

# Ancient genomes indicate population replacement in Early Neolithic Britain

Selina Brace<sup>1,15</sup>, Yoan Diekmann<sup>2,15</sup>, Thomas J. Booth<sup>1,15</sup>, Lucy van Dorp<sup>3</sup>, Zuzana Faltyskova<sup>2</sup>, Nadin Rohland<sup>4</sup>, Swapan Mallick<sup>3,5,6</sup>, Iñigo Olalde<sup>4</sup>, Matthew Ferry<sup>4,6</sup>, Megan Michel<sup>4,6</sup>, Jonas Oppenheimer<sup>4,6</sup>, Nasreen Broomandkshobacht<sup>4,6</sup>, Kristin Stewardson<sup>4,6</sup>, Rui Martiniano<sup>7</sup>, Susan Walsh<sup>8</sup>, Manfred Kayser<sup>9</sup>, Sophy Charlton<sup>10</sup>, Garrett Hellenthal<sup>3</sup>, Ian Armit<sup>11</sup>, Rick Schulting<sup>12</sup>, Oliver E. Craig<sup>10</sup>, Alison Sheridan<sup>13</sup>, Mike Parker Pearson<sup>14</sup>, Chris Stringer<sup>1</sup>, David Reich<sup>4,5,6,16</sup>, Mark G. Thomas<sup>2,3,16\*</sup> and Ian Barnes<sup>1,16\*</sup>

**The roles of migration, admixture and acculturation in the European transition to farming have been debated for over 100 years. Genome-wide ancient DNA studies indicate predominantly Aegean ancestry for continental Neolithic farmers, but also variable admixture with local Mesolithic hunter-gatherers. Neolithic cultures first appear in Britain circa 4000 bc, a millennium after they appeared in adjacent areas of continental Europe. The pattern and process of this delayed British Neolithic transition remain unclear. We assembled genome-wide data from 6 Mesolithic and 67 Neolithic individuals found in Britain, dating 8500–2500 bc. Our analyses reveal persistent genetic affinities between Mesolithic British and Western European hunter-gatherers. We find overwhelming support for agriculture being introduced to Britain by incoming continental farmers, with small, geographically structured levels of hunter-gatherer ancestry. Unlike other European Neolithic populations, we detect no resurgence of hunter-gatherer ancestry at any time during the Neolithic in Britain. Genetic affinities with Iberian Neolithic individuals indicate that British Neolithic people were mostly descended from Aegean farmers who followed the Mediterranean route of dispersal. We also infer considerable variation in pigmentation levels in Europe by circa 6000 bc.**

The transition to farming marks one of the most important ecological shifts in human evolution. The processes by which this transition occurred have been a matter of intense debate for over a century<sup>1–3</sup>, although across continental Europe ancient DNA studies indicate a predominant role for expanding Neolithic farmer populations of mostly Aegean ancestry (Aegean Neolithic Farmers (ANF))<sup>4–15</sup>. ANF-derived populations dispersed throughout Europe via two major routes: one along the Mediterranean and the other through Central and into Northern Europe<sup>7,11</sup>. Both dispersing populations introgressed repeatedly with local Mesolithic foragers, which gradually increased their proportion of European Mesolithic ancestry<sup>7,13–15</sup>.

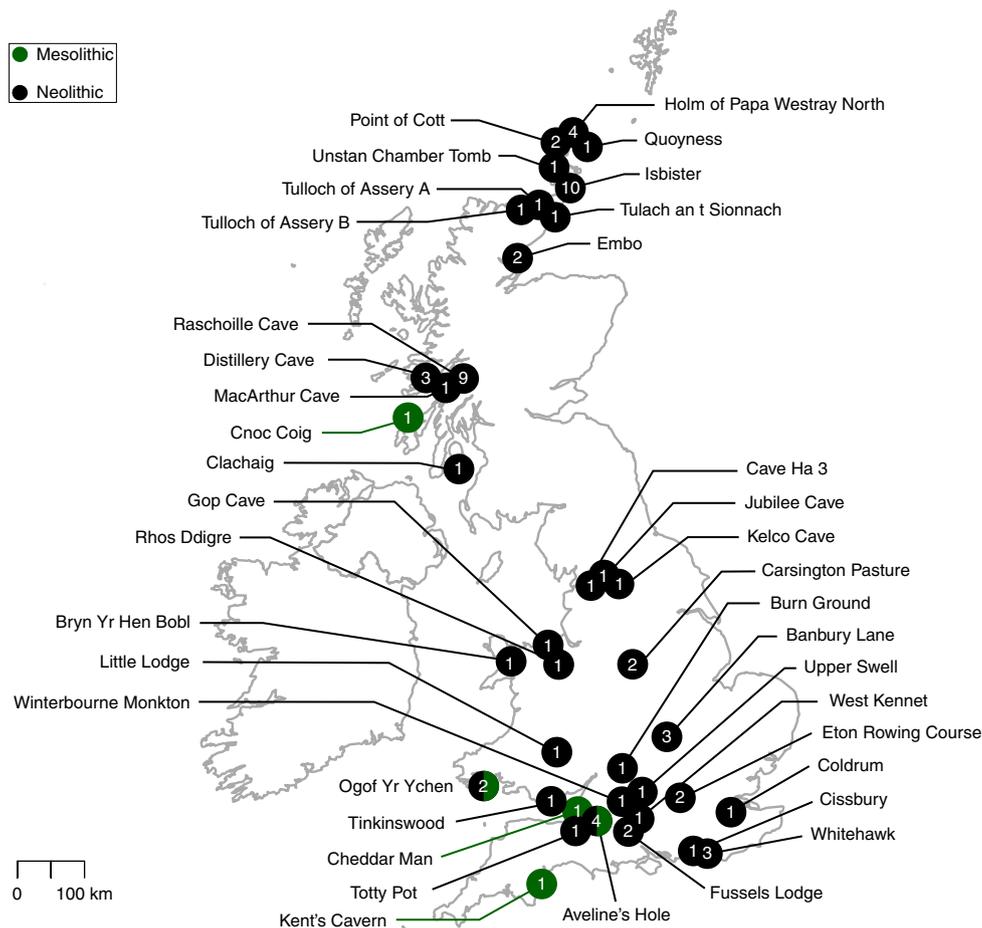
The nature of the Neolithic transition in Britain remains unclear because of the millennium-long delay in its appearance after the establishment of farming in adjacent regions of continental Europe<sup>1–3</sup>, and the lack of genome-wide data from British Mesolithic hunter-gatherers. Although there is universal agreement among archaeologists that there was a dramatic change in material culture in Britain around 4000 bc, there are divergent views regarding the extent to which this change was influenced by cultural or demographic processes<sup>1–3</sup>. The British Isles lie furthest from the Aegean origin<sup>4–15</sup> of the migrating farmers that influenced the

development of the Neolithic across Europe, are geographically isolated from continental Europe by large bodies of water and had maritime climates which differ from the majority of mainland Europe—all factors that may have altered the nature of the adoption of farming. The relationship between British and continental European Mesolithic populations is also of interest, as Britain geographically abuts two genetically distinct but contemporaneous populations, Western European and Scandinavian Mesolithic hunter-gatherers (WHGs and SHGs, respectively), and could have potentially harboured ancestry from earlier (~19000–15000 bc) Magdalenian Palaeolithic hunter-gatherer populations<sup>16–18</sup>.

