

RESEARCH ARTICLE

# Imputed genomes and haplotype-based analyses of the Picts of early medieval Scotland reveal fine-scale relatedness between Iron Age, early medieval and the modern people of the UK

Adeline Morez<sup>1\*</sup>, Kate Britton<sup>2,3</sup>, Gordon Noble<sup>2</sup>, Torsten Günther<sup>4</sup>, Anders Götherström<sup>5</sup>, Ricardo Rodríguez-Varela<sup>5</sup>, Natalija Kashuba<sup>6</sup>, Rui Martiniano<sup>1</sup>, Sahra Talamo<sup>3,7</sup>, Nicholas J. Evans<sup>2</sup>, Joel D. Irish<sup>1</sup>, Christina Donald<sup>8</sup>, Linus Girdland-Flink<sup>1,2\*</sup>

**1** School of Biological and Environmental Sciences, Liverpool John Moores University, Liverpool, United Kingdom, **2** Department of Archaeology, School of Geosciences, University of Aberdeen, Aberdeen, United Kingdom, **3** Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, **4** Department of Organismal Biology, Uppsala University, Uppsala, Sweden, **5** Department of Archaeology and Classical Studies, Stockholm University, Stockholm, Sweden, **6** Department of Archaeology and Ancient History, Uppsala University, Uppsala, Sweden, **7** Department of Chemistry, University of Bologna, Bologna, Italy, **8** The McManus: Dundee's Art Gallery and Museum, Dundee, United Kingdom

\* [adelinemorez@gmail.com](mailto:adelinemorez@gmail.com) (AM); [linus.girdlandflink@abdn.ac.uk](mailto:linus.girdlandflink@abdn.ac.uk) (LGF)



## OPEN ACCESS

**Citation:** Morez A, Britton K, Noble G, Günther T, Götherström A, Rodríguez-Varela R, et al. (2023) Imputed genomes and haplotype-based analyses of the Picts of early medieval Scotland reveal fine-scale relatedness between Iron Age, early medieval and the modern people of the UK. *PLoS Genet* 19(4): e1010360. <https://doi.org/10.1371/journal.pgen.1010360>

**Editor:** Jonathan Marchini, Regeneron Genetics Center, UNITED STATES

**Received:** July 29, 2022

**Accepted:** February 7, 2023

**Published:** April 27, 2023

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pgen.1010360>

**Copyright:** © 2023 Morez et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All raw and mapped sequence data generated for this project are

## Abstract

There are longstanding questions about the origins and ancestry of the Picts of early medieval Scotland (*ca.* 300–900 CE), prompted in part by exotic medieval origin myths, their enigmatic symbols and inscriptions, and the meagre textual evidence. The Picts, first mentioned in the late 3<sup>rd</sup> century CE resisted the Romans and went on to form a powerful kingdom that ruled over a large territory in northern Britain. In the 9<sup>th</sup> and 10<sup>th</sup> centuries Gaelic language, culture and identity became dominant, transforming the Pictish realm into Alba, the precursor to the medieval kingdom of Scotland. To date, no comprehensive analysis of Pictish genomes has been published, and questions about their biological relationships to other cultural groups living in Britain remain unanswered. Here we present two high-quality Pictish genomes (2.4 and 16.5X coverage) from central and northern Scotland dated from the 5<sup>th</sup>–7<sup>th</sup> century which we impute and co-analyse with >8,300 previously published ancient and modern genomes. Using allele frequency and haplotype-based approaches, we can firmly place the genomes within the Iron Age gene pool in Britain and demonstrate regional biological affinity. We also demonstrate the presence of population structure within Pictish groups, with Orcadian Picts being genetically distinct from their mainland contemporaries. When investigating Identity-By-Descent (IBD) with present-day genomes, we observe broad affinities between the mainland Pictish genomes and the present-day people living in western Scotland, Wales, Northern Ireland and Northumbria, but less with the rest of England, the Orkney islands and eastern Scotland—where the political centres of Pictland were located. The pre-Viking Age Orcadian Picts evidence a high degree of IBD sharing across modern

available from the European Nucleotide Archive under the study accession number PRJEB58104. All imputed data generated for this project are available from the European Variation Archive under the study accession number PRJEB60907.

**Funding:** AM was supported by ECR strategic support of early career researchers in the faculty of science at LJMU, awarded to LGF. LGF was supported by the School of Geoscience, University of Aberdeen. KB was supported by the Leverhulme Trust (PLP-2019-284) during production of this manuscript. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

Scotland, Wales, Northern Ireland, and the Orkney islands, demonstrating substantial genetic continuity in Orkney for the last ~2,000 years. Analysis of mitochondrial DNA diversity at the Pictish cemetery of Lundin Links ( $n = 7$ ) reveals absence of direct common female ancestors, with implications for broader social organisation. Overall, our study provides novel insights into the genetic affinities and population structure of the Picts and direct relationships between ancient and present-day groups of the UK.

## Author summary

We report two high-quality autosomal and eight mitochondrial genomes sequenced from individuals associated with the Pictish period of early medieval Scotland (*ca.* 300–900 CE). We demonstrate genetic affinities between the Pictish genomes and Iron Age people who lived in Britain, which supports current archaeological theories of a local origin. The autosomal genomes also allowed us to detect haplotype sharing between the Pictish genomes and present-day Europeans. Our results demonstrate a proportionally higher degree of haplotype sharing, and thus genetic affinity, between the Pictish genomes and individuals from western Scotland, Wales, Northern Ireland and Northumbria. We also detected genetic structure in Scotland during the Iron Age, likely driven by the combination of genetic drift and small population size, which we also detect in present-day Orcadians. Lastly, the seven mitochondrial DNA from the Lundin Links cemetery showed that these individuals had no direct maternal ancestors which could suggest exchanges of people, or at least females, between groups during the Pictish period, challenging older ideas that the Picts were a matrilineal society. Overall, our results show that high-quality ancient genomes combined with haplotype imputation are highly informative for obtaining novel insights to population structure and migration over the past 2,000 years.

## Introduction

The genetic origins of the present-day populations of the UK have been extensively studied and can broadly be modelled as a mixture of three deep genetic ancestries, mirroring western European ancestry: western European Mesolithic hunter-gatherer ancestry, Early European farmer ancestry derived from Anatolian Neolithic farmers, and Late Neolithic steppe-related ancestry [1–10]. Our understanding of more recent demographic changes in the British and Irish Isles—referring to Britain, Ireland and associated smaller islands—has also been expanded via large-scale sequencing of ancient genomes, revealing extensive gene flow from mainland Europe into southern Britain during the Middle Bronze Age, which contributed to genetic differentiation between Iron Age groups from southern and northern Britain [9, 10]. Present-day genetic diversity in Wales, Cornwall, Devon and western Ireland indicates a long-standing genetic structure, possibly already present during the Iron Age [11], but the lack of ancient samples especially from Scotland limits our ability to directly test this hypothesis, and ‘pockets’ of older ancestries could have survived regionally in isolated populations for extended periods [12,13].

The British and Irish Isles witnessed a complex cultural turnover from the Iron Age to the early medieval period. The Romans occupied part of Britain to southern Scotland from 43 to *ca.* 410 CE; however, the single aDNA study of Roman Britain suggests this occupation resulted in little detectable gene flow from mainland Europe [14]. Multiple episodes of long-

distance migration across western and central Eurasia [9,15,16] intensified during the Late Antique period (*ca.* 300–800 CE), before and following the collapse of the Western Roman Empire. In Britain, Angles, Saxons and other Germanic-speaking peoples, likely originating in Scandinavia, the Low Countries and parts of Germany, settled predominantly in south-eastern and central Britain with genetic evidence of extensive admixture with local populations carrying genetic ancestry from the Iron Age [14,17,18]. During the so-called ‘Viking Age’ (starting about 800 CE), Scandinavians settled in the ‘Danelaw’ in northern and eastern England, as well as in the coastal areas of Ireland and northern and western Britain [19], which led to admixture with the inhabitants of Ireland and western and northern Britain over nearly four centuries [9]. In addition, local, culturally distinct groups lived in Britain around the end of the Roman period, before major Anglo-Saxon settlement: the Britons (speaking the ancestral language of Welsh, as well as Latin) inhabited the island south of the Firth of Forth, the Gaels (Dál Riata) occupied Argyll and the southern Hebrides in Scotland, and the Picts lived in the rest of Britain north of the Forth [20,21]. The genetic diversity between and within these groups is poorly understood. In particular, the lack of genomes from Scotland has limited our ability to understand how the genetic structure changed between the Iron Age and the early medieval period.

Among the peoples present during the first millennium CE in Britain, the Picts (*ca.* 300–900 CE) are one of the most enigmatic. Their unique cultural features (e.g. Pictish symbols) and the scarcity of direct writing resulted in many diverse hypotheses about their origin, life-style and culture, the so-called ‘Pictish problem’ [22]. Other than a list of kings and difficult-to-decipher ogham and alphabetic inscriptions, the only written evidence comes to us from their neighbours—the Romans and later the Gaels, Britons and Anglo-Saxons. This deficiency has been compounded by a sparse archaeological record with few settlements and fewer cemeteries from this period [23].

In the modern era perceptions of Pictish origins have varied, often according to cultural and political biases, with the Picts and their languages regarded as Germanic, Gaelic, Brittonic, Basque, and Illyrian, among other theories. In the 1950s Jackson influentially argued that the Picts spoke a non-Indo-European language and a Celtic language akin to ancient Gaulish [24,25]. The current consensus is that they spoke a Celtic language closest to that of neighbouring Britons from which Cornish, Welsh, and Breton derive [26–28]. However, some still argue from undeciphered inscriptions and other words that some Picts spoke an otherwise unknown language, presumably derived from a pre-Celtic population [29, 30]. This theory is underpinned by recent findings arguing that Late Bronze Age migrations in southern Britain may have introduced Celtic or Celtic-related languages in this area; however, these migrations were seemingly not as influential in northern Britain [10]. Thus, the question remains of whether the Picts were somehow fundamentally different from their neighbours.

In the medieval period, the Picts were considered immigrants from Thrace (north of the Aegean Sea), Scythia (eastern Europe), or isles north of Britain [31]. However, Irish accounts and the Northumbrian scholar Bede added that, before settling in Britain, the Picts first gained wives in Ireland, on the condition that Pictish succession passed through the female line. This is the origin of the theory that the Picts practised a form of matriliney, with succession and perhaps inheritance going to the sister’s son rather than directly through the male line. However, our earliest source for this practice, Bede’s ‘Ecclesiastical History of the English People’ (finished in 731 CE), stated that it was limited to occasions when the succession was in dispute. It is now argued that the origin-legend was intended to reinforce Pictish identity and legitimise particular kings whose claims to the throne were through their mothers [32, 33]. Nevertheless, matriliney remains one potential explanation for the absence of father-to-son succession in at least one Pictish royal dynasty before the mid-8<sup>th</sup> century [34].

Here, we aim to provide new insights into the genetic diversity of the Picts through analysis of two ancient whole genomes (2.4 and 16.5X coverage) sequenced from individuals excavated from two Pictish-era cemeteries: Lundin Links (Fife, Southern Pictland) and Balintore (Easter Ross, Northern Pictland). We imputed diploid genotypes alongside published medium-to-high coverage ancient genomes, including individuals from England dating from the Iron Age, Roman and early medieval periods [14,17], and co-analysed these with previously imputed ancient genomes from the Orkney islands dating from the Late Iron Age (Pictish period) and Viking Age [9]. Using allele frequency and haplotype-based methods, we aim to determine the genetic relationships between the Picts and neighbouring modern-day and ancient populations. In addition, using the mitogenomes of seven individuals from Lundin Links, we explore how differences in female mobility due to possible post-marital residence customs, i.e., matrilocality (female endogamy), may have shaped the genetic diversity at this possible high-status cemetery and discuss its implication for our understanding of Pictish elite descent systems.

