Large-Scale Migration into Southern Britain During the Middle to Late Bronze Age

Nick Patterson[1,2,*], Michael Isakov[3,*], Thomas Booth[4,*], Lindsey Büster[5,*], Claire-Elise Fischer[5,*], Iñigo Olalde[6,7,8], Harald Ringbauer[1,2], Ali Akbari[1,6], Olivia Cheronet[9], Madeleine Bleasdale[5], Nicole Adamski[6,10], Eveline Altena[11], Rebecca Bernardos[6], Selina Brace[12], Nasreen Broomandkhoshbacht[6,10], Kimberly Callan[6,10], Francesca Candilio[13], Brendan Culleton[14], Elizabeth Curtis[6,10], Lea Demetz[9], Kellie Sara Duffett Carlson[9], Daniel M. Fernandes[9,15], M. George B. Foody[16], Suzanne Freilich[9], Helen Goodchild[5], Aisling Kearns[6], Ann Marie Lawson[6,10], Iosif Lazaridis[1,6], Matthew Mah[2,6,10], Swapan Mallick[2,6,10], Kirsten Mandl[9], Adam Micco[2,6], Megan Michel[6,10], Guillermo Bravo Morante[9], Jonas Oppenheimer[6,10], Kadir Toykan Özdoğan[9], Lijun Qiu[6], Constanze Schattke[9], Kristin Stewardson[6,10], J. Noah Workman[6], Fatma Zalzala[6,10], Zhao Zhang[6], Bibiana Agustí[17], Tim Allen[18], Katalin Almássy[19], Luc Amkreutz[20,21], Abigail Ash[22], Christèle Baillif-Ducros[23], Alistair Barclay[24], László Bartosiewicz[25], Katherine Baxter[26], Zsolt Bernert[27], Jan Blažek[28], Mario Bodružić[29], Philippe Boissinot[30], Clive Bonsall[31], Pippa Bradley[24], Marcus Brittain[32], Alison Brookes[33], Fraser Brown[18], Lisa Brown[34], Richard Brunning[35], Chelsea Budd[36], Josip Burmaz[37], Sylvain Canet[23], Silvia Carnicero-Cáceres[38], Morana Čaušević-Bully[39], Andrew Chamberlain[40], Sébastien Chauvin[23], Sharon Clough[24], Natalija Čondić[41], Alfredo Coppa[6,9,42], Oliver Craig[5], Matija Črešnar[43,44], Vicki Cummings[45], Szabolcs Czifra[46], Alžběta Danielisová[47], Robin Daniels[48], Alex Davies[18], Philip de Jersey[49], Jody Deacon[50], Csilla Deminger[51], Peter W. Ditchfield[52], Marko Dizdar[53], Miroslav Dobeš[47], Miluše Dobisíková[54], László Domboróczki[55], Gail Drinkall[56], Ana Đukić[57], Ceiridwen J. Edwards[16], Michal Ernée[47], Christopher Evans[32], Jane Evans[58], Manuel Fernández-Götz[31], Slavica Filipović[59], Andrew Fitzpatrick[60], Harry Fokkens[21], Chris Fowler[61], Allison Fox[62], Zsolt Gallina[63], Michelle Gamble[64], Manuel R. González Morales[65], Borja González-Rabanal[66], Adrian Green[67], Katalin Gyenesei[68], Diederick Habermehl[69], Tamás Hajdu[27,70], Derek