.36-1.18	2.43 (3)	1.43-3.07	0.236	0.28	0.16	**0.1879**	0.26	0.27	**0.1951**	XY
SBT-A1	66_P1 66_P2	1775	5	0.75 (1)	0.37-1.48	2.53 (5)	2.0-2.66	2.422	1.08	0.04	<2.20E-16	1.01	0.07	8.46E-13	XY
SSG-A2	67_P2 67_S1	423	5	0.32 (2)	0.09-1.05	0.5 (1)	0.09-1.98	5.247	1.2	0	<2.20E-16	1.32	0	<2.20E-16	XY
SSG-A3	58_P1	3744	5	0.21 (2)	0.11-0.38	2.51 (1)	2.06-3.07	0.151	0.56	0.51	**0.3928**	0.15	1.14	**0.07615**	XY
SSG-A4	63_P1 63_P2 63_S1	441	5	0.29 (1)	0.02-1.83	2.77 (1)	1.46-5.05	6.81	NA	NA	NA	NA	NA	NA	XX
SSJ-A2	77_S1	1306	10	0.33 (2)	0.15-0.76	3.09 (10)	2.75-3.39	0.223	0.34	0.23	**0.05999**	0.06	0.11	1	XY
STT-A2	74_P1 74_S1	2618	5	0.13 (3)	0.06-0.25	1.23 (13)	1.13-1.37	5.541	0.43	0	<2.20E-16	3.80E-01	0	5.09E-05	XY
SUB-B1	100_P1*	264	5	6.32 (1)	3.84-10.11	8.62 (7)	7.45-10.08	0.475	2.45	0.38	9.14E-13	2.11	0.58	7.37E-05	XY
SVK-A1	70_S1	5982	5	1.15 (5)	0.52-0.74	4.87 (9)	4.26-4.66	0.751	0.56	0.1	8.31E-05	0.67	0.16	0.00337	XY
TGS-A1	8_P1	457	5	2.04 (2)	0.88-3.06	0	NA	0.576	0.59	0.17	0.00019	0.63	0.26	0.01147	XY
VDP-A5	97_P1 97_S1	4899	5	0	NA	4.42 (1)	3.86-5.04	1.01	NA	NA	NA	NA	NA	NA	XX
VDP-A6	101_P1 101_S1	1416	5	1.38 (1)	0.79-2.36	3.25 (5)	2.84-3.76	0.599	0.92	0.02	<2.20E-16	0.95	0.05	8.02E-14	XY
VDP-A7	94_S1	7623	5	0.06 (3)	0.91-1.08	1.58 (4)	1.47-1.75	1.006	0.6	0.02	<2.20E-16	0.73	0.05	1.05E-10	XY

| YGS-B2 | 92_P1 | 763 | 5 | 0.66 (4) | 0.21-1.79 | 4.53 (4) | 3.82-5.3 | 0.237 | NA | NA | NA | NA | NA | NA | XXY |
| ÞSK-A26 | 9_P1 | 1673 | 5 | 1.63 (1) | 0.97-2.69 | 3.43 (1) | 2.61-4.47 | 0.423 | 1.58 | 0.26 | 1.57E-15 | 1.18 | 0.4 | 5.64E-05 | XY |

Table S5. **Mitochondrial motifs based on WGS and PCR amplification.**

Sample	WGS-based motif 16055-16410	PCR-based motif 16055-16410	WGS-based motif 16517-409	PCR-based motif 16517-409	WGS-based full haplogroup motifs	Haplogroup assignment
DAV-A8			16519C 263G	16519C 263G	263G 750G 1438G 3010A 4769G 8860G 9163C 15326G 16519C	H1
DAV-A9			16519C 263G	16519C 263G	263G 750G 1438G 3010A 4769G 8860G 9163C 15326G 16519C	H1
DKS-A1	16192T 16239T 16256T 16270T 16399G	16192T 16239T 16256T 16270T 16399G	73G 150T 263G	73G 150T 263G	73G 150T 263G 750G 1303A 1438G 2706G 3192T 3197C 3591A 4592C 4769G 7028T 8860G 9477A 11296T 11467G 11719A 11938T 12308G 12372A 12618A 13617C 14766T 14793G 15218G 15326G 16192T 16239T 16256T 16270T 16399G	U5a1h
FOV-A-1	16111T 16311C	16111T 16311C	263G	263G	263G 549T 750G 1438G 2706G 3397G 4769G 7028T 8460G 8860G 15326G 16111T 16311C	HV17a
FSS-A1	16356C	16356C	16519C 73G 195C 263G	16519C 73G 195C 263G	73G 195C 263G 499A 750G 1438G 1811G 2706G 4646C 4703C 4769G 5999C 6047G 7028T 8818T 8860G 11332T 11467G 11719A 12308G 12372A 14620T 14766T 15326G 15693C 16356C 16519C	U4a2
GRS-A-1	16093C 16224C 16311C	16093C 16224C 16311C	16519C 73G 114T 263G	16519C 73G 114T 263G	73G 114T 263G 497T 593C 750G 1189C 1438G 1811G 2483C 2706G 3480G 4769G 7028T 8860G 9055A 9698C 10398G 10550G 11299C 11467G 11470G 11620G 11719A 11914A 12308G 12372A 14167T 14766T 14798C 15924G 16093C 16224C 16311C 16519C	K1a1b1b
GTE-A1			195C* 263G	263G	195C 263G 750G 1438G 3992T 4024G 4769G 5004C 8269G 8860G 9123A 12642G 14365T 14569A 14582G 15326G 15884A	H4a1a4b
HSJ-A1			16519C 46C 152C	16519C 46C 152C	46C 152C 750G 1438G 3992T 4418C 4769G 6776C 8860G 9755A 10754C 15326G 16519C	H3g1
KNS-A1	16304C 16362C	16304C 16362C	263G	263G	263G 456T 750G 1438G 4769G 8860G 15326G 16304C 16362C	H5
KOV-A2	16234T	16189C 16234T	16519C 263G	16519C 263G	263G 750G 1438G 3010A 4769G 8860G 12561A 15047A 15326G 16189C 16234T 16519C	H1
MKR-A1	16224C 16311C	16224C 16311C	16519C 73G 146C 152C 263G	16519C 73G 146C 152C 263G	73G 146C 152C 263G 750G 1189C 1438G 1811G 2706G 3480G 4769G 7028T 8860G 9055A 9093G 9698C 10398G 10550G 11299C 11377A 11467G 11719A 12308G 12372A 13105G 14167T 14766T 14798C 15326G 15900C 16224C 16311C 16519C	K1c1b