## Results and discussion

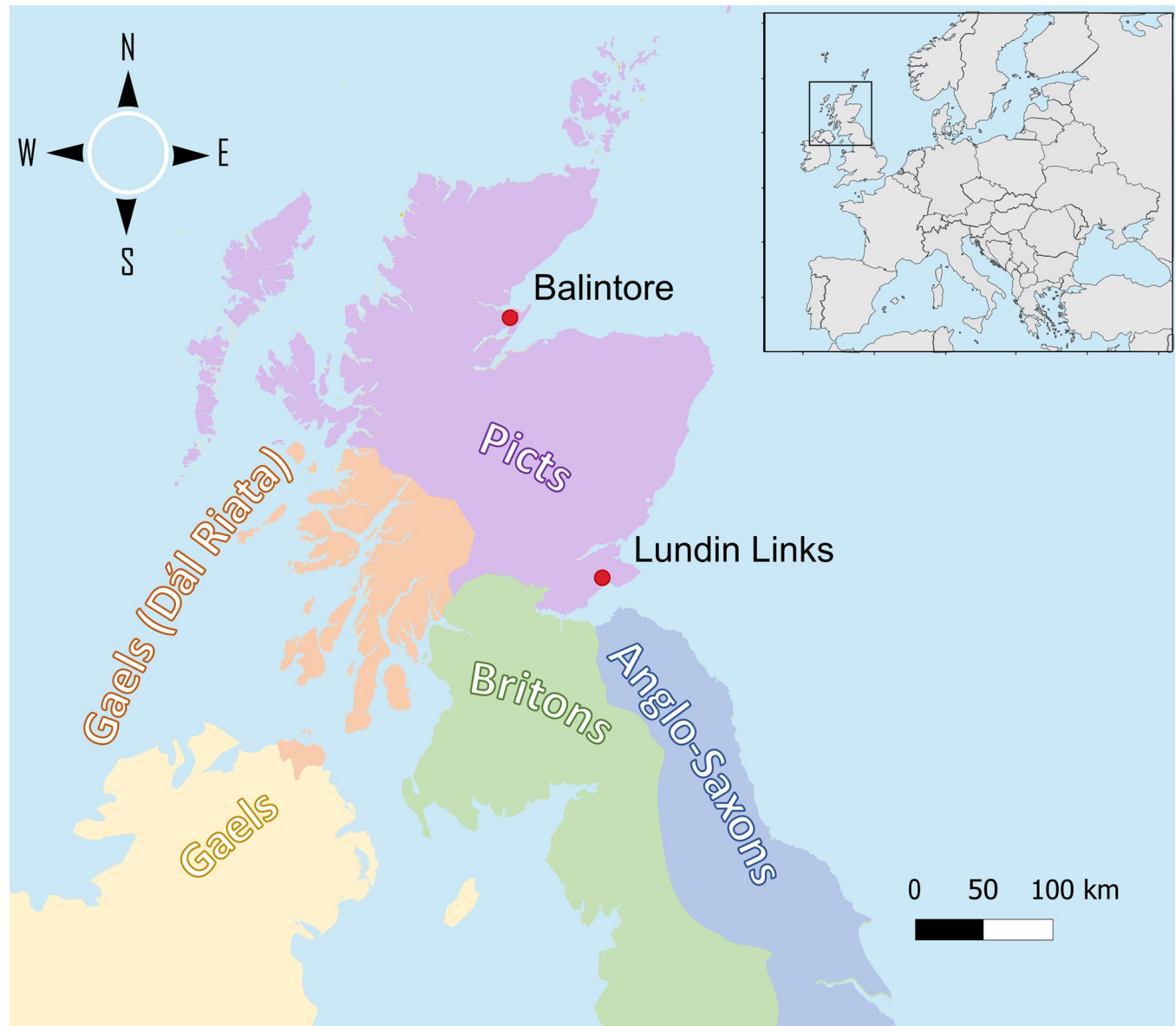
### DNA extraction, sequencing, and radiocarbon dating

We retrieved DNA from eight individuals, one from Balintore (Easter Ross) and seven from Lundin Links (Fife), representing the northern and southern parts of Pictland (Fig 1 and Tables 1 and S1 and section S1.1 in S1 Text). Two individuals, BAL003 and LUN004 were treated with the USER enzyme to remove post-mortem deamination [35] and were shotgun sequenced to medium and high coverage (2.4 and 16.5X) (Tables 1 and S1). Seven individuals from Lundin Links were shotgun sequenced to a sufficient mitochondrial DNA (mtDNA) coverage for subsequent analyses (3.47–195.05X, Table 1). LUN001 and LUN003 are excluded from the population genetics analysis involving autosomal DNA as we found evidence for library index misassignment in the autosomal data, but not the mitochondrial data (section S1.2 in S1 Text). Ten samples from Lundin Links (including LUN001, LUN002, LUN003 and LUN009) and three from Balintore were radiocarbon dated to the 5<sup>th</sup>–7<sup>th</sup> century CE (Tables 1 and S4 and S2 Fig and section S1.1 in S1 Text) [36]. The individuals will be referred to as Pictish throughout the text as they lived during the period that Pictish identity existed and are from areas likely to have been within Pictish territories *ca.* 700 CE.

All samples evidenced deamination patterns at the fragment termini characteristic of ancient DNA (aDNA), except BAL003 and LUN004 which underwent damage repair enzymatic treatment, for which the respective unrepaired screening libraries BAL003-b1e1l1 and LUN004-b1e1l1 did show deamination (S3 Fig). Contamination estimates based on mtDNA, X and Y-chromosome reads are low (Tables 1 and S1 and section S1.3 in S1 Text). The genetic sex determination (S4 Fig) agrees with the morphological sex determination (Tables 1 and S1).

### Demographic history

**Analysis of uniparental genetic markers.** The mitochondrial haplogroups observed in the samples are common in present-day north-western Europeans, with the sub-clade J1c3 being identified in three individuals out of eight (Tables 1 and S5). In terms of paternal Y-chromosomal lineages, we assigned LUN004 to R1b-DF49 (Tables 1 and S6), which is predominantly distributed in the UK and Ireland and a sub-clade to R1b-P312/S116 haplogroup introduced to Britain by Bell Beaker peoples during the Chalcolithic, alongside steppe-related ancestry [7]. During the Chalcolithic, R1b-derived haplogroups largely replaced the predominant I2a Y-chromosome lineage in the British Neolithic [7,8,40] except in Orkney where I2a persisted into the Bronze Age [13]. R1b sub-clades are extremely common across Britain and western Europe from the Iron Age onwards [13,14,17,41].



**Fig 1. Sampling location and the regions under ancient Brittonic, Irish and Anglo-Saxon control around the 7<sup>th</sup> century [37–39].**

<https://doi.org/10.1371/journal.pgen.1010360.g001>

**Allele frequency-based genomic affinities in ancient Britain.** To investigate population affinities of the individuals from Pictland, we performed Principal Component Analysis (PCA) and ADMIXTURE analyses on a dataset comprising present-day Europeans, the newly imputed genomes and the imputed ancient genomes from Margaryan et al. [9] (S7 Table). The PCA shows that the ancient individuals from Britain broadly fit with present-day individuals whose four grandparents lived in Britain (Fig 2A). However, we notice some variability among these individuals as BAL003 and LUN004 fall within the modern Welsh cluster, but with BAL003 being notably closer to the present-day Scottish, Orcadians, English and Northern Irish clusters, suggesting some degree of genetic differentiation amongst individuals from Pictland. The Iron Age and Roman period individuals from England are spread across the modern English, Northern Irish, Scottish and Welsh clusters. Four ancient Orcadians from the Iron



Table 1. Summary information from the eight samples investigated in this study.

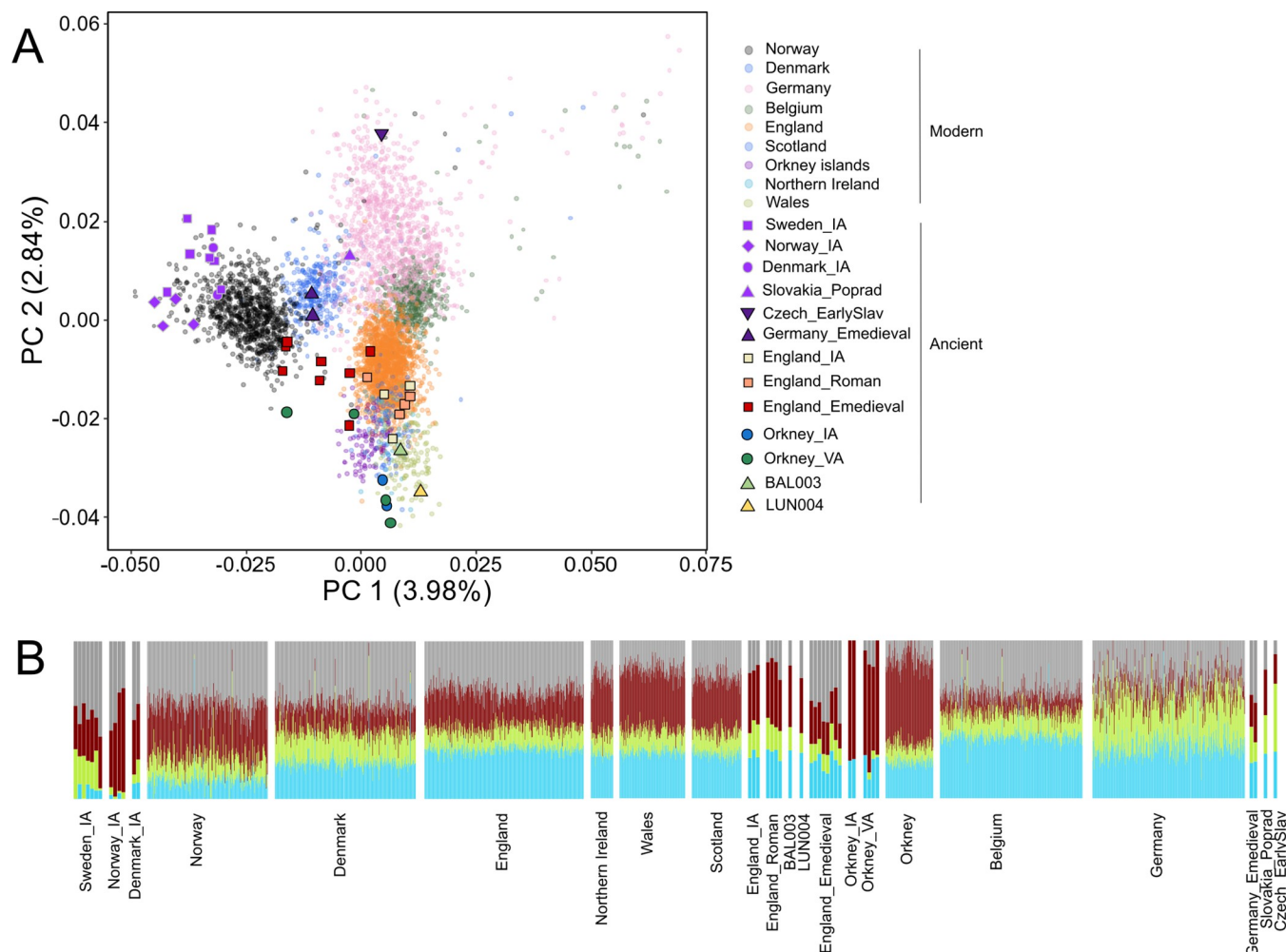
Sample	Site	Radiocarbon calibrated date (95.4% confidence, AD)	Genome coverage (X)	mtDNA Coverage (X)	Sex	mtDNA / Y haplogroup	Contamination Estimate (%)		
							Based on mtDNA	Based on X	Based on Y
BAL003	Balintore	419–538	16.54	294.97	XX	H2a1e	1 ± 1	-	0.30
LUN001	Lundin Links	416–545	0.18	195.05	XX	T2a1a	2 ± 1	-	3.78
LUN002	Lundin Links	563–653	2.24x10 <sup>-3</sup>	3.47	-	H1c20	1 ± 1	-	-
LUN003	Lundin Links	384–562	0.32	154.31	XX	T2b11	2 ± 1	-	1.70
LUN004	Lundin Links	406–542	2.43	121.32	XY	J1c3g / R1b- DF49	1 ± 1	0.80 ± 0.10	-
LUN005	Lundin Links		7.02x10 <sup>-3</sup>	3.99	XY	K1c2	1 ± 1	-	-
LUN006	Lundin Links		1.48x10 <sup>-3</sup>	8.13	-	J1c3b1	2 ± 1	-	-
LUN009	Lundin Links	430–590	6.11x10 <sup>-3</sup>	4.26	XX	J1c3b	1 ± 1	-	0.59

<https://doi.org/10.1371/journal.pgen.1010360.t001>

and Viking Ages fit with present-day Welsh, Northern Irish and Scottish populations. However, two Viking Age Orcadians (VK204 and VK205) are intermediates between the British and Scandinavian clusters, consistent with previous results finding evidence of admixture in these individuals between British-like and Scandinavian-like ancestries [9]. The early medieval individuals from England are intermediate between modern English people and Scandinavians, which is consistent with various degrees of admixture between Iron Age groups from England and immigrants from northern/central Europe [14,17,18]. These results agree with the pseudo-haploid-based analyses of the BAL003 and LUN004 genomes, showing a broad affinity to modern western Europeans (section S1.3 in [S1 Text](#) and [S10](#), [S12–S15](#) and [S18 Figs](#)), but with a much-improved resolution.