## Results

Here, we report whole-genome data from 6 Mesolithic (including ‘Cheddar Man’ from Gough’s Cave, Somerset, England) and 16 Neolithic British individuals, and combine these with data from 51 previously published Neolithic British individuals<sup>12</sup> to characterize the Mesolithic to Neolithic transition in Britain (Fig. 1 and Supplementary Fig. 16). Our Mesolithic samples date from 8750–8459 calibrated (cal.) bc (Early Mesolithic Aveline’s Hole, Somerset, England) to 4256–3803 cal. bc (Late Mesolithic Cnoc Coig, Oronsay, western Scotland). Our Neolithic samples date

<sup>1</sup>Department of Earth Sciences, Natural History Museum, London, UK. <sup>2</sup>Research Department of Genetics, Evolution and Environment, University College London, London, UK. <sup>3</sup>UCL Genetics Institute, University College London, London, UK. <sup>4</sup>Department of Genetics, Harvard Medical School, Boston, MA, USA. <sup>5</sup>Broad Institute of MIT and Harvard, Cambridge, MA, USA. <sup>6</sup>Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA. <sup>7</sup>Department of Genetics, University of Cambridge, Cambridge, UK. <sup>8</sup>Department of Biology, Indiana University-Purdue University Indianapolis, Indianapolis, IN, USA. <sup>9</sup>Department of Genetic Identification, Erasmus University Medical Centre Rotterdam, Rotterdam, the Netherlands. <sup>10</sup>Bioarch, University of York, York, UK. <sup>11</sup>School of Archaeological and Forensic Sciences, University of Bradford, Bradford, UK. <sup>12</sup>Institute of Archaeology, University of Oxford, Oxford, UK. <sup>13</sup>National Museums Scotland, Edinburgh, UK. <sup>14</sup>Institute of Archaeology, University College London, London, UK. <sup>15</sup>These authors contributed equally: Selina Brace, Yoan Diekmann, Thomas J. Booth. <sup>16</sup>These authors jointly supervised this work: David Reich, Mark G. Thomas, Ian Barnes. \*e-mail: m.thomas@ucl.ac.uk; I.Barnes@nhm.ac.uk



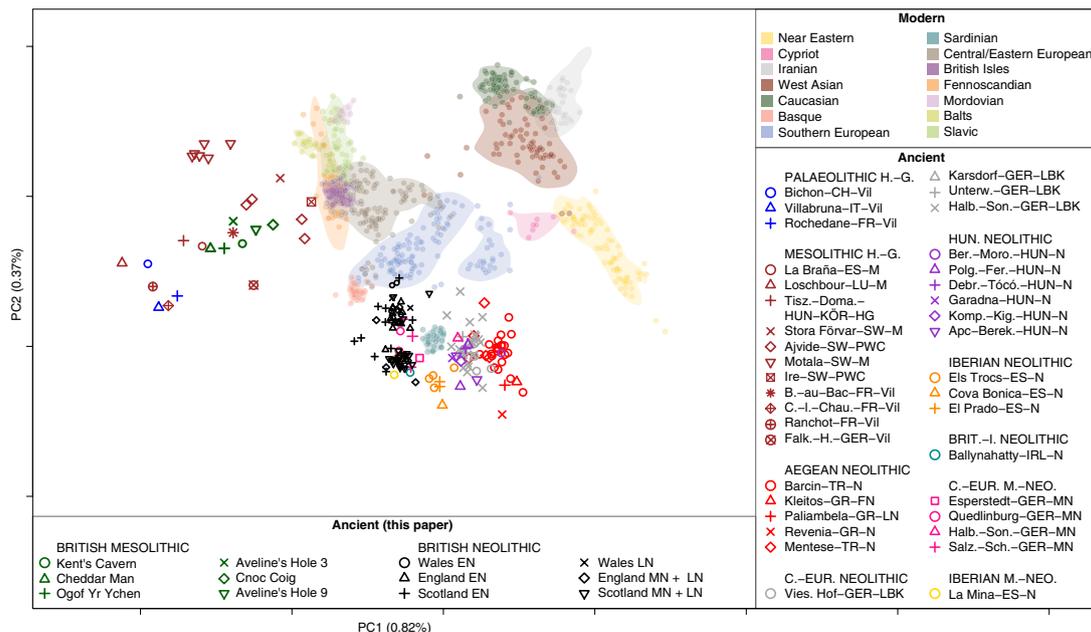
**Fig. 1 | Map of sample locations.** Geographical locations of British samples analysed here. Numbers indicate total samples obtained from a given location.

from 3951–3780 cal. BC (Early Neolithic McArthur Cave, western Scotland) to 2570–2347 cal. BC (Late Neolithic Isbister, Orkney, Scotland). We combined data generated in two different ways. For 35 individuals, we generated new whole-genome shotgun sequencing data, including full genomes from British Mesolithic (at  $\times 2.3$ ) and Neolithic (at  $\times 10.7$ ) individuals. For all samples we enriched next-generation sequencing libraries for approximately 1.24 million single nucleotide polymorphisms (SNPs) (median coverage  $\times 0.88$ ). When available, we merged data obtained from both methods and identified the most likely allele at each locus (see Methods). These were combined with ancient genomic data from 67 previously reported individuals<sup>4–7,9–12,14,16–22</sup> (see Supplementary Table 1) and modern genomic data from diverse global populations<sup>23</sup>.