NNM-A1	16235G 16291T	16235G 16291T	263G	263G	263G 8860G 11914A 12954C 14305A 15326G 16235G 16291T	H2a2b5a
NÞR-A2	16224C 16249C 16311C	16224C 16249C 16311C	16519C 73G 263G	16519C 73G 263G	73G 263G 497T 750G 1189C 1438G 1811G 2706G 3480G 4769G 5773G 6032A 7028T 8860G 9055A 9698C 10398G 10550G 11025C 11299C 11467G 11719A 12308G 12372A 12373G 13650T 14167T 14766T 14798C 15326G 15495C 16224C 16249C 16311C 16519C	K1a2a
ORE-A1	16224C 16311C 16362C	16224C 16311C 16362C	16519C 73G 263G	16519C 73G 263G	73G 263G 497T 750G 1189C 1438G 1811G 2251G 2706G 3316A 3480G 4769G 7028T 7559G 8860G 9055A 9698C 10398G 10550G 11299C 11467G 11719A 11902A 12308G 12372A 13117G 14167T 14766T 14798C 15326G 16224C 16311C 16362C 16519C	K1a3a
SBT-A1			16519C 152C 263G	16519C 152C 263G	152C 263G 750G 1438G 3992T 4418C 4769G 6776C 8860G 10754C 12965C 15326G 15377G 16519C	H3g1a
SSG-A2	16126C	16069T 16126C	73G 185A 263G 295T	73G 185A 263G 295T	73G 185A 263G 295T 462T 489C 750G 1438G 2706G 3010A 4216C 4769G 6380G 7028T 8860G 9755A 10398G 11251G 11719A 12612G 13708A 13934T 14766T 14798C 15326G 15452A 16069T 16126C	J1c3g
SSG-A3	16126C	16126C 16294T	16519C 73G 152C 263G	16519C 73G 152C 263G	73G 152C 263G 709A 750G 930A 1438G 1888A 2706G 4216C 4769G 4917G 5147A 7028T 8697A 8860G 10463C 11242G 11251G 11719A 11812G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16519C	T2b2b
SSG-A4	16069T 16126C 16145A 16172C 16192T 16261T	16069T 16126C 16145A 16172C 16192T 16261T	73G 242T 263G 295T	73G 242T 263G 295T	73G 242T 263G 295T 462T 489C 750G 1438G 2158C 2706G 3010A 4216C 4769G 5460A 5463T 6911C 7028T 8269A 8557A 8860G 10398G 11251G 11719A 12007A 12612G 13708A 13879C 14766T 15326G 15452A 16069T 16126C 16145A 16172C 16192T 16261T	J1b1a1a
SSJ-A2	16172C 16256T 16399G	16172C 16256T 16399G	73G 263G	73G 263G	73G 263G 750G 1438G 1700C 2706G 3197C 4769G 5495C 7028T 8705C 8860G 9477A 11467G 11719A 12308G 12372A 13617C 14766T 14793G 15218G 15326G 15924G 16172C 16256T 16399G	U5a1a1
STT-A2	16356C	16356C	16519C 73G 152C 195C 263G	16519C 73G 152C 195C 263G	73G 152C 195C 263G 499A 750G 1438G 1811G 2706G 4646C 4769G 5999C 6047G 7028T 7705C 8308G 8860G 9389G 10819G 11332T 11339C 11467G 11719A 12308G 12372A 13528G 13565T 14620T 14766T 15326G 15373G 15693C 15758G 16356C 16519C	U4b1b1