### Haplotype-inferred genetic structure in early medieval Britain

Haplotype-based methods have been shown to outperform conventional unlinked SNP approaches in the detection of population substructure [42–44]. To make use of the additional power provided by haplotype information, we conducted a FineSTRUCTURE clustering analysis and Identity-By-Descent (IBD) analysis on the imputed diploid dataset. Our analysis shows that the genomes from Lundin Links and Balintore form a genetic cluster together with genomes from the Iron Age and Roman period from England (except 6DT3 –who instead shows strong affinity to western/central Europe or Scandinavia based on the IBD analysis, [Fig 3](#) and section S1.6 in [S1 Text](#)–and I0160), and from the Late Iron Age to Viking Age from Orkney (except VK204 and VK205 who carried substantial Scandinavian-like ancestry; ‘Pop12’, [Figs 3](#) and [S27](#)). Included in this cluster are also Viking Age individuals from Britain, Iceland and Scandinavia; the latter likely corresponds to individuals buried in Scandinavia but whose parents were from a British-like gene pool, consistent with results in Margaryan et al. [9]. Based on *outgroup-f3*, the individuals from Orkney, Scotland, and England, dated from the Iron Age to medieval period are symmetrically related to each other ([S5 Fig](#)). However, we also note that BAL003, but not LUN004, show multiple instances of IBD sharing >4 cM with early medieval individuals from England ([S21 Fig](#)), which is also reflected in their relative position in the PCA ([Fig 2A](#)), implying substantial shared ancestry and possibly recent gene-flow

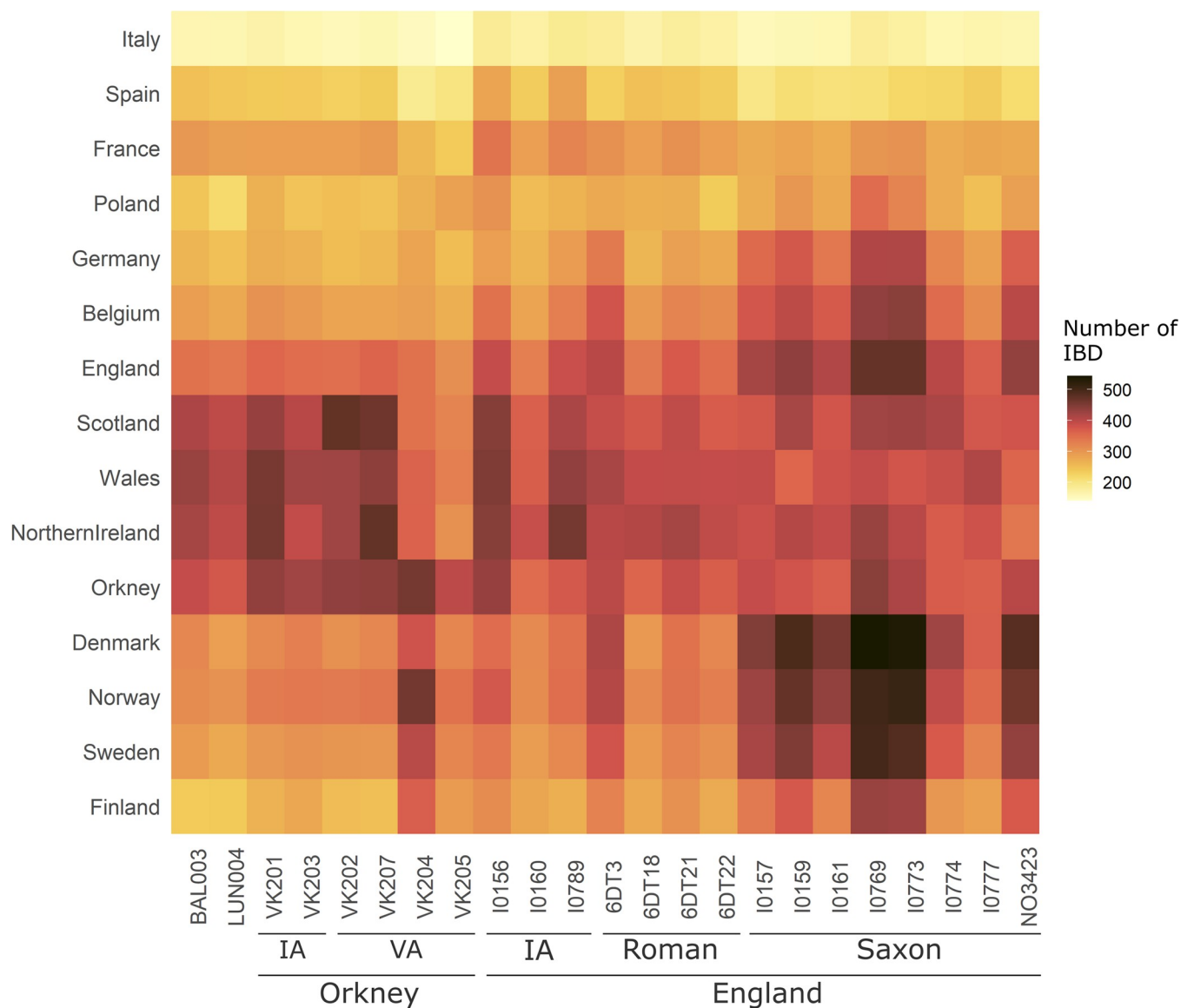


**Fig 2. Genetic diversity of Iron Age, early medieval and present-day individuals from northern and central Europe.** A) A Principal Component Analysis of 4,914 individuals and 87,518 SNPs. B) An ADMIXTURE ancestry component ( $K = 4$ ) of these same genomes (see S7 Fig for the complete analysis from  $K = 2$  to  $K = 10$ ).

<https://doi.org/10.1371/journal.pgen.1010360.g002>

from a source genetically similar to those samples. We therefore suggest that individuals from Pictland should not be considered a homogenous genetic group, but instead a complex mixture of contemporary genetic ancestries.

The unlinked approach implemented in the ADMIXTURE analysis also reveals a minor but detectable genetic structure consistent with results from the PCA (Fig 2A) but not evident in the FineSTRUCTURE analysis (Figs 2B and S7). While the proportion of ancestry components are similar across BAL003, LUN004, Iron Age and the Roman period in England, the Late Iron Age and unadmixed Viking Age Orcadians are differentiated from this group (Figs 2B and S7). They show an absence of the grey and green ancestry components, likely first introduced by Scandinavian migrants as they are first observed in VK204 and VK205 and then in modern Orcadians. These components are also carried at a high proportion in modern Norwegians and Danes. Iron and Viking Age individuals from Orkney with an absence of Scandinavian-like admixture share more genetic drift with modern Orcadians [ $f_3(\text{Orkney Iron or Viking Age (unadmixed), modern Orcadians; Yoruba}) < 0.13$ ] than do BAL003 and LUN004 [ $f_3(\text{Picts, modern Orcadians; Yoruba}) = 0.03$ ]; we calculated this on ancient pseudo-haploid



**Fig 3. Shared Identity-By-Descent (IBD) segments >1 cM between the ancient genomes from Britain and present-day European populations.** IA, Iron Age. VA, Viking Age.

<https://doi.org/10.1371/journal.pgen.1010360.g003>

genomes to avoid allele frequency bias originating from the imputation. The results confirm the relatively high proportion of shared genetic drift in Orkney for the last 2,000 years, which supports our interpretation of the ADMIXTURE plot. Nevertheless, a high relative count of IBD sharing (>1 cM, >4 cM and >6 cM) between LUN004 and Late Iron Age or Viking Age Orcadians (S21 Fig) demonstrates that gene flow between Orkney and mainland Scotland—or a source outside Scotland that also contributed ancestry to LUN004—likely occurred.

The Iron Age and unadmixed Viking Age Orcadians also show the highest degree of a red ancestry component (Figs 2B and S7), which is inconsistent with having originated from direct gene flow from any population included in this study and instead likely reflects retention of a less diverse pre-Iron Age ancestry in Orkney and/or strong genetic drift (such as a bottleneck or founder effect) (see [45] for insight on how a bottleneck or admixture can result in a similar