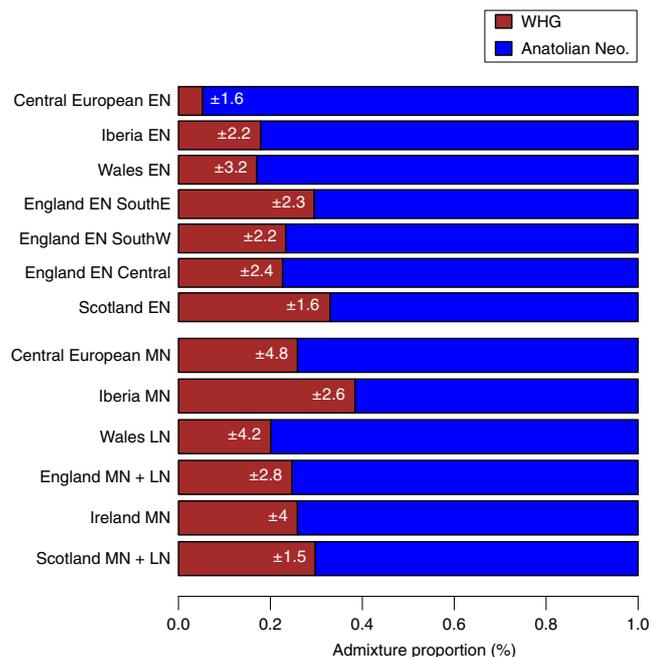
All British Mesolithic individuals cluster with Western and Scandinavian hunter-gatherers in a principal components analysis (Fig. 2). By contrast, all directly dated individuals who post-date 4000 BC and undated individuals associated with Neolithic monuments cluster tightly near Iberian and Central European Middle Neolithic individuals. By examining the degree of allele sharing of British Mesolithic individuals with various European hunter-gatherer individuals/groups (SHG, Eastern Hunter Gatherers (EHG) and El Mirón, see Supplementary Figs. 1–4), we were able to attribute these confidently to the WHG group. Comparison of British Mesolithic individuals to different Mesolithic WHGs (Berry au Bac, France; Ranchot88, France; Loschbour, Luxembourg; La Braña, Spain; KO1, Hungary; Supplementary Figs. 5, 6 and 11–14) indicates that all most closely resemble Loschbour. When we compared the remaining British Mesolithic genomes to Loschbour and Cheddar Man

(our highest-coverage British Mesolithic sample, approximately  $\times 2.3$ ), we found no major excess of shared drift for either individual, indicating that Loschbour, Ranchot88 and the British Mesolithic samples do not form separate clusters (Supplementary Fig. 7).

To investigate the proportions of Aegean farmer-related ancestry in the British samples, we modelled these as mixtures of ANFs and European WHGs using the qpAdm method, which studies ensembles of  $f_4$  statistics (Fig. 3 and Supplementary Fig. 8)<sup>24</sup>. The genomes of all British Mesolithic individuals can be explained almost entirely by WHG ancestry, the remainder ( $< 7.3\%$ ) probably stemming from poorly matching portions of the genome. Most of the ancestry in all British Neolithic individuals could be attributed to ANFs ( $> 56\%$ ,  $\sim 74\%$  on average), indicating a substantial shift in ancestry with the transition to farming. To investigate the proximate source of ANF ancestry in British Neolithic individuals, we examined affinities with Early Neolithic individuals from Iberia and Central Europe. We compare Early to Middle Neolithic individuals as the latter are contemporary with the British Early Neolithic, making these an unlikely direct source. For all British Neolithic individuals considered we inferred more shared drift with Early Neolithic Iberians (Fig. 4a and Supplementary Fig. 9). However, these  $f_4$  statistic-based inferences may be sensitive to levels of WHG admixture, such that the similarity in WHG admixture proportions in Early Neolithic Iberian and British samples, but lower estimates in Central European Early Neolithic individuals, is driving the inference of an Iberian rather than Central European source for Early British farmers. To examine this possibility in more detail, we performed a more powerful haplotype-based analysis.



**Fig. 2 | Principal component analysis of modern and ancient West Eurasians.** British and additional ancient samples are projected onto the reference space computed on present-day West Eurasian populations. See Methods for computational details and Supplementary Table 1 for information on the samples. PC, principal component. Site Abbreviations: Tisz. Doma., Tiszaszőlös-Domaháza; B.-au-Bac, Berry-au-Bac; C.-l.-Chau., Cuiry-les-Chaudardes; Falk., Falkensteiner Höhle; Vies. Hof, Viesenhauser Hof; Unterw., Unterwechel; Halb.-Son., Halberstadt-Sonntagsfeld; Ber.-Moro., Berettyóújfalu-Morotva-liget; Polg.-Fer., Polgár-Ferenci-hát; Debr.-Tócó., Debrecen Tócsópart Erdoalja; Komp.-Kig., Kompolc-Kigyóser; Apc-Berek., Apc-Berekalja I.; Salz.-Sch., Salzmuende-Schiebzig. Region: Brit-I, British Isles; C.-Eur., Central European; CH, Switzerland; ES, Spain; FR, France; GER, Germany; GR, Greece; HUN, Hungary; IRL, Ireland; IT, Italy; LU, Luxembourg; SW, Sweden; TR, Turkey. Plei., Pleistocene; H.-G., hunter-gatherer; Vil., Villabruna Cluster. Cultural Horizon: KÖR, Körös; EN, MN and LN, Early, Middle and Late Neolithic, respectively; M, Mesolithic; N, Neolithic; PWC, Pitted Ware Culture; LBK, Linearbandkeramik.