SVK-A1	16129A 16223T 16391A	16129A 16223T 16391A	16519C 73G 152C 199C 204C 207A 250C 263G 315_1C	16519C 73G 152C 199C 204C 207A 250C 263G 315_1C	73G 152C 199C 204C 207A 250C 263G 750G 1438G 1719A 2706C 4529T 4769G 7028T 8251A 8860G 9653C 10034C 10238C 10398C 11719A 12501A 12705T 13780G 14766T 15043A 15326G 15758G 15924G 16129A 16223T 16391A 16519C	I2
TGS-A1	16126C 16153A 16294T	16126C 16153A 16294T	16519C 41T 73G 150T 263G	16519C 41T 73G 150T 263G	41T 73G 150T 263G 709A 750G 1438G 1888A 2706C 4216C 4769G 4917G 7028T 8697A 8860G 10463C 11251G 11719A 11812G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16153A 16294T 16519C	T2e1
VDP-A5			16519C 263G	16519C 263G	263G 750G 1438G 4769G 6776C 8860G 15326G 16519C	H3
VDP-A6	16176T 16219G	16176T 16219G	16519C 146C 257G 263G	16519C 146C 257G 263G	146C 257G 263G 477C 750G 1438G 3010A 4769G 8473C 8860G 15326G 15758G 16176T 16219G 16519C	H1c3a
VDP-A7			263G	263G	263G 750G 1438G 3992T 4024G 4769G 5004C 8269A 8860G 9123A 10044G 14365T 14582G 15326G	H4a1a1
YGS-B2	16126C	16069T 16126C	73G 228A 263G 295T	73G 228A 263G 295T	73G 228A 263G 295T 462T 482C 489C 750G 1438G 2706G 3010A 3394C 4216C 4769G 7028T 8860G 9635C 10398G 11251G 11623T 11719A 12612G 13708A 13899C 14766T 14798C 15326G 15452A 16069T	J1c1a
ÞSK-A26	16126C 16145A 16172C 16222T 16261T	16069T 16126C 16145A 16172C 16192T 16222T 16261T	73G 242T 263G 295T	73G 242T 263G 295T	73G 242T 263G 295T 462T 489C 750G 1438G 2158C 2706G 3010A 4216C 4769G 5460A 5463T 6911C 7028T 8269A 8557A 8860G 10398G 11251G 11719A 12007A 12612G 13708A 13879C 14766T 15326G 15452A 16069T 16126C 16145A 16172C 16192T 16222T	J1b1a1a

The position marked with an asterisk symbol (*) is the only one that differs between Sanger sequence and WGS data, which could be explained by the fact that it is heteroplasmic.

Table S6. **Inference of biological sex**. Sequences with mapping quality of at least 30 were used. Statistics were computed using the *Ry* approach.

Sample	Number of alignments to chrX and chrY	Number of alignments to chrY	R_y	SE	95% CI	Inferred sex chr. karyotype	Inferred biological sex
DAV-A8	8962496	18392	0.0021	1.00E-04	0.002-0.0021	XX	Female
DAV-A9	619099	68225	0.1102	0.0004	0.1094-0.111	XY	Male
DKS-A1	1168134	143320	0.1227	0.0003	0.1221-0.1233	XY	Male
FOV-A1	968134	107001	0.1105	0.0003	0.1099-0.1111	XY	Male
FSS-A1	2335627	5080	0.0022	1.00E-04	0.0021-0.0022	XX	Female
GRS-A1	797055	77530	0.0973	0.0003	0.0966-0.0979	XY	Male
GTE-A1	361200	40198	0.1113	0.0005	0.1103-0.1123	XY	Male
HSJ-A1	27974756	3045520	0.1089	0.0001	0.1088-0.109	XY	Male
KNS-A1	1063798	121868	0.1146	0.0003	0.114-0.1152	XY	Male
KOV-A2	862231	92555	0.1073	0.0003	0.1067-0.108	XY	Male
MKR-A1	425384	46790	0.11	0.0005	0.1091-0.1109	XY	Male
NNM-A1	699499	79886	0.1142	0.0004	0.1135-0.1149	XY	Male
NÞR-A2	1025445	2624	0.0026	1.00E-04	0.0025-0.0027	XX	Female
ORE-A1	480412	62833	0.1308	0.0005	0.1298-0.1317	XY	Male
SBT-A1	7731038	803053	0.1039	0.0001	0.1037-0.1041	XY	Male
SSG-A2	10354341	1007151	0.0973	0.0001	0.0971-0.0974	XY	Male
SSG-A3	295126	30886	0.1047	0.0006	0.1035-0.1058	XY	Male
SSG-A4	14111563	25079	0.0018	0	0.0018-0.0018	XX	Female
SSJ-A2	406220	52075	0.1282	0.0005	0.1272-0.1292	XY	Male
STT-A2	12970412	1394719	0.1075	0.0001	0.1074-0.1077	XY	Male
SVK-A-1	1733259	205367	0.1185	0.0002	0.118-0.119	XY	Male
TGS-A1	1348489	135371	0.1004	0.0003	0.0999-0.1009	XY	Male
VDP-A5	6727199	19049	0.0028	1.00E-04	0.0028-0.0029	XX	Female
VDP-A6	4364988	485617	0.1113	0.0002	0.111-0.1115	XY	Male
VDP-A7	4281340	569778	0.1331	0.0002	0.1328-0.1334	XY	Male
YGS-B2	558974	29906	0.0535	0.0003	0.0529-0.0541	XXY	Male Klinefelter syndrome
ÞSK-A26	1072265	114549	0.1068	0.0003	0.1062-0.1074	XY	Male