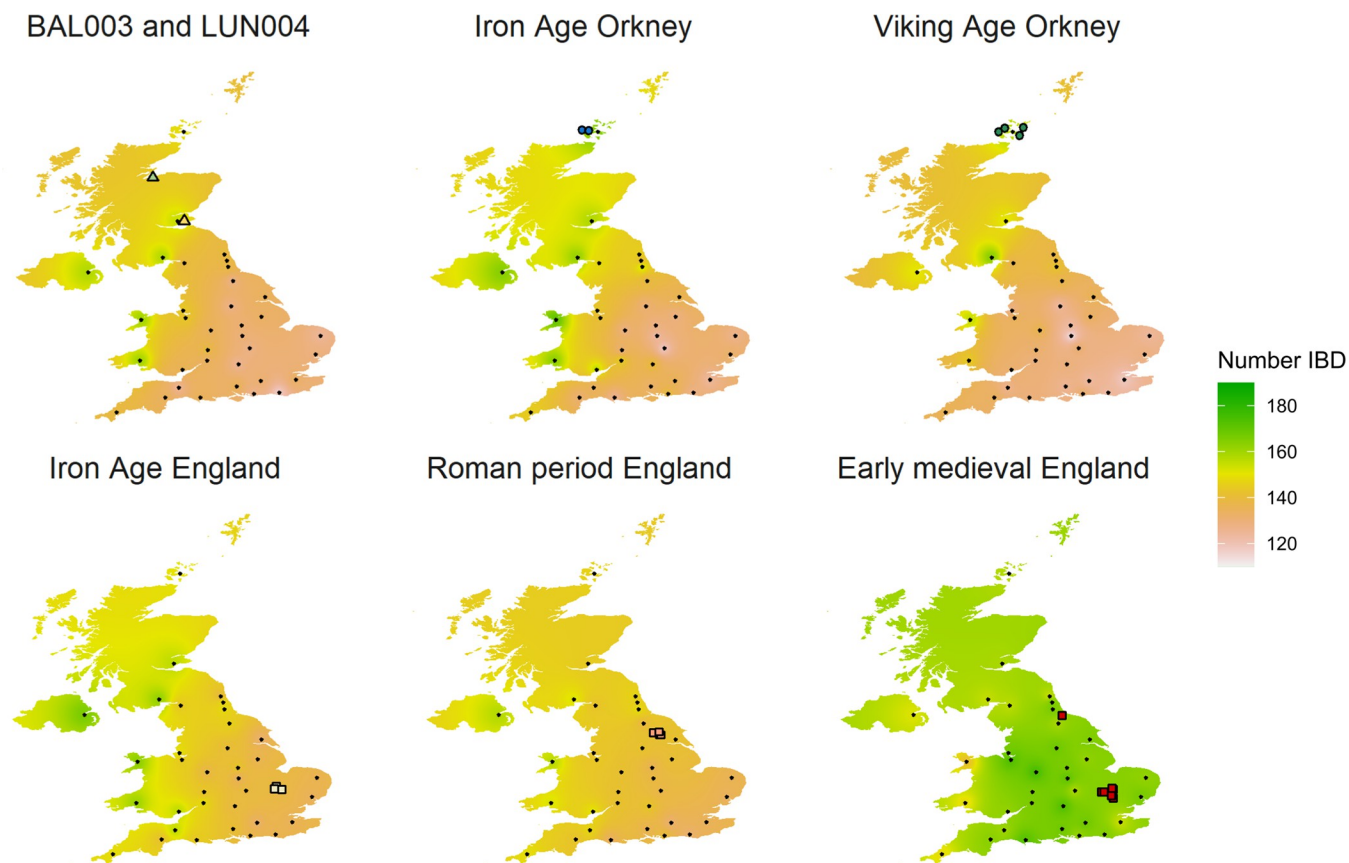


ADMIXTURE profile). In fact, recent research show that Bronze Age populations in Orkney were differentiated from their counterparts on mainland Britain because of strong genetic drift as a result of a small ancestral population size, which was deduced from the observation of a relatively high number of short runs of homozygosity [13]. The Bronze Age Orkney population also retained the Y-chromosomal haplogroup I2, which is associated with Neolithic ancestry, while the R1b haplogroup associated with Bell Beaker expansion largely replaced the I2 haplogroups in the rest of Britain [13]. Moreover, although modern Orcadians are differentiated from the rest of the British Isles due to extensive admixture with Scandinavians, recent genomic research shows that genetic drift also played an important role [11,46]. This is consistent with our results that show a high proportion of shared IBD segments among modern Orcadians ( $>1$  to  $>6$  cM, S19 Fig), meaning they share a high proportion of recent common genetic ancestors relative to most modern European populations, typical of small or genetically isolated populations. Three Orcadians dated from the Late Iron Age and Viking Age also displayed the highest number of small HBD  $<1.5$  cM (S26 Fig), typical of individuals descending from a small population. One ancient individual from Orkney (VK201) evidenced a long Homozygosity-By-Descent (HBD) segment ( $>9.5$  cM), the longest observed amongst all ancient individuals and indicative of small population size or consanguinity (S26 Fig). Overall, these data indicate a long-term small population size, which likely contributed to the extensive genetic drift observed in modern Orcadians. The genetic differentiation between populations living in Orkney and Scotland during the Late Iron Age and early medieval period could thus be partially explained by different degrees of genetic drift.

### Analysis of genetic continuity across Britain

The Pictish data allow us to obtain a transect of Iron Age/early medieval genomes across Britain and directly look at the pattern of haplotype sharing between them and present-day genomes. The Iron Age and Roman period (except 6DT3) individuals from England and Scotland share more IBD segments  $>1$  cM (both in terms of number and length) with present-day individuals from Scotland (including Orkney), Northern Ireland and Wales than with any other European populations included in our analyses (Figs 3 and S20), consistent with the structure observed in the PCA analysis (Fig 2A). We also show that all early medieval individuals (excluding I0777) share more IBD with modern Danish than with any other present-day population (Fig 3), suggesting genetic continuity between modern-day Danish and the ancestors of these individuals (section S1.6 in S1 Text).

The analysis also revealed high IBD sharing between early medieval individuals from England and present-day people across Britain following a southeast/northwest cline (Figs 4 and S22). This pattern suggests that northern continental European ancestry associated with Anglo-Saxon migrations expanded out of south-eastern England followed by admixture with local populations, a scenario consistent with previous research [11,14,17,18,46,47]. BAL003 and LUN004 share a high proportion of IBD segments with present-day people from western Scotland, Wales and Northern Ireland, similar to the individuals from Late Iron Age Orkney and England (Figs 4 and S22). However, unlike these individuals, LUN004, and to a lesser extent BAL003, shares fewer IBD segments with the present-day eastern Scottish population sample (Figs 4 and S22). Byrne et al. [47] and Gilbert et al. [46] previously suggested that the genetic structure between western and eastern Scotland could result from the divide between the kingdoms of the Gaelic-speaking Dál Riata in the west and Picts in the east, which is seemingly in contradiction with the results presented here. Instead, the present-day genetic structure in Scotland likely results from more complex demographic processes that cannot be reduced to a single model.



**Fig 4. Average IBD sharing >1 cM between present-day and ancient groups from the UK.** IBD sharing between each of the ancient genomes and modern samples is illustrated in S22 Fig. Ancient individuals are indicated with coloured symbols. The black dots represent the geographic location of present-day people from 35 regions of the UK [11, 53], by the county town.

<https://doi.org/10.1371/journal.pgen.1010360.g004>

We propose two non-exclusive processes that might explain the observed pattern of IBD sharing between the Iron Age and early medieval populations and the present-day Scottish population. The first is substantial admixture from immigrants that brought Iron Age Orkadian-, and England-like ancestries (likely independently), which partially replaced the eastern Scottish early medieval gene pool. Indeed, in the following centuries (1,100–1,300 CE), eastern Scotland received substantial immigration, such as settlers from Britain south of the Forth, France, and the Low Countries [48–50]. Under this scenario, BAL003 and LUN004 are good representatives of the broader ancestry present in Scotland during the Pictish period. Alternatively, the ancestors of BAL003 and LUN004 share more IBD segments with present-day people from western Scotland, Wales, and Northern Ireland because they (or their direct ancestors) migrated from these regions but did not contribute substantially to later generations via admixture with local groups in eastern Scotland. This scenario is consistent with an emerging picture of west-east lifetime mobility of both males and females in the early medieval period in Scotland [51,52]. Under such a model, it may be feasible that there are indeed still undiscovered ‘pockets’ of eastern Pictish-period ancestry, that was differentiated from ancestry carried by BAL003 and LUN004 and which contributed significantly to present-day populations from eastern Scotland. Oxygen and strontium isotope analysis of teeth from these individuals holds promise to characterise this further. Importantly, we also emphasise that

stochasticity likely affected the pattern of IBD sharing in such a small sample size. Indeed, high variability in IBD sharing is observed amongst individuals from the early medieval and Iron Age groups, and to some extent between BAL003 and LUN004 (S22Fig).

Our results also show substantial IBD sharing between Iron Age, Viking Age and present-day Orcadians, supporting our observations using allele-frequency based methods of strong genetic continuity in this region over time (Figs 2, 4, and S22). Therefore, the marked genetic differentiation between Orkney and mainland Britain is not only a result of Scandinavian admixture, as previously hypothesised [11,46,54–57] but also pronounced genetic continuity that persisted for at least 2,000 years. The relatively low IBD sharing between BAL003 and LUN004 and modern-day Orcadians (Fig 4) suggests the emergence of Pictish culture in Orkney [22,23,37] was not associated with population replacement but largely due to cultural diffusion and connections.

IBD segments in Iron Age individuals from south-eastern England are widespread throughout western and northern Britain compared to the more recent Romano-British individuals from northern England; the latter, however, do not share substantial IBD with any present-day people of the British Isles (Figs 4 and S22). The only exception is 6DT3 who was from the same genetic population as two early medieval individuals (I0159 and I0773) with Scandinavian-, and northern European-like ancestry ('pop12', S22 Fig and section S1.6 in S1 Text). 6DT3 also share relatively more IBD segments  $>1$  cM with the present-day population from Scandinavia, Belgium and the UK (Fig 3), suggesting that Scandinavian-like ancestry could have spread to the British Isles before the early medieval period.

These results need to be interpreted with caution because of existing flaws in the IBD segment detection. First, it has previously been demonstrated that the false positive rate (genomic regions falsely attributed to IBD segments) is high, especially for small IBD segments [43]. They are usually created from the concatenation of shorter IBD segments into a false longer one. The false positive rate is around 10–40% for  $\leq 2$  cM IBD segments, but decreases with the length of the segment and is close to 0% for IBD segments  $\geq 5$  cM. Several factors can influence this rate, the main ones being data quality and population history. The detection of false positives varies with the density and informativeness of genetic markers. For example, segments of the genome containing many rare alleles are easier to identify as IBD. Regions with low SNP density will be associated with a higher false positive rate. Ralph and Coop [43] also detected significant variation in the false positive rate between populations for IBD segments  $<2$  cM, with populations from Spain, Portugal and Italy showing significantly higher false positive rates than other European populations. A likely explanation given by the authors is that IBD detection algorithms are based on haplotype frequency, so the false positive rate should increase in populations that are more differentiated from the rest of the sample; in this case, they are the populations sharing the least common genetic ancestors both within their respective populations and with other European populations. A second flaw is the low power to detect true small IBD  $<2$  cM, as demonstrated in [43] using fastIBD [58]. They estimate that the power of detecting  $<2$  cM IBD segments is  $<50\%$ . However, in the present study, we used RefinedIBD [59], which was shown to be more powerful in detecting true 1–2 cM IBD segments, reaching 40% power for 1 cM IBD segments (with segment LOD  $>3$ ), against only 20% using fastIBD. One possible reason that can break down haplotypes is genotyping error [43, 60]; however, we merged segments having at most one discordant homozygote  $<0.6$  cM apart (see section S1.6 in S1 Text) to overcome this issue. Other effects could influence IBD detection, such as a higher chance of difference in recombination rates among ancient samples and the HapMap dataset or presence of structural variants [61].

## Social organisation

Seven mtDNA genomes were retrieved at Lundin Links, which allows us to address questions about the Pictish social organisation reflected in the individuals interred at the site. The use of the cemetery was relatively short, likely around 130 years (S2 Fig), and the individuals excavated were adults (S1 Table) [36]. The diversity of mtDNA lineages was high, and none of the individuals shared an immediate maternal ancestor (S5 Table). It is worth noting that individuals retrieved from the horned cairns complex (S1 Fig) show evidence of familial links based on the presence of highly heritable skeletal traits (congenitally missing teeth, torus mandibularis, and septal apertures on the humeri) [35]. However, the two individuals yielding mtDNA (LUN001 and LUN009) are not maternally related (S5 Table). In a matrilineal system, which is typical of matrilineal descent, low female post-marital migration and high male migration decrease female mtDNA diversity [62–65]. This result suggests the individuals buried at Lundin Links were unlikely to have been practicing matrilocality; this assumes that this cemetery is a snapshot of the community near Lundin Links, and individuals did not return to their natal place for burial after having lived in their marital community. Ongoing isotope analyses focused on the movement histories of the Lundin Links individuals using strontium, oxygen and other isotope approaches may further characterise sex-specific mobility. Additional Y-chromosome analyses will also help confirm whether patrilocality or neolocality was more common in Pictish society [62].