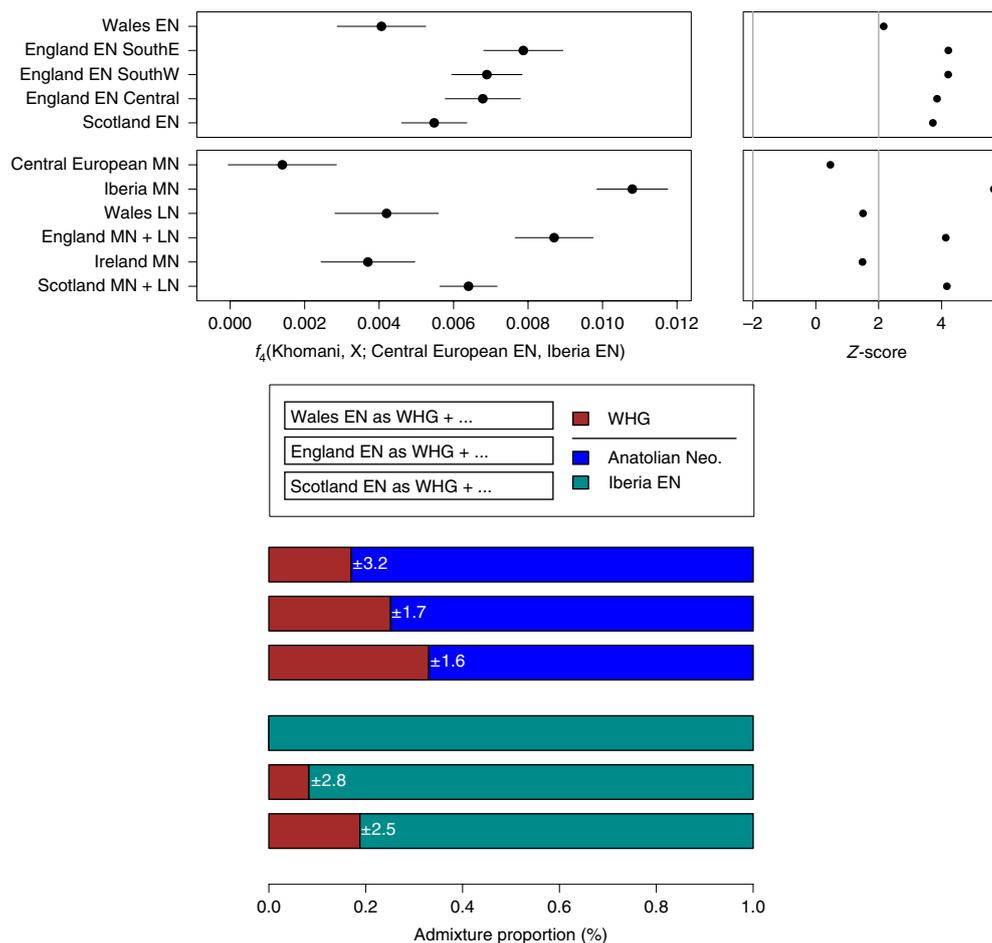


**Fig. 3 | WHG and ANF ancestry components of British and Central European Neolithic populations.** Relative WHG and ANF ancestry in Early and Middle Neolithic British and continental European populations quantified by qpAdm. Percentages indicate error estimates computed by block jack-knifing with a block size of 5 centimorgans (cM) (ref. 24). See Methods for computational details and Supplementary Table 1 for lists of samples grouped into WHG and the different Neolithic populations. Neo., Neolithic; SouthE, South-East; SouthW, South-West.

Using a chromosome-painting approach<sup>25</sup>, we obtained patterns of haplotype matching between our high-coverage British Neolithic sample and a global modern reference panel (Supplementary Data 7). We found similar patterns of donor haplotype matching in the British Neolithic genome to those inferred for other high-coverage Neolithic genomes from Ireland and Iberia. These were more similar than the same profiles obtained for high-coverage Neolithic genomes from Central Europe (Fig. 5a). Inferred ancestry coefficients (see Methods) further support this connection between the British, Irish and Iberian Neolithic<sup>6</sup> and are consistent with the same ancestral populations bringing the Neolithic to Britain and Ireland (Fig. 5b,c and Supplementary Table 8). Additional modelling using global modern populations<sup>26</sup> as ancestry surrogates suggests that this population is best represented today by components found in French and Spanish peoples (Fig. 5c and Supplementary Table 9).

To test for a potential second ANF ancestry stream from Central Europe, we explicitly modelled WHG and Early Neolithic populations in qpGraph (see Supplementary Fig. 23 and Supplementary Table 10). The results suggest that the limited Central European Neolithic admixture we find in British Neolithic populations is regionally structured, with populations from England showing the highest levels of admixture followed by those from Scotland. We infer no Central European admixture in Neolithic farmers from Wales. However, we caution that the model fits are poor and so these inferences should be considered preliminary.

We inferred some notable geographic structure in WHG admixture proportions among the British Early Neolithic individuals (see Supplementary Table 4 for statistical comparison of inferred WHG admixture proportions); those from Wales retain the lowest levels of WHG admixture, followed by those from South-West and Central England. Neolithic individuals from



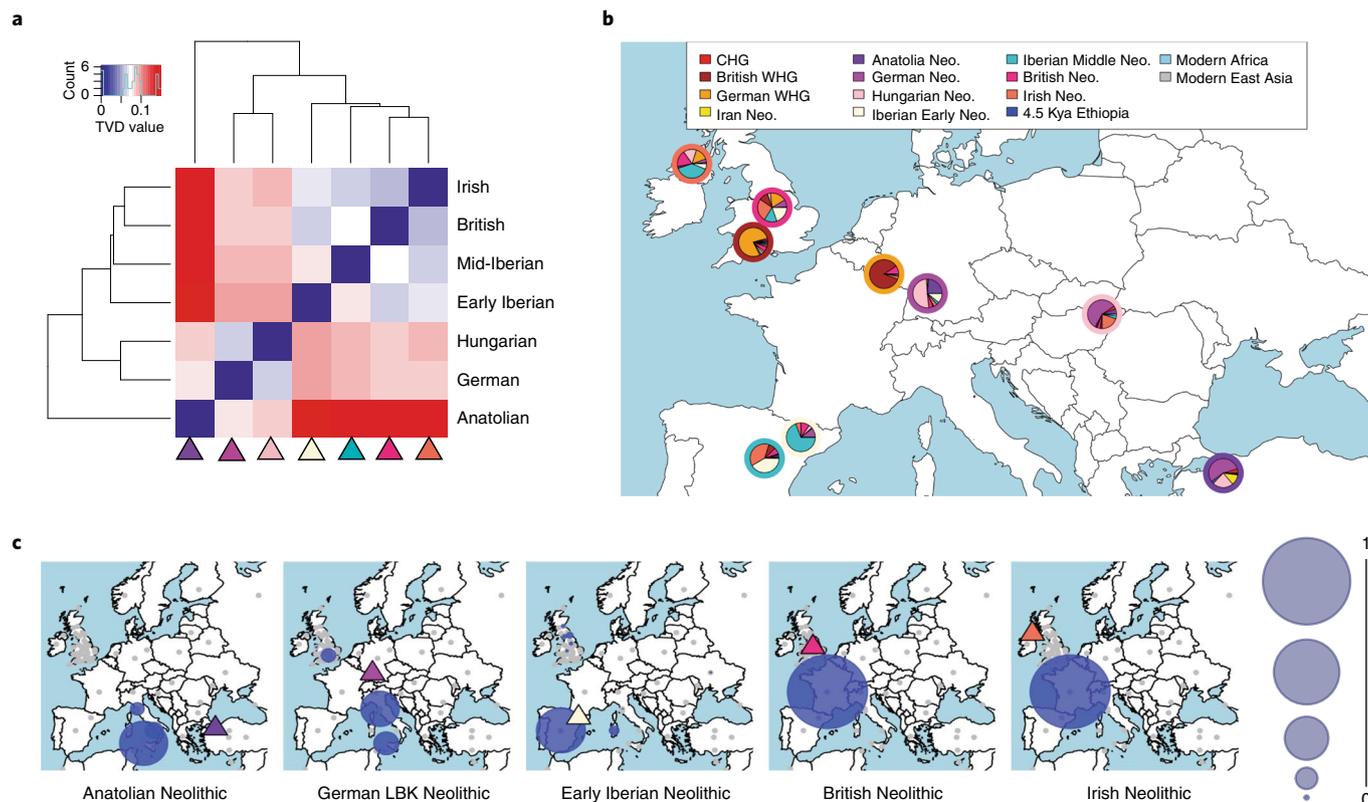
**Fig. 4 | Affinities of British and continental Neolithic populations.** Top, We computed  $f_4$ -statistics of the form  $f_4$  (Khomani, test; Central European EN, Iberia EN) for different British EN, MN and LN and continental MN populations to compare shared drift to Central European EN and Iberian EN populations. A positive Z-score  $> 2$  corresponds to a significant affinity to the Iberian EN over Central European EN population. Bottom, Quantification of excess WHG ancestry in British EN compared to the Iberian EN population. We computed qpAdm estimates of WHG and Anatolian and Iberian ANF populations in EN samples from Wales, England and Scotland. See Supplementary Table 1 for lists of samples grouped into WHG and the different Neolithic populations. The three white bars and the coloured boxes indicate how the bars below them are derived. Percentages and bars indicate error estimates computed by block jack-knifing with a block size of 5 cM (ref. <sup>24</sup>).