Table S7. **Summary of reference data used in this study.**

Study and Chip type	European panel. no. of samp.	North European panel. no. of samp.
This study	*973*	*4224*
Illumina Human Omni Express chip		
Sweden	194	1,103
Norway	193	1,035
Denmark	199	980
Iceland	197	916
The "Irish DNA Atlas"	*190*	*190*
Illumina Human Omni Express chip		
Ireland	190	190
Leslie et al. (2015)	*562*	*1246*
Illumina Human 1M chip and Affymetrix 500K SNP		
Wales	130	130
Gwynedd	*76*	*76*
Dyfed	*54*	*54*
England	200	882
Midlands	*45*	*219*
North East	*24*	*96*
North West	*23*	*95*
South East	*45*	*203*
South West	*31*	*174*
Yorkshire	*32*	*95*
Northern Ireland	40	40
Scotland	192	194
Dumfries and Galloway	*42*	*42*
Fife	*57*	*57*
Orkney Islands	*93*	*95*
The International HapMap 3 Consortium	NA	*91*
Illumina Human 1M chip		
Yoruba	NA	91
Li et al. 2008	*125*	NA
Illumina HumanHap 650 K		
France	47	NA
Bergamo Italy	13	NA
Sardinia Italy	21	NA
Tuscany Italy	8	NA
Russia	25	NA
Adygei Russia	11	NA
Yunusbayev et al. 2012	*33*	NA
Illumina 610 K SNP array		
Bulgaria	13	NA
Ukraine	20	NA

Kovacavic et al. 2014	***85***	NA
Illumina 660 K SNP array		
Macedonia	14	NA
Kosovo	9	NA
Montenegro	14	NA
Mordovia	15	NA
Bosnia-Herzegovina (Bosnian)	15	NA
Serbia/Bosnia-Herzegovina (Serbian)	18	NA
Yunusbayev et al. 2015	***83***	NA
Illumina 650 K/610K/550K SNP array		
Germany	13	NA
Karelia	15	NA
Gagauz	12	NA
Russia	32	NA
Veps	11	NA
The 1000 Genomes Project Consortium	***278***	NA
Illumina Omni2.5 array		
Tuscan Italy	98	NA
Finland	83	NA
Spain	97	NA

Table S8. **Ancient individuals from the British Isles dating from the Neolithic to the Anglo-Saxon period.**

Time period	¹⁴C cal	Lab ID	Sample ID	Country or region	Site	EuroSNPs [%]†	NEurSNP [%]‡	Study
Middle to Late Neolithic	3343-3020 BC	BA64*	Ballynahatty	Ireland	Ballynahatty	95.69	NA	Cassidy et al. (2015)
Early Bronze Age	2026-1885 BC	RM127	Rathlin1	Ireland	Glebe, Rathlin Island	95.67	99.92	Cassidy et al. (2015)
	2024-1741 BC	RSK1	Rathlin2	Ireland	Glebe, Rathlin Island	73.87	77.07	Cassidy et al. (2015)
	1736-1534 BC	RSK2	Rathlin3	Ireland	Glebe, Rathlin Island	53.31	55.58	Cassidy et al. (2015)
Late Iron Age	210 BC-40 AD	1489	M1489	Northern Britain	Melton	37.25	38.92	Martiniano et al. (2016)
	160 BC-50 AD	SK1964	HI1	East England	Hinxton	18.99	19.67	Schiffels et al. (2016)
	170 BC-80 AD	SK1231	H12	East England	Hinxton	80.18	83.3	Schiffels et al. (2016)
Roman	71-400 AD	3DT16, 3DT26*, 6DT3, 6DT18, 6DT21, 6DT22, 6DT23	3DRIF-16, 3DRIF-26, 6DRIF-3, 6DRIF-18, 6DRIF-21, 6DRIF-22, 6DRIF-23	Northern Britain	Driffield Terrace	48.15, 64.63, 78.82, 64.22, 67.04, 65.71, 47.39	50,15, NA, 82,29, 67.14, 69.9, 68,66, 49.32	Martiniano et al. (2016)
Anglo-Saxon	550-650 AD	NO3423	NO3423	Northern Britain	Norton on Tees	63.39	66.3	Martiniano et al. (2016)
	631-776 AD	SK5518	HS2	East England	Hinxton	42.28	43.61	Schiffels et al. (2016)
	666-770 AD	SK241	HS1	East England	Hinxton	37.98	39.05	Schiffels et al. (2016)
	690-881 AD	SK355	HS3	East England	Hinxton	11.52	11.88	Schiffels et al. (2016)

Samples marked with an asterisk symbol (*) are not of North European ancestry (Fig. S8A) and were therefore excluded from further analysis. †Overlapping SNPs with European reference data. ‡Overlapping SNPs with North European reference data.

Table S9. **D-statistics of the form** $D(Yoruba, X; Norse, Celtic)$**.** These results are plotted in main Fig. 2B. The table is sorted by D-statistic value, from lowest to highest.