Hamilton[71], James Harris[33], Chris Hayden[18], Joep Hendriks[72], Bénédicte Hernu[73], Gill Hey[18], Milan Horňák[43], Gábor Ilon[74], Eszter Istvánovits[75,] Andy M. Jones[76], Martina Blečić Kavur[77], Kevin Kazek[78], Robert A. Kenyon[79], Amal Khreisheh[35], Viktória Kiss[80], Jos Kleijne[81], Mark Knight[32], Lisette M. Kootker[82], Péter F. Kovács[83], Anita Kozubová[84], Gabriella Kulcsár[85], Valéria Kulcsár[86], Christophe Le Pennec[87], Michael Legge[88], Matt Leivers[89], Louise Loe[18], Olalla López-Costas[90], Tom Lord[91], Dženi Los[37], James Lyall[92], Ana B. Marín-Arroyo[66], Philip Mason[44], Damir Matošević[93], Andy Maxted[94], Lauren McIntyre[18], Jacqueline McKinley[89], Kathleen McSweeney[31], Bernard Meijlink[95], Balázs G. Mende[80], Marko Menđušić[96], Milan Metlička[97], Sophie Meyer[98], Kristina Mihovilić[99], Lidija Milasinovic[100], Steve Minnitt[35], Joanna Moore[101], Geoff Morley[102], Graham Mullan[103], Margaréta Musilová[104], Benjamin Neil[32], Rebecca Nicholls[105], Mario Novak[106], Maria Pala[16], Martin Papworth[107], Cécile Paresys[23], Ricky Patten[32], Domagoj Perkić[108], Krisztina Pesti[109], Alba Petit[110], Katarína Petriščáková[111], Coline Pichon[112], Catriona Pickard[31], Zoltán Pilling[113], T. Douglas Price[114], Siniša Radović[115], Rebecca Redfern[116], Branislav Resutík[104], Daniel T. Rhodes[117], Martin B. Richards[16], Amy Roberts[118], Jean Roefstra[119], Pavel Sankot[120], Alena Šefčáková[121], Alison Sheridan[122], Sabine Skae[123], Miroslava Šmolíková[111], Krisztina Somogyi[124], Ágnes Somogyvári[125], Mark Stephens[126], Géza Szabó[127], Anna Szécsényi-Nagy[80], Tamás Szeniczey[27,70], Jonathan Tabor[32], Károly Tankó[128], Clenis Tavarez Maria[129], Rachel Terry[130], Biba Teržan[43], Maria Teschler-Nicola[9,131], Jesús F. Torres-Martínez[132], Julien Trapp[78], Ross Turle[133], Ferenc Ujvári[109], Menno van der Heiden[134], Petr Veleminsky[54], Barbara Veselka[135,136], Zdeněk Vytlačil[54], Clive Waddington[137], Paula Ware[126], Paul Wilkinson[138], Linda Wilson[103], Rob Wiseman[32], Eilidh Young[139], Joško Zaninović[140], Andrej Žitňan[141], Carles Lalueza-Fox[142], Peter de Knijff[11], Ian Barnes[12], Peter Halkon[143], Mark G. Thomas[144], Douglas J. Kennett[145], Barry Cunliffe[146], Malcolm Lillie[36,147], Nadin Rohland[2,6], Ron Pinhasi[9], Ian Armit[5], David Reich[1,2,6,10]