South-East England and Scotland show considerably higher WHG admixture proportions. These proportions remain stable from the Early into the Middle/Late Neolithic. To infer levels of WHG introgression occurring between Iberian Early Neolithic populations and early British farmers, we estimated admixture proportions using qpAdm<sup>24</sup>. We detected little excess (~10%) WHG ancestry beyond that already present in Iberian Early Neolithic individuals, supporting little or no additional admixture with British hunter-gatherers, particularly in Wales and South-West and Central England (Figs. 3, 4b and Supplementary Table 4). This result appears to be slightly at odds with the  $f_4$  results presented in Supplementary Fig. 7, which indicate that some British Neolithic samples share genetic affinities with Cheddar Man over Loschbour, although it is difficult to say in these cases whether this is due to genuine substantial admixture with British WHGs or with other WHGs in northern Europe. We regressed individual WHG ancestry proportions in British Neolithic farmers (shown in Supplementary Fig. 8) against latitude and longitude and found a notably positive south-west to north-east cline (Supplementary Fig. 15).

To further explore WHG introgression in Britain we applied ALDER<sup>27</sup> to pairs of Early Neolithic regional samples to estimate the timing of WHG/ANF admixture events (Supplementary Table 3).

Only Early Neolithic farmers from western Scotland show evidence of WHG introgression within ten generations. Two individuals from Raschoille Cave had estimated introgression events occurring  $4.0 \pm 3.4$  generations before they lived, which is sufficiently recent in their past that it probably occurred in Britain. The elevated levels of WHG ancestry we see in Neolithic samples from South-East England are older, and therefore probably a result of farmer–forager interactions in mainland Europe. Chronological modelling (using OxCal 4.3 (ref. <sup>28</sup>)) of available direct Early Neolithic radiocarbon data from individuals showing ANF ancestry suggests that continental farmers arrived in Britain by 3975–3722 cal. BC (95% confidence interval), 481 years after to 27 years before (95% confidence interval) the death of our latest Mesolithic individual showing no ANF ancestry (Supplementary Data 6). Our model suggests that continental farmers arrived marginally earlier in the west (although see the discussion in Supplementary Data), and rapidly dispersed to other regions of Britain (including the Orkney Islands). The latest regional appearance of ANF ancestry is in Central England and occurs 59 to 386 years (95% confidence interval) after it first appears in Britain.

To explore variation in pigmentation of European populations, we predicted pigmentation in higher-coverage Mesolithic and Neolithic Europeans using HIrisplex-S<sup>29</sup>. We infer that Cheddar Man mostly probably had blue/green eyes, dark brown (possibly



**Fig. 5 | Patterns of haplotype sharing across high-coverage aDNA samples. a**, Hierarchical clustering of total variation distance (TVD) between CHROMOPAINTER inferred haplotype-sharing profiles of seven high-coverage Neolithic individuals when compared to a global modern reference panel. **b**, Inferred ancestry proportions (SOURCEFIND inferred mixing coefficients) of high-coverage ancient genomes, coloured as per the legend and outer pie ring colour, relative to a panel of ancient genomes, plus modern Yoruba and Han (as given in the legend at the top). Raw proportions and standard errors are provided in Supplementary Table 8. CHG, Caucasus hunter-gatherer. **c**, Inferred ancestry proportions of five high-coverage Neolithic individuals (triangles coloured as in **b**) relative to a global modern reference panel. The size of the blue circle provides the majority of inferred contributions, as given by the scale at bottom right, with all possible modern contributors denoted by a black dot. Raw proportions and standard errors are provided in Supplementary Table 9.

black) hair and dark or dark to black skin, whereas our highest-coverage Early Neolithic individual had brown eyes, black (possibly dark brown) hair and dark to intermediate skin (see Supplementary Data 3). Together with the pigmentation prediction outcomes we obtained for Loschbour and La Braña, these results imply that different pigmentation levels co-existed in Europe by around 6000 BC.

## Discussion

The six British Mesolithic genomes examined here are typical of WHGs, indicating that this population spread to the furthest north-western point of Early Holocene Europe after moving from south-eastern Europe, or further east, from approximately 12000 BC<sup>17</sup>. This genetic similarity among British and European Mesolithic individuals spans a period in Britain (circa 8500–4000 BC) that includes the cultural transition to the Late Mesolithic and the separation of Britain from continental Europe. Our analyses indicate that the appearance of Neolithic practices and domesticates in Britain circa 4000 BC was mediated overwhelmingly by immigration of farmers from continental Europe<sup>1,2</sup>, and strongly reject the hypothesized adoption of farming by indigenous hunter-gatherers as the main process<sup>3</sup>. British farmers were substantially descended from Iberian Neolithic-related populations whose ancestors had expanded along a Mediterranean route<sup>6,11</sup>, although with a minority portion of their ancestry from populations who took the Danubian route<sup>12</sup>. The affinities we find between Neolithic individuals from the British Isles and modern individuals from France are consistent with populations sharing ancestry with Neolithic groups in Iberia moving into

northern France via the Atlantic seaboard and/or southern France, mixing to a limited degree with Neolithic populations from Central Europe before travelling across the English Channel<sup>1,2,30</sup>.