Outgroup	Sample (X)	Pop1	Pop2	D-stat	SD	Z	BABA	ABBA	No. of SNPs
Yoruba	DAV-A8	Norse	Celtic	-0.0023	0.00031	-7.432	27702	27830	379483
Yoruba	SVK-A1	Norse	Celtic	-0.0023	0.00033	-6.917	18768	18854	256472
Yoruba	DAV-A9	Norse	Celtic	-0.0016	0.000392	-4.14	6617	6639	91550
Yoruba	NNM-A1	Norse	Celtic	-0.0016	0.000427	-3.705	6886	6908	94136
Yoruba	DKS-A1	Norse	Celtic	-0.0013	0.000386	-3.381	8760	8783	119576
Yoruba	SSJ-A2	Norse	Celtic	-0.0011	0.000489	-2.302	4335	4344	59687
Yoruba	TGS-A1	Norse	Celtic	-0.0011	0.000348	-3.16	16785	16822	230869
Yoruba	VDP-A5	Norse	Celtic	-0.001	0.000335	-3.044	25834	25887	351726
Yoruba	FOV-A1	Norse	Celtic	-0.0009	0.000376	-2.339	11451	11472	157205
Yoruba	VDP-A6	Norse	Celtic	-0.0009	0.00035	-2.451	19521	19554	266003
Yoruba	MKR-A1	Norse	Celtic	-0.0007	0.000538	-1.374	3470	3475	47201
Yoruba	KOV-A2	Norse	Celtic	-0.0003	0.000371	-0.814	10378	10384	143139
Yoruba	HSJ-A1	Norse	Celtic	-0.0002	0.000332	-0.602	29920	29932	404010
Yoruba	FSS-A1	Norse	Celtic	0.0001	0.000337	0.44	17253	17247	236555
Yoruba	VDP-A7	Norse	Celtic	0.0003	0.000318	0.849	25736	25722	350236
Yoruba	GRS-A1	Norse	Celtic	0.0004	0.000389	0.918	8531	8525	117195
Yoruba	GTE-A1	Norse	Celtic	0.0005	0.000499	0.915	3904	3900	53242
Yoruba	ÞSK-A26	Norse	Celtic	0.0005	0.000354	1.393	13600	13586	187705
Yoruba	SSG-A3	Norse	Celtic	0.0008	0.000572	1.33	2511	2507	34574
Yoruba	YGS-B2	Norse	Celtic	0.001	0.000463	2.248	4558	4549	63332
Yoruba	SBT-A1	Norse	Celtic	0.0011	0.000299	3.531	29254	29192	397830
Yoruba	STT-A2	Norse	Celtic	0.0014	0.000314	4.468	29866	29783	403679
Yoruba	NÞR-A2	Norse	Celtic	0.0017	0.00039	4.324	9563	9531	133068
Yoruba	ORE-A1	Norse	Celtic	0.0018	0.000414	4.422	7046	7020	96146
Yoruba	SSG-A2	Norse	Celtic	0.0018	0.00031	5.787	29770	29663	403271
Yoruba	KNS-A1	Norse	Celtic	0.0025	0.000374	6.591	11131	11076	152735
Yoruba	SSG-A4	Norse	Celtic	0.0026	0.000315	8.251	29470	29317	400347

Table S10. **Haplogroup assignments of North European reference individuals.**

chrY haplogroup	Iceland CNT (%)	Norway CNT (%)	Sweden CNT (%)	Denmark CNT (%)	Ireland CNT (%)	Scotland CNT (%)	Wales CNT (%)	England CNT (%)	Norse CNT (%)	Celtic CNT (%)
R1b1	154 (33.41)	74 (30.2)	48 (19.92)	236 (39.14)	112 (81.75)	48 (52.17)	48 (81.36)	261 (61.27)	122 (25.10)	197 (72.43)
R1a1	113 (24.51)	61 (24.9)	46 (19.09)	62 (10.28)	2 (1.46)	7 (7.61)	2 (3.39)	14 (3.29)	107 (22.02)	6 (2.21)
I1	139 (30.15)	80 (32.65)	99 (41.08)	192 (31.84)	0	5 (5.43)	2 (3.39)	37 (8.69)	179 (36.83)	8 (2.94)
I1a1	0	0	0	0	0	2 (2.17)	0	18 (4.23)	0	2 (0.74)
I2	0	2 (0.82)	2 (0.83)	6 (1.0)	0	1 (1.09)	0	7 (1.64)	4 (0.82)	1 (0.37)
I2a	0	0	0	0	8 (5.84)	0	0	0	0	8 (2.94)
I2a1	6 (1.3)	0	1 (0.41)	0	0	0	0	2 (0.47)	0	0
I2a2	12 (2.6)	14 (5.71)	6 (2.49)	45 (7.46)	8 (5.84)	3 (3.26)	2 (3.39)	29 (6.81)	20 (4.12)	13 (4.78)
Q	0	0	0	1 (0.17)	1 (0.73)	0	0	0	0	0
Q1a2	32 (6.94)	3 (1.22)	8 (3.32)	8 (1.33)	0	0	1 (1.69)	1 (0.23)	11 (2.26)	1 (0.37)
E	0	0	0	0	0	1 (1.09)	0	3 (0.70)	0	1 (0.37)
E1b1	2 (0.43)	0	4 (1.66)	12 (1.99)	1 (0.73)	23 (25.0)	2 (3.39)	32 (7.51)	4 (0.82)	26 (9.56)
E2b1	0	0	0	0	0	0	0	2 (0.47)	0	0
J	0	0	0	4 (0.66)	0	0	0	2 (0.47)	0	0
J2	0	0	0	0	0	1 (1.09)	1 (1.69)	0	0	2 (0.74)
J2a1	1 (0.22)	1 (0.41)	1 (0.41)	6 (1.0)	0	0	0	4 (0.94)	2 (0.41)	0
J2b2	0	1 (0.41)	1 (0.41)	1 (0.17)	0	0	0	4 (0.94)	2 (0.41)	0
G	0	0	0	0	1 (0.73)	0	0	0	0	1 (0.37)
G2	0	0	1 (0.41)	1 (0.17)	1 (0.73)	0	0	0	1 (0.21)	1 (0.37)
G2a2	0	1 (0.41)	3 (1.24)	10 (1.66)	3 (2.19)	1 (1.09)	1 (1.69)	7 (1.64)	4 (0.82)	5 (1.84)
T	0	1 (0.41)	1 (0.41)	0	0	0	0	1 (0.23)	2 (0.41)	0
T1a2	0	0	0	2 (0.33)	0	0	0	0	0	0
DE	0	1 (0.41)	0	0	0	0	0	0	1 (0.21)	0
H1	0	0	0	0	0	0	0	1 (0.23)	0	0
L1	0	0	0	1 (0.17)	0	0	0	0	0	0
N1c1	2 (0.43)	6 (2.45)	21 (8.71)	16 (2.65)	0	0	0	1 (0.23)	27 (5.56)	0