Seventy per-cent of matrilineal societies are associated with a matrilineal system [66]. Insofar as funerary treatment can tell us about the social organisation in life, it is unlikely that the community at Lundin Links followed a matrilineal inheritance system. This interpretation challenges older arguments for matrilineal succession among Pictish rulers [34]. However, while some individuals buried at Lundin Links may have been of elevated social status, the relationship between people buried in monuments such as these and the Pictish uppermost elite is uncertain. The cemetery evidences a wide diversity of cultural practices [36], mirrored in the high mitochondrial diversity, suggesting relatively high levels of mobility within the Pictish social structure at this level of society. The burials are organised in complex and stand-alone graves, made of round and square cairns and long cists (S1 Fig). This complexity suggests that, as social practices influence the genetic structure of populations, the social status of archaeological sites can, in turn, bias our understanding of population structure; in this case, the samples may only be representative of a small proportion of the overall Pictish population. Non-harmonious kinship systems (i.e., patrilineal and patrilineal or matrilineal and matrilineal societies) may also impact the genome in different ways. The lack of broad sample size and useful markers (Y-chromosome) to enhance kinship- and mtDNA-based findings remains an obstacle to illuminate further Pictish descent patterns.

## Conclusions

Our study provides novel insight into genetic affinity between ancient and modern populations of the British Isles, a rare opportunity to directly observe micro-scale evolution. High-quality genomes of two individuals buried in Scotland from the Pictish period, one from Balintore (BAL003) and one from Lundin Links (LUN004), reveal a close genetic affinity to Iron Age populations from Britain but with evidence of some genetic differentiation between samples. Overall, our data supports the current archaeological consensus arguing for regional continuity between the Late Iron Age and early medieval periods, but likely with complex patterns of migration, lifetime mobility and admixture. We also show that BAL003 and LUN004 were genetically differentiated from the pre-Viking Age Picts from Orkney, which suggests that Pictish culture spread to Orkney from Scotland primarily via cultural diffusion rather than direct population movement or inter-marriage. We detect strong continuity between local Iron Age and present-day peoples in

Orkney but less pronounced affinity between early medieval and modern people in eastern Scotland. More ancient genomes from the Iron Age and early medieval periods in the UK are necessary to illuminate these relationships further, combined with analyses of lifetime mobility using complementary approaches (e.g., isotope analysis). On a more local level, our mtDNA analysis of individuals interred at Lundin Links is arguably inconsistent with matrilocality. This finding challenges the older hypothesis that Pictish succession was based on a matrilineal system, assuming that wider Pictish society was organised in such a manner.

## Materials and methods

### DNA extraction, library preparation and sequencing

All aDNA work was carried out in dedicated facilities at Stockholm University. The samples were decontaminated by removing the outer surfaces via abrasion using a Dremel drill after a thorough cleaning in 1% sodium hypochlorite, followed by wiping with molecular biology-grade water and ethanol, and UV irradiation at 254 nm in a crosslinker (Ultra-Violet Products Ltd., Cambridge, UK) for 10 min each side at a distance <10 cm. Approximately 100–200 mg of bone or tooth (dentine) powder was extracted from each specimen using a Dremel drill at the lowest possible rotation-per-minute (5000 rpm).

DNA was extracted using 1 mL extraction buffer consisting of 0.45 mL EDTA (pH8), 1 M urea, and 10 uL of 10 mg/ml proteinase K. The mixture was incubated overnight (~18 hrs) at 37°C and purified on Qiagen MinElute columns following the manufacturer's recommendation, but with an additional wash step. DNA was eluted in a 63 uL Qiagen Elution buffer. Illumina-compatible sequencing libraries were constructed following Meyer and Kircher (2010) [67] as outlined in Rodríguez-Varela et al. (2017) [68] and sequenced on an Illumina HiSeq2000 platform. Non damage repair libraries from BAL003-b1e1l1 and LUN004-b1e1l1 were sequenced over three and one lanes, respectively. BAL003 and LUN004 libraries were generated using enzymatic damage repair [35], from the same extracts as BAL003-b1e1l1 and LUN004-b1e1l1, respectively, and sequenced over 5 lanes for LUN004 and 15 lanes for BAL003. The libraries for the remaining samples were not damage-repaired. LUN001 and LUN003 libraries were sequenced over seven lanes each. LUN005, LUN006, LUN009 libraries were sequenced over one lane.

### Sequence processing and alignment

We discarded reads with indexes showing at least one mismatch. Read pairs were merged and adapter sequence removed using Adapter Removal v2.1.7 [69], with a minimum overlap of 11 bp and summing base qualities in overlapping regions. Merged read pairs were mapped as single-end reads to the human reference genome build 37 with decoy sequences (hs37d5) using BWA aln v0.7.8 [70] with the non-default parameters -n 0.01 (maximum edit distance) and -o 2 (maximum number of gap opens), allowing more mismatches and indels, and disabled seeding with -l 16500 as in Lazaridis et al. (2014) [1] and Skoglund et al. (2014) [71]. We collapsed duplicate reads having identical start and end coordinates into consensus reads using FilterUniqueSAMCons.py [72]. Finally, we filtered the alignment so that only reads longer than 35 bp, with mapping quality >30, not containing indels, and with more than 90% matches with the reference were retained. We merged libraries sequenced over several lanes using SAMTOOLS v1.9 [70]. Summary statistics of the obtained reads are presented in S1 Table.

### Identification of authentic aDNA molecules and contaminant DNA

We used MapDamage v2.0 [73] to visualise the substitution distribution along the reads and evidence the presence of damaged aDNA molecules. Contamination was estimated using three



different data sources, namely: 1) the mitochondrial genome, 2) X-chromosome contamination in males, and 3) Y-chromosome contamination in females. We estimated present-day mtDNA-based contamination using Schmutzi [74]. For males we used ANGSD [75], which utilises heterozygous calls on the X-chromosome of male samples, expected to be haploid, to estimate contamination. For females, Y-chromosome contamination was assessed by comparing the observed number of reads mapped on the non-pseudo-autosomal region of the Y-chromosome with the expected number if the sample was male. The expected number of Y-chromosome reads is approximated as half the number of reads mapping to the autosomes multiplied by the Y-chromosome fraction of the genome; this assumes the alignment efficiency for the Y-chromosome and autosomes are similar. The Y-chromosome makes up 2% of the genome [76].

### Sex determination

The biological sex of sequenced individuals was determined using the  $R_y$  parameters [77].  $R_y$  is the fraction of the Y-chromosome alignments ( $n_y$ ) compared to the total number of reads aligned to the X- and Y-chromosomes ( $n_x + n_y$ ). The 95% confidence interval (CI) was computed as  $R_y \pm \frac{1.96 \times R_y \times (1-R_y)}{(n_y + n_x)}$ . This method determines whether an ancient individual can be determined to be male or female. If the lower CI limit of  $R_y$  is  $>0.077$  the individual is assigned as male. If the upper CI limit of  $R_y$  is  $<0.016$  the individual is assigned as female.

### Mitochondrial and Y-chromosome haplogroups

We obtained the consensus mitochondrial DNA from endogenous reads, removing the bases with quality  $<30$  ( $-q\ 30$ ) using Schmutzi [74]. The mitochondrial haplogroups were assigned using Haplogrep2 [78]. The Y-chromosome haplogroup was obtained using pathPhynder with the parameter  $-m$  'no-filter' [79], based on approximately 120,000 SNPs present in a dataset of Y-chromosomal variation in worldwide present-day and ancient males and the International Society Of Genetic Genealogy (ISOGG, <http://www.isogg.org>).

### Genomic analysis of pseudo-haploid sequences

The pseudo-haploid genomes of BAL003 and LUN004 were analysed with a set of published pseudo-haploid ancient genomes, the Human Origins dataset and the Simon Genome Diversity Project from the Allen Ancient DNA Resource (<https://reich.hms.harvard.edu/>) using PCA [80], ADMIXTURE analysis [81],  $D$ - and  $f$ -statistics [80, 82] and  $qpAdm$  [4] (section S1.3 in S1 Text).

### Imputation

Genomes of 33 individuals with a coverage  $>0.7X$  from nine geographic regions were imputed using GLIMPSE [83], namely: 1) two early medieval individuals from Scotland dated from the Pictish period (BAL003 and LUN004, this study); 2) three Iron Age (I0789, I0156 and I0160) [17], four Roman British (6DT3, 6DT18, 6DT21 and 6DT22) [14] and eight early medieval individuals from England (I0769, I0773, I0774, I0777, I0157, I0159, I0161 and NO3423) [14, 17]; 3) one Iron Age individual from Norway, 4) six Iron Age individuals from Sweden; 5) three early medieval individuals from Hungary (SZ15, SZ43 and SZ45) [15]; 6) two early medieval individuals from Germany (Alh1 and Alh10) [16]; 7) one Iron Age/early medieval individual from Slovakia (DA119) [84]; 8) one early medieval individual from the Czech Republic (RISE569) [3]; 9) two pre-Christian individuals from Iceland (SBT-A1 and DAV-A8) [85] (section S1.4 in S1 Text and S7 Table).

## Genotype phasing

The EU and UK datasets were phased together with the genomes from Margaryan et al. [9] (S7 Table and section S1.5 in S1 Text) and the re-imputed and newly imputed ancient genomes using BEAGLE 5.2 [86]. We restricted the phasing on the intersections of the genotypes newly imputed in this study and those imputed in Margaryan et al. [9] to prevent sporadic missing genotypes imputation by BEAGLE 5.2 [86]. The window and overlap lengths were set as wider than any chromosome (window length 380 cM and overlap length 190 cM) to maximise the information used for phasing the genomes. The 1000 genomes phase 3 dataset ([http://bochet.gcc.biostat.washington.edu/beagle/1000\\_Genomes\\_phase3\\_v5a](http://bochet.gcc.biostat.washington.edu/beagle/1000_Genomes_phase3_v5a)) and GRCh37 genomic maps ([http://bochet.gcc.biostat.washington.edu/beagle/genetic\\_maps/](http://bochet.gcc.biostat.washington.edu/beagle/genetic_maps/)) provided by BEAGLE were used for phasing. Imputation and genotyping errors can increase phasing errors. However, the BEAGLE phasing algorithm (Hidden Markov Model-based haplotype clustering) improves widely as the sample size increases. The improvement due to the sample size minimises the phasing error from possible genotyping and imputation biases.

## Reference panel

We compiled four datasets (section S1.5 in S1 Text), 1) a set of 1,764 modern individuals from 18 worldwide populations from the 1000 genomes project phase 3 [87], 2) a set of 10,299 modern individuals living in Europe from the EGAD00000000120 International Multiple Sclerosis Genetics Consortium & The Wellcome Trust Case Control Consortium 2 dataset [88], 3) a set of 2,578 modern individuals born in the UK and whose four grandparents were born in the UK within 80 km of each other from the EGAD00010000632 People of the British Isles dataset [11, 53] and 4) a set of 252 ancient European genomes dated from around the Viking Age period [9]. The ancient genomes in this dataset, together with those selected for imputation, were also pseudo-haploidised prior to imputation following the S1.3 in S1 Text method (i.e., to cross-check some results based on the imputed genomes).

## Principal Component Analysis

A PCA was generated using the re-imputed and newly imputed ancient genomes, the ancient genomes from Margaryan et al., (2020) [9] and the modern EU and UK dataset using PLINK v1.9 [89, 90]. SNPs with minor allele frequency (maf) <5% and in linkage disequilibrium (—indep-pairwise 100 10 0.2) were excluded with PLINK v1.9 [89,90]. The PCA was generated on 8,389 individuals and 88,040 SNPs.

A second PCA was generated to increase variability around the ancient British genomes. It was restricted to the modern individuals from Belgium, Denmark, Germany, Norway and the UK, the ancient genomes from the Iron Age and Viking Age individuals from Orkney and re-imputed and the newly imputed ancient genomes in this study from the British Isles, Iceland, Scandinavia, Germany, Slovakia and the Czech Republic following the same method as above. The PCA was generated on 4,914 individuals and 87,518 SNPs.

## ADMIXTURE

Ancestry components were estimated in modern individuals from Belgium, Denmark, Germany, Norway, and ancient genomes from the Iron Age and Viking Age individuals from Orkney and the newly imputed ancient genomes from the British Isles, Scandinavia, Germany, Slovakia and Czech Republic using the program ADMIXTURE v1.2 [81]. Sites with maf <5% and SNPs in linkage disequilibrium (—indep-pairwise 100 10 0.2) were removed with PLINK

v1.9 [89,90], leaving 4,912 individuals and 75,295 SNPs. ADMIXTURE was run with cross-validation (CV) enabled using -cv flag and 50 bootstraps from  $K = 2$  to  $K = 10$ .