One explanation for the British Neolithic cline in WHG ancestry is that a single population moved across Britain from a western entry point and progressively admixed with local hunter-gatherers. This scenario is consistent with the western distribution of megalithic cultures along the Atlantic seaboard<sup>31</sup>, and is supported by radiocarbon evidence suggesting a marginally earlier date for the arrival of ANF ancestry in the west of Britain<sup>1</sup>. However, the lack of evidence for substantive WHG introgression into British Neolithic populations—outside of western Scotland—favours this cline, reflecting multiple source populations with variable proportions of WHG admixture having entered different parts of Britain. This interpretation is consistent with archaeological evidence for regional British Neolithic cultures showing links to varied parts of mainland Europe<sup>2</sup> and our qpGraph analysis indicating geographically structured Neolithic Central European admixture. Overall, the regional variation in ancestry of British Neolithic populations probably reflects both differing degrees of admixture between farmers and local foragers (for example, western Scotland) and multiple continental source populations carrying variable WHG and Neolithic Central European ancestry.

Evidence for only low levels of WHG introgression among British Early Neolithic people is striking given the extensive and complex admixture processes inferred for continental Neolithic populations<sup>7,13–15,32,33</sup>. Low levels of admixture between these

two groups on the wave front of farming advance in continental Europe have been attributed to the maintenance of cultural and reproductive boundaries for up to two millennia after initial contact, before more extensively mixing<sup>32</sup>. Similarly, isotopic and genetic data from the west coast of Scotland suggest the potential co-existence of genetically distinct hunter-fisher-gatherers and farmers, albeit for a maximum of a few centuries<sup>34</sup>. However, there is no evidence for a resurgence of WHG ancestry in the British Neolithic, consistent with limited evidence for Mesolithic cultural artefacts in Britain beyond 4000 BC<sup>1,2</sup>, and with a major dietary shift from marine to terrestrial resources at this time (see Supplementary Data 5)<sup>35</sup>.

## Conclusion

In contrast to other European regions, the transition to farming in Britain occurred with little introgression from resident foragers—either during initial colonization or throughout the Neolithic. This may reflect low Late Mesolithic population density in Britain and/or an introduction of farming by populations who had mastered the technologies needed to thrive in northern and western continental Europe during the previous two millennia<sup>1,2</sup>.

## Methods

**Ancient DNA extraction and sequencing.** The DNA extractions and library preparations for all samples with newly reported data were conducted in a dedicated ancient DNA laboratory (NHM, London). We used approximately 25 mg of finely drilled bone powder and followed the DNA extraction protocol described in ref. <sup>36</sup>, but replaced the Zymo-Spin V column binding apparatus with a high pure extender assembly from the High Pure Viral Nucleic Acid Large Volume Kit (Roche). Library preparations followed the partial uracil–DNA–glycosylase treatment described in ref. <sup>37</sup> and a modified version of the protocol described in ref. <sup>38</sup>. Library modifications: the initial DNA fragmentation step was not required, and all clean-up steps used MinElute PCR purification kits (Qiagen). The index PCR step included double indexing<sup>39</sup>, the polymerase AmpliTaq Gold and the addition of 0.4 mg ml<sup>-1</sup> bovine serum albumin. The index PCR was set for 20 cycles with three PCR reactions conducted per library. Libraries were screened for DNA preservation on an Illumina NextSeq platform, with paired-ends reads. Promising libraries were further enriched at the NHM using in-solution hybridization capture enrichment kits (Mybaits-3) from MYcroarray. The baits were designed to cover around 20,000 SNPs (5,139 functional and 15,002 neutral SNPs) at ×4 tiling. The capture protocol followed the manufacturer's instructions in Mybaits manual v.3. Captured libraries were sequenced on an Illumina NextSeq platform (NHM) using paired-ends reads. Newly reported data from 36 of these libraries were also obtained at the dedicated ancient DNA laboratory in Harvard Medical School by enrichment in solution for approximately 1.24 million targeted SNPs. We sequenced these libraries on an Illumina NextSeq500 instrument, iteratively sequencing more until we estimated that the additional number of targeted SNPs hit per newly generated sequence was fewer than one per 100.

**Bioinformatics.** All sequence reads underwent adapter and low-quality base trimming, and overlapping reads pairs were collapsed with AdapterRemoval<sup>40</sup>. Non-collapsed reads and those of length less than 30 base pairs were discarded, and the remaining aligned against the hs37d5 human reference genome with the Burrows–Wheeler Aligner<sup>41</sup>. Mapped reads with a mapping quality of at least 30 were merged per individual and realigned around InDels with GATK<sup>42</sup>. Resulting BAM files were split by flowcell and lane, and empirical ATLAS<sup>43</sup> post-mortem damage patterns estimated per individual per lane for lanes with at least 5.5 million reads, otherwise per individual per flowcell. ATLAS BQSR (base quality score recalibration) tables were generated per lane for lanes with at least 5.5 million reads, otherwise per flowcell. We generated recalibrated BAM files per individual with ATLAS recalBAM, and used these to estimate mitochondrial contamination and determine mitochondrial and Y-chromosome haplogroups with ContamMix<sup>44</sup>, Yleaf<sup>45</sup> and Phy-Mer<sup>46</sup>. We considered mitochondrial contamination to be tolerable if 0.98 was included in the confidence intervals. Haploid genotypes were called with ATLAS 'allelePresence' with theta fixed at 0.001, determining the most likely base at a given position. Heterozygosity estimates, shown in Supplementary Fig. 10, were computed with ATLAS 'estimate Theta' and a default window size of 1 megabase pair, excluding windows that overlap with telo- or centromeres.

**Principal component analysis.** Principal component analysis was performed with LASER<sup>47</sup> following the approach described previously<sup>9</sup>. After generating a reference space of modern West Eurasian individuals<sup>10</sup>, we projected the BAM files of ancient reference individuals (see Supplementary Table 1 for references) and the British individuals presented here into the reference space via Procrustes analysis implemented in LASER.