Table S11. **D-statistics of the form** *D*(*Yoruba, X; Norse, Icelandic*) **and** *D*(*Yoruba, X; Celtic, Icelandic*)**.** These results are plotted in main Fig. 3B. The table is sorted by D-statistic value, from lowest to highest.

Outgroup	Sample (X)	Pop1	Pop2	D-stat	SD	Z	BABA	ABBA	No. of SNPs
Yoruba	DKS-A1	Norse	Iceland	0.0002	0.000352	0.616	8766	8762	119576
Yoruba	SK355	Norse	Iceland	0.0002	0.000513	0.45	3241	3240	48003
Yoruba	DAV-A8	Norse	Iceland	0.0003	0.000313	1.088	27755	27736	379483
Yoruba	SSJ-A2	Norse	Iceland	0.0003	0.00046	0.656	4338	4336	59687
Yoruba	NO3423	Norse	Iceland	0.0005	0.000307	1.557	19523	19504	267891
Yoruba	ÞSK-A26	Norse	Iceland	0.0005	0.000342	1.382	13583	13571	187705
Yoruba	6DT21	Norse	Iceland	0.0006	0.000311	1.901	20597	20572	282442
Yoruba	6DT3	Norse	Iceland	0.0007	0.000313	2.271	24316	24282	332500
Yoruba	MKR-A1	Norse	Iceland	0.0007	0.000485	1.446	3471	3467	47201
Yoruba	HSJ-A1	Norse	Iceland	0.0008	0.000325	2.538	29922	29872	404010
Yoruba	NÞR-A2	Norse	Iceland	0.0008	0.000367	2.251	9543	9528	133068
Yoruba	SBT-A1	Norse	Iceland	0.0008	0.000306	2.683	29213	29165	397830
Yoruba	SK241	Norse	Iceland	0.0008	0.000359	2.239	10832	10814	157768
Yoruba	SK5518	Norse	Iceland	0.0008	0.000341	2.199	12098	12079	176220
Yoruba	SSG-A3	Norse	Iceland	0.0008	0.000558	1.479	2509	2505	34574
Yoruba	GRS-A1	Norse	Iceland	0.0009	0.000364	2.465	8528	8513	117195
Yoruba	STT-A2	Norse	Iceland	0.0009	0.000303	3.066	29816	29760	403679
Yoruba	3DT16	Norse	Iceland	0.001	0.000343	2.974	14727	14697	202621
Yoruba	FSS-A1	Norse	Iceland	0.001	0.000323	3.17	17250	17215	236555
Yoruba	VDP-A7	Norse	Iceland	0.001	0.000315	3.111	19538	19499	266003
Yoruba	1489	Norse	Iceland	0.0011	0.000371	2.967	11424	11399	157374
Yoruba	SVK-A1	Norse	Iceland	0.0011	0.000322	3.315	18820	18779	256472
Yoruba	YGS-B2	Norse	Iceland	0.0011	0.00044	2.516	4554	4544	63332
Yoruba	NNM-A1	Norse	Iceland	0.0012	0.000406	2.857	6900	6884	94136
Yoruba	SK1231	Norse	Iceland	0.0012	0.000303	3.842	23823	23768	336593
Yoruba	TGS-A1	Norse	Iceland	0.0012	0.000318	3.707	16808	16769	230869
Yoruba	6DT23	Norse	Iceland	0.0013	0.000331	3.976	14511	14473	199299
Yoruba	DAV-A9	Norse	Iceland	0.0013	0.000403	3.29	6631	6613	91550
Yoruba	RSK2	Norse	Iceland	0.0013	0.000334	3.847	16305	16264	224582
Yoruba	GTE-A1	Norse	Iceland	0.0014	0.000481	2.81	3902	3892	53242
Yoruba	KNS-A1	Norse	Iceland	0.0014	0.000366	3.862	11104	11073	152735