* Co-lead authors: N. Patterson, M. Isakov, T. Booth, L. Büster, C.-E. Fischer

Correspondence to: R. Pinhasi (ron.pinhasi@univie.ac.at), I. Armit (ian.armit@york.ac.uk), and D. Reich (reich@genetics.med.harvard.edu)

[1]Department of Human Evolutionary Biology, Harvard University, Cambridge, MA 02138
[2]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[3]Harvard College, Cambridge, MA 02138, USA
[4]The Francis Crick Institute, London, NW1 1AT, UK
[5]Department of Archaeology, University of York, York YO1 7EP, UK
[6]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA
[7]BIOMICs Research Group, University of the Basque Country UPV/EHU, 01006 Vitoria-Gasteiz, Spain
[8]Ikerbasque – Basque Foundation of Science, 48009 Bilbao, Spain
[9]Department of Evolutionary Anthropology, University of Vienna, 1090 Vienna, Austria
[10]Howard Hughes Medical Institute, Boston, MA 02115, USA
[11]Department of Human Genetics, Leiden University Medical Center, Leiden 2333 ZC, The Netherlands
[12]Department of Earth Sciences, Natural History Museum, London SW7 5BD, UK
[13]Servizio di Bioarcheologia, Museo delle Civiltà, 00144, Rome, Italy
[14]Institutes of Energy and the Environment, The Pennsylvania State University, University Park, PA 16802, USA
[15]CIAS, Department of Life Sciences, University of Coimbra, 3000-456 Coimbra, Portugal
[16]Department of Biological and Geographical Sciences, School of Applied Sciences, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, UK
[17]INSITU S.C.P. Centelles, 08540 Barcelona, Spain
[18]Oxford Archaeology, Oxford, OX2 0ES, UK
[19]Městské Muzeum v Čelákovicích, 250 88 Čelákovice, Czech Republic
[20]National Museum of Antiquities, Leiden, Netherlands
[21]Faculty of Archaeology, Leiden University, 2333 CC Leiden, The Netherlands
[22]Independent Researcher, Soham, Ely, CB7 5DG, UK
[23]Institut National de Recherches Archéologiques Préventives (INRAP), 75685 Paris cedex 14, France
[24]Cotswold Archaeology, Cirencester, Gloucestershire, GL7 6BQ, UK
[25]Department of Archaeology and Classical Studies, Stockholm University, 106 91 Stockholm, Sweden
[26]Leeds Museums and Galleries, Leeds Discovery Centre, Leeds LS10 1LB, UK
[27]Department of Anthropology, Hungarian Natural History Museum, H-1083 Budapest, Hungary
[28]Institute of Preservation of Archaeological Heritage of Northwest Bohemia, 43401 Most, Czech Republic
[29]Stratum Ltd., 21218 Seget Donji, Croatia
[30]TRACES, Ecole des Hautes Etudes en Sciences Sociales, 31058 Toulouse, France
[31]School of History, Classics and Archaeology, University of Edinburgh, Edinburgh EH8 9AG, UK
[32]Cambridge Archaeological Unit, Department of Archaeology, University of Cambridge, Cambridge CB2 3DZ, UK
[33]Corinium Museum, Park Street, Cirencester, GL7 2BX, UK
[34]Wiltshire Museum, Devizes SN10 1NS, UK
[35]South West Heritage Trust, Somerset Heritage Centre, Norton Fitzwarren, Taunton, TA2 6SF, UK
[36]Department of Historical, Philosophical and Religious Studies, Umeå University 901 87, Umeå, Sweden
[37]Kaducej Ltd., 21000 Split, Croatia
[38]Instituto de Medicina Legal y Ciencias Forenses de Cantabria, Santander, Spain

[39]Université de Franche Comté / UMR Chrono-Environnement, 25030 Besançon, France

[40]School of Earth and Environmental Sciences, University of Manchester, Manchester, M13 9PL, UK

[41]Archaeological Museum Zadar, 23000 Zadar, Croatia

[42]Department of Environmental Biology, University of Rome, La Sapienza, 00185 Rome, Italy

[43]Department of Archaeology, Centre for Interdisciplinary Research in Archaeology, Faculty of Arts, 1000 Ljubljana, Slovenija

[44]Institute for the Protection of Cultural Heritage of Slovenia, Conservation Centre, Centre for Preventive Archaeology, SI-1000 Ljubljana, Slovenia

[45]School of Natural Sciences, University of Central Lancashire, Preston, Lancashire PR1 2HE, UK

[46]Hungarian National Museum, 1113 Budapest, Hungary

[47]Institute of Archaeology of the Czech Academy of Sciences, Prague, 11801 Prague 1, Czech Republic

[48]Tees Archaeology, Hartlepool, TS24 8BT, UK

[49]Guernsey Museums & Galleries, Candie Gardens, St Peter Port, Guernsey, GY1 1UG, UK

[50]Amgueddfa Cymru – National Museum Wales, Cardiff CF10 3NP, UK

[51]Kuny Domokos Museum, H-2890 Tata, Hungary

[52]Research Laboratory for Archaeology, School of Archaeology, University of Oxford, Oxford OX1 3TG, UK

[53]Institute of Archaeology, 10000 Zagreb, Croatia

[54]Department of Anthropology, The National Museum, Prague, Prague 9, 19300 Horní Počernice, Czech Republic