**f-statistics.** The *f*-statistics presented here, that is outgroups  $f_3, f_4$ , qpAdm and qpGraph, were computed with qpPop and qpDstat in  $f_4$  mode, and qpAdm and qpGraph from the ADMIXTOOLS<sup>48</sup> package with default parameters on the positions defined by the *H0111* set of SNPs<sup>7</sup>. Ancient individuals analysed here are listed in Supplementary Table 1, including the explanation of all population labels used (WHG, SHG and so on). Modern reference individuals were first published in ref. <sup>23</sup>. All qpAdm runs used the set of outgroups Han, Karitiana, Mbuti, Onge, Papuan, Mota, Ust'-Ishim, MA1, El Mirón, GoyetQ116-1.

**ALDER.** We used ALDER<sup>27</sup> to estimate the dates of admixture between WHG and ANF. All combinations we tested are listed in Supplementary Table 3, which consisted of the pairs or groups of individuals specified in the first column and WHG and ANF (individuals constituting WHG and ANF are given in Supplementary Table 1).

**Chronological modelling.** We used OxCal 4.3 (ref. <sup>28</sup>) to produce chronological models of the arrival and spread of ANF ancestry into Britain (Supplementary Data 6). We used Early Neolithic (4000–3500 BC) radiocarbon dates only from those obtained from material or individuals where there were palaeogenetic data indicating ANF ancestry. We divided these samples into five regional populations: Western Britain, Central England, Eastern England, Western Scotland and the Orkney Isles. Dates associated with each region were grouped as Phases (Supplementary Fig. 19). We used the Boundary function to produce probability distributions for the arrival of ANF ancestry in Britain as a whole, and for each region. We used the Difference function to produce probability distributions for the time between the death of the latest individual with wholly WHG ancestry and the arrival of populations with ANF ancestry, as well as between the arrival of ANF ancestry in Britain as a whole and the different regions of Britain.

**Haplotype-based analyses.** We used CHROMOPAINTER<sup>25</sup> to summarize DNA patterns in our ancient individuals, including other high-coverage, publicly available ancient genomes from relevant cultures and time periods, to infer the proportion of DNA for which ancient individuals most closely matched to those from a global panel of modern donor groups<sup>48–50</sup>. This panel included many population samples from across West Eurasia, as well as 35 labelled groups from within the British Isles. We generated matching profiles when considering SNPs independently (allele sharing) and also when considering the correlations between neighbouring SNPs (haplotype sharing). To do so we first merged high-quality diploid calls for our selected high-coverage ancient genomes and jointly phased the resultant dataset of 159,287 SNPs using SHAPEIT v.2 (ref. <sup>51</sup>). We performed additional mixture modelling on our generated allele- and haplotype-sharing profiles implemented in SOURCEFIND<sup>26</sup> to form target groups as mixtures of the DNA-sharing profiles of other included groups. We performed two sets of analyses: (1) using all modern groups (or a subset of) to model the ancestry of ancient individuals and (2) using different sets of ancient individuals, plus the modern Yoruba and Han, to model the ancestry of modern worldwide groups. Further details are provided in Supplementary Data 7.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

BAM files (one file per library, before realigning around InDels; see Supplementary Table 1) have been deposited at the European Nucleotide Archive under study accession PRJEB31249.

Received: 17 December 2018; Accepted: 6 March 2019;

Published online: 15 April 2019

## References

- Collard, M., Edinborough, K., Shennan, S. & Thomas, M. G. Radiocarbon evidence indicates that migrants introduced farming to Britain. *J. Arch. Sci.* **37**, 866–870 (2010).
- Sheridan, J. A. in *Landscapes in Transition* (eds Finlayson, B. & Warren, G.) 89–105 (Oxbow, 2010).
- Thomas, J. *The Birth of Neolithic Britain: an Interpretive Account* (Oxford University Press: 2013).
- Skoglund, P. et al. Genomic diversity and admixture differs for Stone Age Scandinavian foragers and farmers. *Science* **344**, 747–750 (2014).
- Gamba, C. et al. Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 5257 (2014).
- Cassidy, L. M. et al. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proc. Natl Acad. Sci. USA* **113**, 368–373 (2015).
- Haak, W. et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
- Broushaki, F. et al. Early Neolithic genomes from the eastern Fertile Crescent. *Science* **353**, 499–503 (2016).