### D- and F-statistics

Individual or population relatedness was performed using *outgroup-f3* statistics with the qp3Pop function from ADMIXTOOLS 2 R package (<https://github.com/uqrmaie1/admixtools>). *F3-statistics*  $f3(\text{Yoruba}; A, B)$  measure allele frequency correlation between populations. When X is an equidistant outgroup to A and B, *outgroup-f3* becomes a genetic drift measure between A and B. The outgroup Yoruba is expected to be equidistant from all tested samples.

To obtain information on individuals or population admixture, we performed outgroup *D-statistics* of the form  $D(A, B; C, \text{Yoruba})$  using the qpDstats function from ADMIXTOOLS 2 R package (<https://github.com/uqrmaie1/admixtools>). A, B and C are either present-day or ancient populations/individuals. A result equal to 0 means that the proposed tree  $((A, B), C)$ , Yoruba) is consistent with the data. If D deviates from 0, there are more alleles shared than expected given the proposed tree either between A and C ( $D > 0$ ) or between B and C ( $D < 0$ ).

### Identity-By-Descent and Homozygosity-By-Descent

The identification of IBD and HBD segments was done using RefinedIBD [59]. The window size was set to 3 cM. The minimal size for a segment to be considered shared by IBD or HBD is 1 cM, the same threshold used in Margaryan et al. [9]. We decided to consider segments  $>1$  cM as shared by IBD since 1 cM corresponds to the timespan between the oldest samples from the Iron Age and present-day populations. A common ancestor  $n$  generations in the past ( $2n$  meiosis) results on average in  $100/2n$  cM segment length [91]. Thus, IBD segments longer than 1 cM derive from common ancestors living  $\sim 50$  generations in the past, or  $\sim 1,500$  years ago; this assumes an average human generation time of 30 years [92], which is the period roughly including samples from the Iron Age until the present. The total number and total length of shared IBD segments were generated. To avoid sample size bias, we used bootstrapping to compare total number and length of shared IBD over an equal number of individuals when using modern populations with unequal numbers of individuals. We randomly selected an equal number of modern individuals across populations to calculate the total number and length of shared IBD and averaged using 100 replicates with replacement (section S1.6 in S1 Text).

Additionally, we generated interpolated frequency maps of the total number of shared IBD between modern populations and ancient individuals/populations from the UK with QGIS v3.14.1 [93] using distance coefficient  $P = 2$  and pixel size = 0.01. We used the county town as a proxy for the county geographic coordinate.

### ChromoPainter and FineSTRUCTURE

ChromoPainter is a tool for finding the closest haplotypes in sequence data where each individual is painted as a combination of all other genomes using the Li-Stephen Model [94]. ChromoPainter paints the genome of each 'recipient' using all the remaining individuals as 'donors'. We used ChromoPainter v2.9.0 [42] to paint all of the ancient individuals ( $n = 284$ ). We realised the PCAs based on the ChromoPainter co-ancestry matrix using FineSTRUCTURE GUI.

FineSTRUCTURE v2.9.0 [42] was then used to perform population assignment based on the ChromoPainter coancestry matrix. In brief, FineSTRUCTURE is similar in concept to ADMIXTURE but accesses a large number of SNPs and linkage disequilibrium (haplotype) information. FineSTRUCTURE uses an MCMC approach to partition the dataset into  $K$

groups with indistinguishable genetic ancestry, interpreted as individual populations. The program was run using 100,000 iterations and 100,000 burn-in.

## Supporting information

**S1 Table. DNA sample information from eight Pictish genomes in this study.**

(XLSX)

**S2 Table. Sample misassignment estimate in LUN001 and LUN003 using qpAdm.**

(XLSX)

**S3 Table. Sample misassignment rate in LUN001 and LUN003 mtDNA reads.**

(XLSX)

**S4 Table. Calibrated  $^{14}\text{C}$  radiocarbon dates.**

(XLSX)

**S5 Table. Mitochondrial DNA mutations and haplogroup attribution.**

(XLSX)

**S6 Table. Y-chromosome mutations on LUN004.**

(XLSX)

**S7 Table. Comparison datasets for the analysis of the pseudo-haploid genomes and imputed genomes from Margaryan et al. [9].**

(XLSX)

**S8 Table. Shared genetic drift between the BAL003, LUN004 and modern populations as  $f_3(\text{BAL003/LUN004, modern; Mbuti})$ .**

(XLSX)

**S9 Table.  $D(\text{Iron Age/Roman English, Iron Age/Medieval European; BAL003/LUN004, Mbuti})$ .**

(XLSX)

**S10 Table.  $D(\text{Iron Age/Roman English, BAL003/LUN004; modern and ancient populations, Mbuti})$ .**

(XLSX)

**S11 Table. Coancestry matrix obtained with ChromoPainter for the analysis of 284 ancient European genomes dated from the Iron Age to medieval period.**

(XLSX)

**S12 Table. Genotypes associated with lactase persistence and skin, eyes and hair pigmentation with subsequent prediction performance metrics from HIrisPlex-S.**

(CSV)

**S1 Fig. Plan of cists at Lundin Links.** The plan was published in Greig et al. [36]. The samples yielding DNA are in red.

(TIF)

**S2 Fig. Calibrated  $^{14}\text{C}$  radiocarbon dates.** Radiocarbon determinations calibrated using OxCal. v4.4 [95] and the IntCal20 atmospheric curve [96].

(PNG)

**S3 Fig. Deamination pattern along the reads.**

(TIF)

**S4 Fig. Biological sex.** We determined the biological sex based on Y-chromosome alignments (ny) compared to the total number of reads of the X and Y-chromosomes (nx + ny). The 95% confidence interval (CI) was computed as  $Ry \pm (1.96 \times Ry \times (1 - Ry)) / ((ny + nx))$ .

(TIF)

**S5 Fig. Genetic distance between BAL003, LUN004, Iron Age and Roman individuals from England and Iron and Viking Age individuals from Orkney (excluding VK204 and VK205) measured using outgroup-f3 statistics.**

(TIF)

**S6 Fig. Cross-validation error of the ADMIXTURE analyses.** The cross-validation was done using the—cv option of admixture runs for K = 1 to K = 10 with 4,914 individuals and 87,518 SNPs using 50 bootstraps (-B 50).

(TIF)

**S7 Fig. ADMIXTURE Model-based clustering analysis of 4,914 modern and ancient Northern and Central Europeans from the Iron Age, Early Medieval period and present-day (K = 2 to K = 10).**

(TIF)

**S8 Fig. Principal Component Analysis of the 8,389 present-day and ancient individuals from this study.** SNPs with maf < 5% were removed and pruned (88,040 SNPs remained). PC3 is impacted by allele frequency bias differentiating the data imputed in Margaryan et al. [9] and the genomes newly imputed in this study.

(TIF)

**S9 Fig. Correlation between coverage and PCs coordinates from the PCA presented in S8 Fig.** The genomes newly imputed in this study are in blue and the genomes imputed in Margaryan et al. [9] are in pink.

(TIFF)

**S10 Fig. Genetic affinity of BAL003 and LUN004 compared to western Eurasians using PCA.**

Pseudo-haploid genotypes from ancient samples are projected onto the first two principal components defined by 1,056 present-day West Eurasians from the 'HO' dataset (S). VA, Viking Age.

(TIFF)

**S11 Fig. Cross-validation error of the ADMIXTURE analysis.** The cross-validation was done using the—cv option of admixture runs for K = 3 to K = 20 with 3,591 individuals and 85,655 markers.

(TIF)

**S12 Fig. ADMIXTURE ancestry component (K = 15) of ancient shotgun and present-day worldwide individuals.** Full displays of the ADMIXTURE analysis are in S13–S15 Figs.

(TIFF)

**S13 Fig. Model-based clustering analysis of 3,594 individuals (K = 3 to K = 20).** Only modern individuals are represented.

(TIFF)

**S14 Fig. Model-based clustering analysis of 3,594 individuals (K = 3 to K = 20).** Only ancient individuals from western Eurasia are represented.

(TIFF)



**S15 Fig. Model-based clustering analysis of 3,594 individuals ( $K = 3$  to  $K = 20$ ).** Only ancient individuals from Eastern Eurasia are represented.

(TIFF)

**S16 Fig. Testing for symmetry between the pseudo-haploid BAL003 and LUN004 genomes and the Iron Age or Roman period individuals from England relative to other ancient and modern populations.** We tested for symmetry as D(Pict, ancient England; X, Mbuti) and plotted the resulting Z-scores. Details on the test results and sample size are in [S10 Table](#).

(TIF)

**S17 Fig. Testing for continuity between European Iron Age/medieval period and BAL003 and LUN004 pseudo-haploid genomes using qpAdm in a one-way admixture model.**

(TIF)

**S18 Fig. Modelling Bronze Age, Iron Age and medieval European populations as a mixture of Western Hunters-Gatherers (WHG), Anatolian Neolithic (Anatolia\_N) and Bronze Age Yamnaya using qpAdm on pseudo-haploid genomes.** Empty bars are target populations for which the three-way models' proportion was impossible to estimate using qpAdm. Vik\_97002.SG, vik\_97026.SG, vik\_urm045.SG and vik\_urm161.SG are individuals buried in a Viking context in Sigtuna, Sweden but likely migrants of diverse origins. G, generation.

(TIF)

**S19 Fig. Shared Identity-By-Descent (IBD) segments between and within present-day European populations.** A) Total length of shared IBD segments  $>1$  cM, B) total number of shared IBD segments  $>1$  cM, C) total number of shared IBD segments  $>4$  cM and D) total number of shared IBD segments  $>6$  cM. The number corresponds to the mean of 100 bootstraps drawing 44 random individuals per population.

(TIF)

**S20 Fig. Shared Identity-By-Descent (IBD) segments between ancient genomes and modern European populations.** The ancient genomes are the newly imputed genomes and the six ancient Orcadians from Margaryan et al. [9]. A) Total length of shared IBD segments  $>1$  cM, B) total number of shared IBD segments  $>1$  cM, C) total number of shared IBD segments  $>4$  cM and D) total number of shared IBD segments  $>6$  cM. The number corresponds to the mean of 100 bootstraps drawing 44 random individuals per European population.

(TIF)

**S21 Fig. Shared Identity-By-Descent (IBD) segments between ancient individuals from Britain.** A) Total length of shared IBD segments  $>1$  cM, B) total number of shared IBD segments  $>1$  cM, C) total number of shared IBD segments  $>4$  cM and D) total number of shared IBD segments  $>6$  cM. IA, Iron Age; VA, Viking Age; Emedieval, early medieval.

(TIF)

**S22 Fig. Number of Identity-By-Descent (IBD) segments shared between present-day and ancient genomes from Britain.**

(TIF)

**S23 Fig. Distribution of Identity-By-Descent (IBD) length as a function of the time difference between pairs of samples.**

(TIF)

**S24 Fig. Principal Component Analysis (PCA) of the 8,044 present-day Europeans individuals.** The pseudo-haploid ancient genomes ([S7 Table](#)) were projected. This analysis replicates

[S8 Fig](#) but only using modern genomes for the PCs computation and projected pseudo-haploid ancient genomes. SNPs with maf <5% were removed and pruned (88,369 SNPs remained). (TIFF)

**S25 Fig. Principal Component Analysis (PCA) of the 4,863 present-day western Europeans individuals.** The pseudo-haploid ancient genomes ([S7 Table](#)) were projected. This analysis replicates [Fig 2A](#), but only using modern genomes for the PCs computation and projected pseudo-haploid ancient genomes. SNPs with maf <5% were removed and pruned (88,369 SNPs remained). (TIFF)

**S26 Fig. Distribution of Homozygosity-By-Descent (HBD) across ancient individuals from Britain.** (TIF)

**S27 Fig. PCA based on ChromoPainter coancestry matrix and FineSTRUCTURE clustering generated for 284 ancient European genomes dated to the Iron Age and medieval period.** (TIF)

**S1 Text. Supplementary information.** (DOCX)

## Acknowledgments

We thank Gareth Weedall at Liverpool John Moores University for his valuable help with editing the manuscript, Jeanette Pearson at Inverness Museum & Art Gallery for facilitating sampling of Balintore material, Jean-Jacques Hublin, Mike Richards (at the time both MPI-EVA), and the Max Planck Society for laboratory access and resources in kind and Orsolya Czére at University of Aberdeen for support with C:N determination.