[55]Department of Archaeology, Dobó István Castle Museum, 3300 Eger, Hungary

[56]The Orkney Museum, Kirkwall KW15 1DH, UK

[57]Archaeological Museum in Zagreb, 10000 Zagreb, Croatia

[58]British Geological Survey, Keyworth Nottingham, Nottinghamshire NG12 5GG, UK

[59]Archaeological Museum Osijek, 31000 Osijek, Croatia

[60]Archaeology and Ancient History, University of Leicester, Leicester, LE1 7RH, UK

[61]School of History, Classics and Archaeology, Newcastle University, Newcastle upon Tyne NE1 7RU, UK

[62]Manx National Heritage, Eiraght Ashoonagh Vannin, Manx Museum, Douglas IM1 3LY, Isle of Man

[63]Ásatárs Ltd., 6000 Kecskemét, Hungary

[64]Heritage and Archaeological Research Practice (HARP), Edinburgh, EH2 3JG

[65]Instituto Internacional de Investigaciones Prehistóricas de Cantabria, Universidad de Cantabria, 39005 Santander, Cantabria, Spain

[66]Grupo EvoAdapta, Departamento de Ciencias Históricas, Universidad de Cantabria, Santander, Spain

[67]The Salisbury Museum, Salisbury SP1 2EN, UK

[68]Department of Biological Anthropology, Institute of Biology, Faculty of Science, Eötvös Loránd University, Hungary

[69]VUhbs Archaeology, VU University, 1081 HV Amsterdam, The Netherlands

[70]Department of Biological Anthropology, Eötvös Loránd University, H-1117 Budapest, Hungary

[71]Scottish Universities Environmental Research Centre (SUERC), East Kilbride, G75 0QF, UK

[72]Bureau Archeologie en Bodemkwaliteit, Gemeente Nijmegen, 6500 HG Nijmegen, The Netherlands

[73]Musée Saint-Remi, Ville de Reims, 51000 Reims, France

[74]Savaria Museum, 9700 Szombathely, Hungary

[75]Jósa András Museum, H-4400 Nyíregyháza, Hungary
[76]Cornwall Archaeology Unit, The, Truro TR1 3AY, UK
[77]Faculty of Humanities, University of Primorska, 6000 Koper, Slovenia
[78]Musée de La Cour d'Or, 57000 Metz, France
[79]East Dorset Antiquarian Society (EDAS), Merrow, West Bexington, Dorset, DT2 9DG, UK
[80]Institute of Archaeogenomics, Research Centre for the Humanities, Eötvös Loránd Research Network, 1097 Budapest, Hungary
[81]Groningen Institute of Archaeology, Groningen University, Groningen, Netherlands
[82]Geology and Geochemistry Cluster, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands
[83]Damjanich János Museum, H-5000 Szolnok, Hungary
[84]Archaeological Institute of the Slovak Academy of Sciences, 949 21 Nitra, Slovakia
[85]Institute of Archaeology, Research Centre for the Humanities, Eötvös Loránd Research Network, 1097 Budapest, Hungary
[86]University of Szeged, Faculty of Arts, Department of Archaeology, 6722 Szeged, Hungary
[87]Musee d'Histoire et d'Archeologie, 56000 Vannes, France
[88]School of History, Archaeology and Religion, Cardiff University, Cardiff, CF10 3EU, UK
[89]Wessex Archaeology, Salisbury SP4 6EB, UK
[90]Group EcoPast. CRETUS. Area of Archaeology, Department of History, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain
[91]Lower Winskill Farm, Langcliffe, Settle, BD24 9PZ, UK
[92]Geophiz.biz, West Heslerton, North Yorkshire, YO17 8RQ, UK
[93]Zavičajni Muzej Grada Rovinja, 52210 Rovinj, Croatia
[94]Royal Pavillion and Museums, Brighton BN1 1EE, UK
[95]Walcherse Archeologische Dienst, 4331 CK Middelburg, The Netherlands
[96]Conservation Department in Šibenik, Ministry of Culture of the Republic of Croatia, 22000 Šibenik, Croatia
[97]Museum of West Bohemia, Department of Prehistory, 30100, Czech Republic
[98]Royal Cornwall Museum, Truro, Cornwall TR1 2SJ, UK
[99]Archaeological Museum of Istria, 52100 Pula, Croatia
[100]National Museum of Kikinda, 23300 Kikinda, Serbia
[101]Department of Archaeology, Durham University, Durham, DH1 3LE, UK
[102]MOLES Archaeology, Courtwood House, Silver Street Head, Sheffield, S1 2DD, UK
[103]University of Bristol Spelaeological Society, University of Bristol Student Union Building, Queens Road, Bristol, BS8 1LN, UK
[104]Municipal Monument Preservation Institute Bratislava, 811 01, Bratislava, Slovakia
[105]Archaeological and Forensic Sciences, University of Bradford, Bradford, BD7 1DP, UK
[106]Centre for Applied Bioanthropology, Institute for Anthropological Research, 10000 Zagreb, Croatia
[107]National Trust, Dorset & Wiltshire, Tisbury, Wiltshire SP3 6LW, UK
[108]Archaeological Museum, Dubrovnik Museums, 20000 Dubrovnik, Croatia
[109]Rómer Flóris Museum of Art and History, 9021 Győr, Hungary
[110]Departament de Biologia Animal, de Biologia Vegetal i d'Ecologia, Universitat Autònoma de Barcelona, Bellaterra, Cerdanyola del Vallés, Spain
[111]The City of Prague Museum, 11000 Prague 1, Czech Republic
[112]Musée Le Vergeur, 51100 Reims, France
[113]Independent researcher, H 1097 Budapest, Hungary
[114]Department of Anthropology, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