65

Yoruba	RM127	Norse	Iceland	0.0014	0.000302	4.679	29755	29672	403724
Yoruba	6DT18	Norse	Iceland	0.0015	0.000312	4.829	19821	19762	271270
Yoruba	6DT22	Norse	Iceland	0.0015	0.000304	4.873	20204	20144	277417
Yoruba	RSK1	Norse	Iceland	0.0015	0.000308	4.844	22677	22610	311415
Yoruba	SSG-A2	Norse	Iceland	0.0015	0.000298	5.002	29725	29637	403271
Yoruba	VDP-A5	Norse	Iceland	0.0015	0.000313	4.749	25877	25800	351726
Yoruba	VDP-A6	Norse	Iceland	0.0015	0.000317	4.863	25743	25664	350236
Yoruba	ORE-A1	Norse	Iceland	0.0017	0.000376	4.441	7035	7011	96146
Yoruba	SK1964	Norse	Iceland	0.0017	0.000429	3.897	5423	5405	79495
Yoruba	FOV-A1	Norse	Iceland	0.002	0.00036	5.626	11476	11429	157205
Yoruba	SSG-A4	Norse	Iceland	0.002	0.000291	6.89	29415	29297	400347
Yoruba	KOV-A2	Norse	Iceland	0.0028	0.000352	8.065	10403	10344	143139
Yoruba	6DT21	Celtic	Iceland	-0.0011	0.000397	-2.73	20567	20612	282442
Yoruba	KNS-A1	Celtic	Iceland	-0.001	0.000465	-2.177	11082	11104	152735
Yoruba	1489	Celtic	Iceland	-0.0009	0.00042	-2.126	11404	11424	157374
Yoruba	NÞR-A2	Celtic	Iceland	-0.0009	0.000462	-1.907	9530	9547	133068
Yoruba	6DT23	Celtic	Iceland	-0.0008	0.000397	-2.082	14487	14511	199299
Yoruba	6DT22	Celtic	Iceland	-0.0007	0.000376	-1.848	20166	20194	277417
Yoruba	SK1964	Celtic	Iceland	-0.0006	0.000534	-1.098	5413	5420	79495
Yoruba	SSG-A4	Celtic	Iceland	-0.0006	0.000374	-1.538	29352	29386	400347
Yoruba	RSK2	Celtic	Iceland	-0.0005	0.000442	-1.088	16280	16296	224582
Yoruba	SK1231	Celtic	Iceland	-0.0005	0.000383	-1.301	23790	23814	336593
Yoruba	STT-A2	Celtic	Iceland	-0.0005	0.000355	-1.318	29783	29811	403679
Yoruba	6DT18	Celtic	Iceland	-0.0004	0.000383	-1.098	19790	19807	271270
Yoruba	RM127	Celtic	Iceland	-0.0004	0.000371	-1.065	29711	29734	403724
Yoruba	RSK1	Celtic	Iceland	-0.0003	0.000401	-0.655	22646	22657	311415
Yoruba	SSG-A2	Celtic	Iceland	-0.0003	0.000372	-0.903	29682	29702	403271
Yoruba	SBT-A1	Celtic	Iceland	-0.0002	0.000365	-0.494	29189	29199	397830
Yoruba	ORE-A1	Celtic	Iceland	-0.0001	0.000493	-0.258	7026	7028	96146
Yoruba	ÞSK-A26	Celtic	Iceland	-0.0001	0.000431	-0.213	13577	13580	187705
Yoruba	3DT16	Celtic	Iceland	0.0001	0.000412	0.192	14717	14715	202621
Yoruba	6DT3	Celtic	Iceland	0.0001	0.00037	0.272	24305	24300	332500
Yoruba	SSG-A3	Celtic	Iceland	0.0001	0.000664	0.148	2506	2506	34574
Yoruba	YGS-B2	Celtic	Iceland	0.0001	0.000553	0.139	4549	4548	63332
Yoruba	SK355	Celtic	Iceland	0.0002	0.000624	0.263	3242	3241	48003

Yoruba	SK241	Celtic	Iceland	0.0003	0.000447	0.769	10826	10819	157768
Yoruba	GRS-A1	Celtic	Iceland	0.0006	0.000468	1.203	8525	8515	117195
Yoruba	SK5518	Celtic	Iceland	0.0008	0.000421	1.833	12099	12080	176220
Yoruba	FSS-A1	Celtic	Iceland	0.0009	0.000414	2.22	17251	17219	236555
Yoruba	GTE-A1	Celtic	Iceland	0.0009	0.000593	1.442	3902	3895	53242
Yoruba	HSJ-A1	Celtic	Iceland	0.0011	0.000389	2.71	29929	29866	404010
Yoruba	NO3423	Celtic	Iceland	0.0011	0.00039	2.853	19533	19489	267891
Yoruba	VDP-A7	Celtic	Iceland	0.0013	0.000401	3.142	25739	25674	350236
Yoruba	MKR-A1	Celtic	Iceland	0.0014	0.0006	2.356	3473	3464	47201
Yoruba	DKS-A1	Celtic	Iceland	0.0015	0.000453	3.394	8777	8750	119576
Yoruba	SSJ-A2	Celtic	Iceland	0.0015	0.0006	2.438	4342	4330	59687
Yoruba	VDP-A6	Celtic	Iceland	0.0019	0.000426	4.386	19553	19480	266003
Yoruba	TGS-A1	Celtic	Iceland	0.0023	0.000421	5.397	16826	16750	230869
Yoruba	VDP-A5	Celtic	Iceland	0.0025	0.000385	6.548	25903	25772	351726
Yoruba	DAV-A8	Celtic	Iceland	0.0027	0.000374	7.116	27811	27664	379483
Yoruba	NNM-A1	Celtic	Iceland	0.0027	0.000487	5.537	6909	6872	94136
Yoruba	FOV-A1	Celtic	Iceland	0.0029	0.000449	6.524	11484	11417	157205
Yoruba	DAV-A9	Celtic	Iceland	0.003	0.000509	5.858	6641	6601	91550
Yoruba	KOV-A2	Celtic	Iceland	0.0031	0.000442	7.074	10406	10341	143139
Yoruba	SVK-A1	Celtic	Iceland	0.0033	0.000402	8.324	18856	18730	256472