## Author Contributions

**Conceptualization:** Adeline Morez, Kate Britton, Gordon Noble, Torsten Günther, Anders Götherström, Ricardo Rodríguez-Varela, Nicholas J. Evans, Linus Girdland-Flink.

**Data curation:** Adeline Morez, Kate Britton, Linus Girdland-Flink.

**Formal analysis:** Adeline Morez, Gordon Noble, Rui Martiniano.

**Funding acquisition:** Anders Götherström, Linus Girdland-Flink.

**Investigation:** Adeline Morez, Kate Britton, Ricardo Rodríguez-Varela, Natalija Kashuba, Nicholas J. Evans, Linus Girdland-Flink.

**Methodology:** Adeline Morez, Torsten Günther, Sahra Talamo, Linus Girdland-Flink.

**Project administration:** Adeline Morez, Linus Girdland-Flink.

**Resources:** Kate Britton, Gordon Noble, Anders Götherström, Rui Martiniano, Sahra Talamo, Christina Donald, Linus Girdland-Flink.

**Software:** Adeline Morez, Rui Martiniano.

**Supervision:** Anders Götherström, Linus Girdland-Flink.

**Validation:** Adeline Morez.

**Visualization:** Adeline Morez.

**Writing – original draft:** Adeline Morez, Joel D. Irish, Linus Girdland-Flink.

**Writing – review & editing:** Adeline Morez, Kate Britton, Gordon Noble, Torsten Günther, Ricardo Rodríguez-Varela, Natalija Kashuba, Rui Martiniano, Sahra Talamo, Nicholas J. Evans, Joel D. Irish, Christina Donald, Linus Girdland-Flink.

## References

1. Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Kirsanow K, et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*. 2014; 513(7518):409–13. <https://doi.org/10.1038/nature13673> PMID: 25230663
2. Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, et al. Genomic insights into the origin of farming in the ancient Near East. *Nature*. 2016; 536(7617):419–24. <https://doi.org/10.1038/nature19310> PMID: 27459054
3. Allentoft ME, Sikora M, Sjögren KG, Rasmussen S, Rasmussen M, Stenderup J, et al. Population genomics of Bronze Age Eurasia. *Nature*. 2015; 522(7555):167–72. <https://doi.org/10.1038/nature14507> PMID: 26062507
4. Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*. 2015; 522(7555):207–11. <https://doi.org/10.1038/nature14317> PMID: 25731166
5. Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*. 2015; 528(7583):499–503. <https://doi.org/10.1038/nature16152> PMID: 26595274
6. Broushaki F, Thomas MG, Link V, López S, van Dorp L, Kirsanow K, et al. Early Neolithic genomes from the eastern Fertile Crescent. *Science*. 2016; 353(6298):499–503. <https://doi.org/10.1126/science.aaf7943> PMID: 27417496
7. Olalde I, Brace S, Allentoft ME, Armit I, Kristiansen K, Booth T, et al. The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature*. 2018; 555(7695):190–6. <https://doi.org/10.1038/nature25738> PMID: 29466337
8. Brace S, Diekmann Y, Booth TJ, van Dorp L, Faltyskova Z, Rohland N, et al. Ancient genomes indicate population replacement in Early Neolithic Britain. *Nature Ecology and Evolution*. 2019; 3(5).
9. Margaryan A, Lawson DJ, Sikora M, Racimo F, Rasmussen S, Moltke I, et al. Population genomics of the Viking world Scandinavian ancestry and Viking Age origins Check for updates. *Nature*. 2020; 585(7825):390.
10. Patterson N, Isakov M, Booth T, Büster L, Fischer CE, Olalde I, et al. Large-scale migration into Britain during the Middle to Late Bronze Age. *Nature*. 2021; 601(7894):1–14. <https://doi.org/10.1038/s41586-021-04287-4> PMID: 34937049
11. Leslie S, Winney B, Hellenthal G, Davison D, Bourmertit A, Day T, et al. The fine-scale genetic structure of the British population. *Nature*. 2015; 519(7543):309–14. <https://doi.org/10.1038/nature14230> PMID: 25788095
12. Armit I, Reich D. The return of the Beaker folk? Rethinking migration and population change in British prehistory. *Antiquity*. 2021; 95(384):1464–77.
13. Dulas K, Foody MGB, Justeau P, Silva M, Martiniano R, Oteo-García G, et al. Ancient DNA at the edge of the world: Continental immigration and the persistence of Neolithic male lineages in Bronze Age Orkney. *Proceedings of the National Academy of Sciences*. 2022; 119(8):e2108001119. <https://doi.org/10.1073/pnas.2108001119> PMID: 35131896
14. Martiniano R, Caffell A, Holst M, Hunter-Mann K, Montgomery J, Müldner G, et al. Genomic signals of migration and continuity in Britain before the Anglo-Saxons. *Nature Communications*. 2016; 7:1–8. <https://doi.org/10.1038/ncomms10326> PMID: 26783717
15. Amorim CEG, Vai S, Posth C, Modi A, Koncz I, Hakenbeck S, et al. Understanding 6th-century barbarian social organization and migration through paleogenomics. *Nature Communications*. 2018; 1–11.
16. Veeramah KR, Rott A, Groß M, Dorp LV, López S, Kirsanow K, et al. Population genomic analysis of elongated skulls reveals extensive female-biased immigration in Early Medieval Bavaria. *Proceedings of the National Academy of Sciences of the United States of America*. 2018; 115(13):3494–9. <https://doi.org/10.1073/pnas.1719880115> PMID: 29531040

17. Schiffels S, Haak W, Paajanen P, Llamas B, Popescu E, Loe L, et al. Iron Age and Anglo-Saxon genomes from East England reveal British migration history. *Nature Communications*. 2016; 7:1–9. <https://doi.org/10.1038/ncomms10408> PMID: 26783965
18. Gretzinger J, Sayer D, Justeau P, Altena E, Pala M, Dulias K, et al. The Anglo-Saxon migration and the formation of the early English gene pool. *Nature*. 2022; 610(7930):112–9. <https://doi.org/10.1038/s41586-022-05247-2> PMID: 36131019
19. Thorpe B, editor. *The Anglo-Saxon Chronicle*. Vol. 1. Cambridge: Cambridge University Press; 2012.
20. Charles-Edwards TM. *Wales and the Britons, 350–1064*. Oxford: Oxford University Press; 2013. (History of Wales).
21. Evans N. Picti: from Roman name to internal identity. *Journal of Medieval History*. 2022; 48(3):291–322.
22. Wainwright FT, editor. *The Problem of the Picts*. Edinburgh: Thomas Nelson, 1995.
23. Noble G, Gondek M, Campbell E, Cook M. Between prehistory and history: The archaeological detection of social change among the Picts. *Antiquity*. 2013; 87(338):1136–50.
24. Jackson KH. The Pictish language. In: *The problem of the Picts*. Frederick T. Wainwright. Edinburgh: Thomas Nelson; 1955. p. 126–60, 161–6, 173–6.
25. Kidd C. The ideological Uses of the Picts, 1707–c. 1990. In: *Scottish History: The Power of the Past*. Edward J. Cowan and Richard J. Finlay, editor. Edinburgh: Edinburgh University Press; 2002. p. 169–90.
26. Forsyth K. *Language in Pictland: the case against “non-Indo-European Pictish.”* Utrecht: De Keltische Draak; 1997.
27. Taylor S. Pictish place-names revisited. In: Driscoll ST, Geddes J, Hall MA, editors. *Leiden, The Netherlands: Brill*; 2011. p. 67–118.
28. Rhys G. *Approaching the Pictish language: historiography, early evidence and the question of Pritenic* [PhD]. University of Glasgow; 2015.
29. Isaac GR. Scotland. In: *New Approaches to Place-Names in Ptolemy’s Geography*. de Hoz J, Lújan ER, Sims-Williams P, editors. Madrid: Ediciones Clásicas; 2005. p. 189–214.
30. Rodway S. The Ogham Inscriptions of Scotland and Brittonic Pictish. *Journal of Celtic Linguistics*. 2020; 21(1):173–234.
31. Evans N. Ideology literacy and matriliney: approaches to medieval texts on the Pictish past. In: *Pictish progress New studies on Northern Britain in the Early Middle Ages*. Driscoll S.T., Hall M. and Geddes J. Leiden, The Netherlands: Brill; 2011. p. 45–65.
32. Woolf A. Pictish matriliney reconsidered. *Innes Review*. 1998; 49(2):147–67.
33. Ross A. Pictish Matriliney. *Northern Studies*. 1999; 34:11–22.
34. Evans N. Royal succession and kingship among the Picts. *Innes Review*. 2008; 59(1):1–48.
35. Briggs AW, Stenzel U, Meyer M, Krause J, Kircher M, Pääbo S. Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Research*. 2010; 38(6):e87. <https://doi.org/10.1093/nar/gkp1163> PMID: 20028723
36. Greig C, Greig M, Ashmore P, Campbell-Wilson M, Lorimer D, Smart I, et al. Excavation of a cairn cemetery at Lundin Links, Fife, in 1965–6. *Proceedings of the Society of Antiquaries of Scotland*. 2000; 130:585–636.
37. Fraser JE. *From Caledonia to Pictland: Scotland to 795*. Edinburgh: Edinburgh University Press; 2009.
38. Taylor S, Clancy TO, McNiven P, Williamson E. *The Place-Names of Clackmannanshire*. Donington: Shaun Tyas; 2020.
39. McNiven P. *Gaelic place-names and the Social History of Gaelic Speakers in Medieval Menteith* [PhD]. University of Glasgow; 2011.
40. Sánchez-Quinto F, Malmström H, Fraser M, Girdland-Flink L, Svensson EM, Simões LG, et al. Megalithic tombs in western and northern Neolithic Europe were linked to a kindred society. *Proceedings of the National Academy of Sciences*. 2019; 116(19):9469–74. <https://doi.org/10.1073/pnas.1818037116> PMID: 30988179
41. Balaresque P, Bowden GR, Adams SM, Leung HY, King TE, Rosser ZH, et al. A Predominantly Neolithic Origin for European Paternal Lineages. *PLOS Biology*. 2010; 8(1):e1000285. <https://doi.org/10.1371/journal.pbio.1000285> PMID: 20087410
42. Lawson DJ, Hellenthal G, Myers S, Falush D. Inference of Population Structure using Dense Haplotype Data. *Copenhagen GP, editor. PLoS Genetics*. 2012; 8(1):e1002453.
43. Ralph P, Coop G. The Geography of Recent Genetic Ancestry across Europe. *PLOS Biol*. 2013 May 7; 11(5):e1001555. <https://doi.org/10.1371/journal.pbio.1001555> PMID: 23667324