[115]Institute for Quaternary Palaeontology and Geology, Croatian Academy of Sciences and Arts, 10000 Zagreb, Croatia

[116]Centre for Human Bioarchaeology, Museum of London, 150 London Wall, London EC2Y 5HN, UK

[117]National Trust For Scotland, Hermiston Quay, Edinburgh EH11 4DF, UK

[118]The Novium Museum, Chichester, PO19 1QH, UK

[119]Provinciaal Archeologisch Depot Noord-Holland, 1901 Castricum, The Netherlands

[120]Department of Prehistory and Classical Antiquity, The National Museum, Prague, 11579 Prague 1, Czech Republic

[121]Slovak National Museum – Natural History Museum, 810 06 Bratislava 16, Slovakia

[122]National Museums Scotland, Edinburgh EH1 1JF, UK

[123]The Dock Museum, Barrow-in-Furness, Cumbria LA14 2PW, UK

[124]Rippl-Rónai Museum, 7400 Kaposvár, Hungary

[125]Katona József Museum Hungary, Kecskemét, H-6000, HUNGARY

[126]MAP Archaeological Practice, Showfield Lane, Malton, North Yorkshire, YO17 6BT, UK

[127]Wosinsky Mór Museum, 7100 Szekszárd, Hungary

[128]MTA - ELTE Research Group for Interdisciplinary Archaeology, Eötvös Loránd University, 1088 Budapest, Hungary

[129]Museo del Hombre Dominicano, Santo Domingo 10204, Dominican Republic

[130]Craven Museum and Gallery, Skipton, BD23 1AH, UK

[131]Department of Anthropology, Natural History Museum Vienna, 1010 Vienna, Austria

[132]Instituto Monte Bernorio de Estudios de la Antigüedad del Cantábrico (IMBEAC) & Universidad Complutense de Madrid, Facultad de Geografía e Historia, Departamento de Prehistoria, Historia Antigua y Arqueología, 28040, Madrid, Spain

[133]Hampshire Cultural Trust, Winchester, SO23 8RB, UK

[134]Rijksdienst voor het Cultureel Erfgoed, Amersfoort, Netherlands

[135]Maritime Cultures Research Institute, Department of Art, Sciences, and Archaeology, Vrije Universiteit Brussel, 1050 Brussels, Belgium

[136]Research Unit: Analytical, Environmental & Geo-Chemistry, Department of Chemistry, Vrije Universiteit Brussel, 1050 Brussels, Belgium