References:

33.	R. A. Bentley, Strontium isotopes from the earth to the archaeological skeleton: A review. *J Archaeol Method Th* **13**, 135-187 (2006).

34.	A. Helgason *et al.*, Sequences from first settlers reveal rapid evolution in Icelandic mtDNA pool. *PLoS genetics* **5**, e1000343 (2009).

35.	H. Gestsdóttir, *Geldingurinn á Öndverðarnesi*. (2000), vol. 1998, pp. 143-150.

36.	H. Stefánsson, Enn um silfursalann og urðarbúann. *Eimreiðin* **4**, (1949).

37.	B. Gíslason, Beinafundurinn hjá Jökulsá. *Lesbók Morgunblaðsins* **21**, 189-195 (1946).

38.	Á. Óla, Lesbók Morgunblaðsins. (1948).

39.	Þ. M., Skeljastaðir, Þjórsárdalur. *Forntida, Gårdar i Island*, (1943).

40.	N. Rohland, M. Hofreiter, Ancient DNA extraction from bones and teeth. *Nature protocols* **2**, 1756-1762 (2007).

41.	J. Dabney *et al.*, Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 15758-15763 (2013).

42.	M. Meyer, M. Kircher, Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor protocols* **2010**, pdb prot5448 (2010).

43.	S. Lindgreen, AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res Notes* **5**, 337 (2012).

44.	H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).

45.	H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).

46.	A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).

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

48.	T. S. Korneliussen, A. Albrechtsen, R. Nielsen, ANGSD: Analysis of Next Generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).

49.	L. Orlando *et al.*, Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499**, 74-78 (2013).

50.	J. Herrero *et al.*, Ensembl comparative genomics resources. *Database (Oxford)* **2016**, (2016).

51.	R. M. Andrews *et al.*, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* **23**, 147 (1999).

52.	M. Rasmussen *et al.*, An aboriginal Australian genome reveals separate human dispersals into Asia. *Science* **334**, 94-98 (2011).

53.	D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**, 1655-1664 (2009).

54.	A. L. Price *et al.*, Long-range LD can confound genome scans in admixed populations. *Am J Hum Genet* **83**, 132-135 (2008).

55.	L. Kovacevic *et al.*, Standing at the Gateway to Europe - The Genetic Structure of Western Balkan Populations Based on Autosomal and Haploid Markers. *PloS one* **9**, (2014).

56.	J. Z. Li *et al.*, Worldwide human relationships inferred from genome-wide patterns of variation. *Science* **319**, 1100-1104 (2008).

57.	B. Yunusbayev *et al.*, The Caucasus as an asymmetric semipermeable barrier to ancient human migrations (vol 29, pg 359, 2012). *Molecular Biology and Evolution* **29**, 1891-1891 (2012).

58.	B. Yunusbayev *et al.*, The genetic legacy of the expansion of Turkic-speaking nomads across Eurasia. *PLoS genetics* **11**, 1005068 (2015).

59.	D. M. Altshuler *et al.*, A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).

60. S. Leslie *et al.*, The fine-scale genetic structure of the British population. *Nature* **519**, 309-314 (2015).

61. The Irish DNA Atlas; revealing fine-scale population structure and history within Ireland. *Scientific Reports*.

62. A. K. Hedstrom *et al.*, Smoking and two human leukocyte antigen genes interact to increase the risk for multiple sclerosis. *Brain* **134**, 653-664 (2011).

63. I. S. o. G. G. (2016). (23 December 2016).

64. N. Patterson, A. L. Price, D. Reich, Population structure and eigenanalysis. *PLoS genetics* **2**, e190 (2006).

65. C. Wang *et al.*, Comparing spatial maps of human population-genetic variation using Procrustes analysis. *Stat Appl Genet Mol Biol* **9**, Article 13 (2010).

66. E. Y. Durand, N. Patterson, D. Reich, M. Slatkin, Testing for ancient admixture between closely related populations. *Mol Biol Evol* **28**, 2239-2252 (2011).

67. R. E. Green *et al.*, A draft sequence of the Neandertal genome. *Science* **328**, 710-722 (2010).

68. D. H. Alexander, K. Lange, Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* **12**, 246 (2011).

69. H. Gestsdóttir, "Kyn- og lífaldursgreiningar á beinum úr íslenskum kumlum". *Fornleifastofnun Íslands ; FS055-98151* (Fornleifastofnun Íslands, Reykjavik, 1998).