44. Al-Asadi H, Petkova D, Stephens M, Novembre J. Estimating recent migration and population-size surfaces. DeGiorgio M, editor. *PLOS Genet.* 2019 Jan 14; 15(1):e1007908. <https://doi.org/10.1371/journal.pgen.1007908> PMID: 30640906
45. Lawson DJ, van Dorp L, Falush D. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nat Commun.* 2018; 9(1):1–11.
46. Gilbert E O'Reilly S, Merrigan M, McGettigan D, Vitart V, Joshi PK, et al. The genetic landscape of Scotland and the Isles. *Proceedings of the National Academy of Sciences.* 2019; 116(38): 19064–19070. <https://doi.org/10.1073/pnas.1904761116> PMID: 31481615
47. Byrne RP, Martiniano R, Cassidy LM, Carrigan M, Hellenthal G, Hardiman O, et al. Insular Celtic population structure and genomic footprints of migration. *PLoS Genetics.* 2018; 14(1):1–22. <https://doi.org/10.1371/journal.pgen.1007152> PMID: 29370172
48. Oram R. *Domination and Lordship. Scotland 1070–1230.* Edinburgh: Edinburgh University Press; 2011.
49. Barrow GWS. *The Kingdom of the Scots, 2nd edn.* Edinburgh: Edinburgh University Press; 2003.
50. Taylor S, Márkus G. *The Place-Names of Fife.* Donington: Shaun Tyas; 2012. p. 169–80.
51. Czére O, Fawcett J, Evans J, Sayle K, Müldner G, Hall M, et al. Multi-isotope analysis of the human skeletal remains from Blair Atholl, Perth and Kinross, Scotland. *Tayside and Fife Archaeological Journal.* 2021; 27:31–44.
52. Czére O, Lawson JA, Müldner G, Evans J, Boyle A, Britton K. The Bodies in the 'Bog': A Multi-Isotope Investigation of Individual Life-Histories at an Unusual 6th/7th AD Century Group Burial from a Roman Latrine at Cramond, Scotland. *Archaeological and Anthropological Sciences.* 2022; 14(4):67.
53. Winney B, Boumertit A, Day T, Davison D, Echeta C, Evseeva I, et al. People of the British Isles: Preliminary analysis of genotypes and surnames in a UK-control population. *European Journal of Human Genetics.* 2012; 20(2):203–10. <https://doi.org/10.1038/ejhg.2011.127> PMID: 21829225
54. Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva I, Blue-Smith J, et al. The Eurasian heartland: A continental perspective on Y-chromosome diversity. *Proceedings of the National Academy of Sciences of the United States of America.* 2001; 98(18):10244–9. <https://doi.org/10.1073/pnas.171305098> PMID: 11526236
55. Wilson JF, Weiss DA, Richards M, Thomas MG, Bradman N, Goldstein DB. Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proceedings of the National Academy of Sciences of the United States of America.* 2001; 98(9):5078–83. <https://doi.org/10.1073/pnas.071036898> PMID: 11287634
56. Capelli C, Redhead N, Abernethy JK, Gratrix F, Wilson JF, Moen T, et al. A Y chromosome census of the British Isles. *Current Biology.* 2003; 13(11):979–84. [https://doi.org/10.1016/s0960-9822\(03\)00373-7](https://doi.org/10.1016/s0960-9822(03)00373-7) PMID: 12781138
57. Goodacre S, Helgason A, Nicholson J, Southam L, Ferguson L, Hickey E, et al. Genetic evidence for a family-based Scandinavian settlement of Shetland and Orkney during the Viking periods. *Heredity.* 2005; 95(2):129–35. <https://doi.org/10.1038/sj.hdy.6800661> PMID: 15815712
58. Browning BL, Browning SR. A Fast, Powerful Method for Detecting Identity by Descent. *Am J Hum Genet.* 2011 Feb 11; 88(2):173–82. <https://doi.org/10.1016/j.ajhg.2011.01.010> PMID: 21310274
59. Browning BL, Browning SR. Improving the Accuracy and Efficiency of Identity-by-Descent Detection in Population Data. *Genetics.* 2013; 194(2):459–71. <https://doi.org/10.1534/genetics.113.150029> PMID: 23535385
60. Ringbauer H, Novembre J, Steinrücken M. Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nat Commun.* 2021 Sep 14; 12(1):5425. <https://doi.org/10.1038/s41467-021-25289-w> PMID: 34521843
61. Albrechtsen A, Moltke I, Nielsen R. Natural Selection and the Distribution of Identity-by-Descent in the Human Genome. *Genetics.* 2010 Sep; 186(1):295–308. <https://doi.org/10.1534/genetics.110.113977> PMID: 20592267
62. Oota H, Settheetham-Ishida W, Tiwawech D, Ishida T, Stoneking M. Human mtDNA and Y-chromosome variation is correlated with matrilineal versus patrilineal residence. *Nature Genetics.* 2001; 29(1):20–1. <https://doi.org/10.1038/ng711> PMID: 11528385
63. Hamilton G, Stoneking M, Excoffier L. Molecular analysis reveals tighter social regulation of immigration in patrilineal populations than in matrilineal populations. *Proceedings of the National Academy of Sciences of the United States of America.* 2005; 102(21):7476–7480. <https://doi.org/10.1073/pnas.0409253102> PMID: 15894624
64. Gunnarsdóttir ED, Nandineni MR, Li M, Myles S, Gil D, Pakendorf B, et al. Larger mitochondrial DNA than Y-chromosome differences between matrilineal and patrilineal groups from Sumatra. *Nature Communications.* 2011; 2(1):1–7. <https://doi.org/10.1038/ncomms1235> PMID: 21407194



65. Tumonggor MK, Karafet TM, Downey S, Lansing JS, Norquest P, Sudoyo H, et al. Isolation, contact and social behavior shaped genetic diversity in West Timor. *Journal of Human Genetics*. 2014; 59(9):494–503. <https://doi.org/10.1038/jhg.2014.62> PMID: 25078354
66. Murdock GP, White DR. Standard Cross-Cultural Sample. *Ethnology*. 1969; 8(4):329.
67. Meyer M, Kircher M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*. 2010; 5(6). <https://doi.org/10.1101/pdb.prot5448> PMID: 20516186
68. Rodríguez-Varela R, Günther T, Krzewińska M, Storå J, Gillingwater TH, MacCallum M, et al. Genomic Analyses of Pre-European Conquest Human Remains from the Canary Islands Reveal Close Affinity to Modern North Africans. *Current Biology*. 2017; 27(21):3396–3402. <https://doi.org/10.1016/j.cub.2017.09.059> PMID: 29107554
69. Schubert M, Lindgreen S, Orlando L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Research Notes*. 2016; 9(1):1–7. <https://doi.org/10.1186/s13104-016-1900-2> PMID: 26868221
70. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009; 25(14):1754–60. <https://doi.org/10.1093/bioinformatics/btp324> PMID: 19451168
71. Skoglund P, Malmström H, Omrak A, Raghavan M, Valdiosera C, Günther T, et al. Genomic diversity and admixture differs for stone-age Scandinavian foragers and farmers. *Science*. 2014; 344(6185):747–50. <https://doi.org/10.1126/science.1253448> PMID: 24762536
72. Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Research*. 2012; 40(1):e3. <https://doi.org/10.1093/nar/gkr771> PMID: 22021376
73. Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*. 2013; 29(13):1682–4. <https://doi.org/10.1093/bioinformatics/btt193> PMID: 23613487
74. Renaud G, Slon V, Duggan AT, Kelso J, Schmutzi: Estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biology*. 2015; 16(1):1–18.
75. Korneliussen TS, Albrechtsen A, Nielsen R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*. 2014; 15(356). <https://doi.org/10.1186/s12859-014-0356-4> PMID: 25420514
76. Howe KL, Achuthan P, Allen J, Allen J, Alvarez-Jarreta J, Amodé MR, et al. Ensembl 2021. *Nucleic Acids Research*. 2021; 49(D1):D884–91. <https://doi.org/10.1093/nar/gkaa942> PMID: 33137190
77. Skoglund P, Storå J, Götherström A, Jakobsson M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. *Journal of Archaeological Science*. 2013; 40(12):4477–82.
78. Weissensteiner H, Pacher D, Kloss-Brandstätter A, Forer L, Specht G, Bandelt HJ, et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Research*. 2016; 44(1):58–63. <https://doi.org/10.1093/nar/gkw233> PMID: 27084951
79. Martiniano R, De Sanctis B, Hallast P, Durbin R. Placing Ancient DNA Sequences into Reference Phylogenies. *Molecular Biology and Evolution*. 2022; 39(2):msac017. <https://doi.org/10.1093/molbev/msac017> PMID: 35084493
80. Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, et al. Ancient admixture in human history. *Genetics*. 2012; 192(3):1065–93. <https://doi.org/10.1534/genetics.112.145037> PMID: 22960212
81. Alexander DH, Novembre J. Fast Model-Based Estimation of Ancestry in Unrelated Individuals. *Genome Research*. 2009; 19(9):1655–64. <https://doi.org/10.1101/gr.094052.109> PMID: 19648217
82. Petr M, Vernot B, Kelso J. admixr—R package for reproducible analyses using ADMIXTOOLS. *Bioinformatics*. 2019; 35(17):3194–5. <https://doi.org/10.1093/bioinformatics/btz030> PMID: 30668635
83. Rubinacci S, Ribeiro DM, Hofmeister RJ, Delaneau O. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nature Genetics*. 2021; 53(1):120–6. <https://doi.org/10.1038/s41588-020-00756-0> PMID: 33414550
84. Damgaard PB, Marchi N, Rasmussen S, Peyrot M, Renaud G, Korneliussen T, et al. 137 ancient human genomes from across the Eurasian steppes. *Nature*. 2018; 557(7705):369–74. <https://doi.org/10.1038/s41586-018-0094-2> PMID: 29743675
85. Ebenesersdóttir SS, Sandoval-Velasco M, Gunnarsdóttir ED, Jagadeesan A, Guðmundsdóttir VB, Thordardóttir EL, et al. Ancient genomes from Iceland reveal the making of a human population. *Science*. 2018; 360(6392):1028–32. <https://doi.org/10.1126/science.aar2625> PMID: 29853688
86. Browning BL, Tian X, Zhou Y, Browning SR. Fast two-stage phasing of large-scale sequence data. *The American Journal of Human Genetics*. 2021; 108(10):1880–90. <https://doi.org/10.1016/j.ajhg.2021.08.005> PMID: 34478634

87. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global reference for human genetic variation. *Nature*. 2015; 526(7571):68–74. <https://doi.org/10.1038/nature15393> PMID: 26432245
88. International Multiple Sclerosis Genetics Consortium, Wellcome Trust Case Control Consortium 2, Sawcer S, Hellenthal G, Pirinen M, Spencer CCA, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*. 2011; 476(7359):214–9. <https://doi.org/10.1038/nature10251> PMID: 21833088
89. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*. 2007; 81(3):559–75. <https://doi.org/10.1086/519795> PMID: 17701901
90. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*. 2015; 4(1):7.
91. Browning SR. Estimation of Pairwise Identity by Descent From Dense Genetic Marker Data in a Population Sample of Haplotypes. *Genetics*. 2008; 178(4):2123. <https://doi.org/10.1534/genetics.107.084624> PMID: 18430938
92. Helgason A, Hrafnkelsson B, Gulcher JR, Ward R, Stefánsson K. A Populationwide Coalescent Analysis of Icelandic Matrilineal and Patrilineal Genealogies: Evidence for a Faster Evolutionary Rate of mtDNA Lineages than Y Chromosomes. *The American Journal of Human Genetics*. 2003. 72(6):1370–88. <https://doi.org/10.1086/375453> PMID: 12721957
93. QGIS Development Team. QGIS Geographic Information System. 2020. Available from: <http://qgis.osgeo.org>
94. Li N, Stephens M. Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*. 2003; 165(4):2213–33. <https://doi.org/10.1093/genetics/165.4.2213> PMID: 14704198
95. Bronk Ramsey C. Bayesian analysis of radiocarbon dates. *Radiocarbon*. 2009; 51(1):337–60.
96. Reimer PJ, Austin WEN, Bard E, Bayliss A, Blackwell PG, Bronk Ramsey C, et al. The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). *Radiocarbon*. 2020; 62(4):725–57.