9. Hofmanová, Z. et al. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc. Natl Acad. Sci. USA* **113**, 6886–6891 (2016).
10. Lazaridis, I. et al. Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).
11. Olalde, I. et al. A common genetic origin for early farmers from Mediterranean cardiac and central European LBK cultures. *Mol. Biol. Evol.* **32**, 3132–3142 (2015).
12. Olalde, I. et al. The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* **555**, 190–196 (2018).
13. González-Fortes, G. et al. Paleogenomic evidence for multi-generational mixing between Neolithic farmers and mesolithic hunter-gatherers in the lower Danube basin. *Curr. Biol.* **27**, 1801–1810 (2017).
14. Lipson, M. et al. Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372 (2017).
15. Mathieson, I. et al. The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).
16. Günther, T. et al. Genomics of Mesolithic Scandinavia reveal colonization routes and high-latitude adaptation. *PLoS Biol.* **16**, e2003703 (2018).
17. Fu, Q. et al. The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
18. Olalde, I. et al. Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225–228 (2014).
19. Mathieson, I. et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
20. Kılınç, G. M. et al. The demographic development of the first farmers in Anatolia. *Curr. Biol.* **26**, 2659–2666 (2016).
21. Jones, E. R. et al. The Neolithic transition in the Baltic was not driven by admixture with early European farmers. *Curr. Biol.* **27**, 576–582 (2017).
22. Lazaridis, I. et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
23. Mallick, S. et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201–206 (2016).
24. Patterson, N. et al. Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
25. Lawson, D. J. et al. Inference of population structure using dense haplotype data. *PLoS Genet.* **8**, e1002453 (2012).
26. Chacon-Duque, J. C. et al. Latin Americans show wide-spread Converso ancestry and the imprint of local Native ancestry on physical appearance. *Nat. Commun.* **9**, 5388 (2018).
27. Loh, P.-R. et al. Inferring admixture histories of human populations using linkage disequilibrium. *Genetics* **193**, 1233–1254 (2013).
28. Bronk Ramsey, C. Bayesian analysis of radiocarbon dates. *Radiocarbon* **51**, 337–360 (2009).
29. Chaitanya, L. et al. The HIRISplex-S system for eye, hair and skin colour prediction from DNA: introduction and forensic developmental validation. *Forensic Sci. Int. Genet.* **35**, 123–135 (2018).
30. Whittle, A. W. R., Healy, F., Bayliss, A. & Allen, M. J. *Gathering Time: Dating the Early Neolithic Enclosures of Southern Britain and Ireland*. (Oxbow Books, 2011).
31. Scarre, C. The early Neolithic of western France and Megalithic origins in Atlantic Europe. *Oxford J. Archaeol.* **11**, 121–154 (1992).
32. Bollongino, R. et al. 2000 years of parallel societies in Stone Age Central Europe. *Science* **342**, 479–481 (2013).
33. Fraser, M. et al. New insights on cultural dualism and population structure in the Middle Neolithic Funnel Beaker culture on the island of Gotland. *Sci. Rep.* **17**, 325–334 (2018).
34. Charlton, S. et al. Finding Britain's last hunter-gatherers: a new biomolecular approach to 'unidentifiable' bone fragments utilising bone collagen. *J. Archaeol. Sci.* **73**, 55–61 (2016).
35. Schulting, R. J. and Borčić, D. in *Neolithic Europe: Essays in Honour of Professor Alasdair Whittle* (eds P. Bickle, V. Cummings, D. Hofmann & J. Pollard) 82–104 (Oxford, 2017).
36. Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–15763 (2013).
37. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. Lond. B* **370**, 20130624–20130624 (2014).
38. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* <https://doi.org/10.1101/pdb.prot5448> (2010).
39. Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* **40**, e3 (2011).
40. Schubert, M., Lindgreen, S. & Orlando, L. Adapter Removal v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88 (2016).
41. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
42. Mckenna, A. et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
43. Link, V. et al. ATLAS: analysis tools for low-depth and ancient samples. Preprint at *Biorxiv* <https://doi.org/10.1101/105346> (2017).
44. Fu, Q. et al. A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23**, 553–559 (2013).
45. Navarro-Gomez, D. et al. Phy-Mer: a novel alignment-free and reference-independent mitochondrial haplogroup classifier. *Bioinformatics* **31**, 1310–1312 (2014).
46. Ralf, A., Montiel González, D., Zhong, K. & Kayser, M. Yleaf: software for human Y-chromosomal haplogroup inference from next generation sequencing data. *Mol. Biol. Evol.* **35**, 1291–1294 (2018).
47. Wang, C., Zhan, X., Liang, L., Abecasis, G. R. & Lin, X. Improved ancestry estimation for both genotyping and sequencing data using projection procrustes analysis and genotype imputation. *Am. J. Hum. Genet.* **96**, 926–937 (2015).
48. Hellenthal, G. et al. A genetic atlas of human admixture history. *Science* **343**, 747–751 (2014).
49. Busby, G. B. et al. The role of recent admixture in forming the contemporary West Eurasian genomic landscape. *Curr. Biol.* **25**, 2518–2526 (2015).
50. Leslie, S. et al. The fine-scale genetic structure of the British population. *Nature* **519**, 309 (2015).
51. Delaneau, O. et al. A linear complexity phasing method for thousands of genomes. *Nat. Methods* **9**, 179–181 (2012).

## Acknowledgements

The authors would like to thank the Longleat Estate, T. Lord at Lower Winskill Farm, B. Chandler at Torquay Museum, A. Chamberlain at the University of Manchester, L. Wilson and G. Mullan at the University of Bristol Speleological Society, E. Walker, A. Gwilt and J. Deacon at the National Museum of Wales, A. Maxted at Brighton Museum, M. Lahr at the Duckworth Laboratory, B. Lane at Wells Museum, M. Smith at Bournemouth University, D. Rice at the Museum of Gloucester and R. Kruszynski at the Natural History Museum for providing access to samples. In addition, Y.D. wishes to thank J. Blöcher, A. Scheu, C. Sell and J. Burger for discussions on the bioinformatic pipeline, and V. Link for help with ATLAS. M.G.T. and I.B. were supported by a Wellcome Trust Investigator Award (project No. 100713/Z/12/Z). S.C. was supported by the Natural Environment Research Council (NE/K500987/1). L.v.D. acknowledges financial support from the Newton Trust (grant No. MR/P007597/1). R.M. was supported by an EMBO Long-Term Fellowship (No. ALTF 133-2017). D.R. was supported by a NIH grant (No. GM100233), by NSF HOMINID (No. BCS-1032255) and by an Allen Discovery Center of the Paul Allen Foundation, and is a Howard Hughes Medical Institute investigator. C.S. is supported by the Calleva Foundation and the Human Origins Research Fund. S.W. was supported by the US National Institute of Justice (grant No. 2014-DN-BX-K031).

## Author contributions

I.B. and M.G.T. conceived the project. Y.D., S.B., Z.F., O.C. and T.B. contributed to the project design. S.B., Y.D., T.B., L.v.D., N.R., S.M., I.O., M.F., M.M., J.O., N.B., K.S., R.M., S.C. and S.W. generated and analysed data. I.B., M.G.T., Y.D., S.B., T.B., M.K., S.W., G.H., I.A., R.S., O.C., A.S., M.P.P., C.S. and D.R. contributed to the sampling strategy and the interpretation of results. I.B., M.G.T., Y.D., S.B. and T.B. wrote the paper, with contributions from all other authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41559-019-0871-9>.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Correspondence and requests for materials** should be addressed to M.G.T. or I.B.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- |                                     |                                     |  |
|-------------------------------------|-------------------------------------|--|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A description of all covariates tested   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated   |

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

No Software used to collect data

Data analysis

Adapter Removal v2.1.7; BWA v0.7.12; GATK v3.6; ATLAS v1.0; ContamMix v1.0-10; Yleaf v1; Phy-Mer; LASER v2.02; ADMIXTOOLS v4.1; ALDER v1.03; OxCal v4.3; CHROMOPAINTER v2; SHAPEIT v2; SOURCEFIND v1 Computer code is available from Yoan Diekmann upon request

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data availability—BAM files (a file per library, before re-aligning around InDels; see Supplementary Table 1) have been deposited at ENA under study accession PRJEB31249.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was determined by sample availability and sample quality. Data consists of 6 Mesolithic samples and 67 Neolithic samples. Genome wide data, enriched for 1.24 million SNPs for each individual were analysed.
Data exclusions	Sequencing reads: Non-collapsed reads, those below 30bp length and those with MAPQ less than 30 were discarded.
Replication	Multiple DNA extractions were carried out and multiple libraries were constructed for a subset of the samples.
Randomization	N/A
Blinding	N/A

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input type="checkbox"/>	<input checked="" type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Palaeontology

Specimen provenance	SI: Table S1.
Specimen deposition	SI: Table S1 and for sequenced data see data availability statement.
Dating methods	SI: Table S7

Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.