[137]Archaeological Research Services, Angel House, Portland Square, Bakewell, Derbyshire, DE45 1HB, UK

[138]Swale and Thames Archaeological Survey Company, Faversham, Kent, ME13 8UP, UK

[139]Keswick Museum, Keswick CA12 4NF, UK

[140]Krka National Park, 22000 Šibenik, Croatia

[141]AA AVALA s.r.o., 821 04 Bratislava, Slovakia

[142]Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, 08003 Barcelona, Spain

[143]Department of History, University of Hull, Hull, HU6 7RX, UK

[144]Research Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, UK

[145]Department of Anthropology, University of California, Santa Barbara, CA 93106, USA

[146]Institute of Archaeology, University of Oxford, Oxford OX1 2PG, UK

[147]Department of Geography, Geology and Environment, University of Hull, Hull, HU6 7RX, UK

Present-day populations from England and Wales harbour more ancestry derived from Early European Farmers (EEF) than did people of the Early Bronze Age. To study how this occurred, we generated genome-wide data from 803 individuals, almost all from the Middle to Late Bronze Age and the Iron Age, increasing data in this period from Britain by 12-fold, and from Western and Central Europe by more than two-fold. Between 1000-875 BCE, EEF ancestry increased in southern Britain (England and Wales) but not in northern Britain (Scotland), due to incorporation into the population of a stream of migrants who arrived at this time and in previous centuries and who were genetically most similar to ancient individuals from France. These migrants cumulatively contributed about half the ancestry of Iron Age people of England and Wales, thereby documenting a previously unknown demographic process that is a plausible vector for the spread of early Celtic languages into Britain. These patterns are part of a broader trend of EEF ancestry proportions becoming more similar across Central and Western Europe in the Middle to Late Bronze Age, coincident with intensification of cultural exchange and highlighting this period as a peak of interaction and mobility. We find no evidence of a comparable rate of migration into Britain in the Iron Age. The distinct genetic trajectories of Britain and continental Europe in the Iron Age are also exemplified by the fact that the allele conferring lactase persistence rose to ~50% frequency in Britain by this time whereas it was only ~7% in Central Europe and underwent a comparable rise in frequency only a millennium later, a pattern that could only occur if there were qualitative differences in how dairy products were used in Britain and in Central Europe.

"Migration" is a central concept in both population genetics and archaeology, but its meaning has evolved in divergent ways in the course of the development of these disciplines[1]. Population geneticists use "migration" to refer to any movement of genetic material from one region to another, but from an archaeological perspective this can be too broad, since this definition would see even low-level symmetrical exchanges of mates between adjacent communities as representing migration. Archaeologists generally use "migration" to refer to processes that result in significant demographic change due to directed and permanent translocation of people from one region to another[2,3]. In European archaeology, discussions of prehistoric migrations have become fraught due to the ways in which theories of migration were exploited for political purposes in the early-mid twentieth century—when movement of large numbers of people over a short period of time was sometimes argued to be a primary mechanism for the spread of ethnic groups[4] and archaeological reconstructions of such events were used to justify claims on territory[5]. Setting a high bar for theorizing migration, however—for example by only using it

6

35   when there is clear evidence for organized movements of people over a short period of time—
36   minimizes the important effects that cumulatively large-scale movements are likely to have had
37   in shaping prehistory[2], such as the westward migration of people from the Steppe beginning in
38   the third millennium BCE that genetic data have unambiguously shown contributed much of the
39   ancestry of later Europeans including eventually people from Britain[6,7,8,9]. We emphasize that
40   the large-scale Middle to Late Bronze Age migrations we document here unfolded over
41   hundreds of years as attested by our radiocarbon dating data and so cannot all be explained by
42   movement over a short time.

43   Whole genome ancient DNA studies have shown that the first Neolithic farmers of the island of
44   Great Britain (hereafter Britain) who lived 3950-2450 BCE derived roughly 80% of their
45   ancestry from Early European Farmers (EEF) who originated in Anatolia more than two
46   millennia earlier, and 20% from Mesolithic hunter-gatherers (Western European Hunter-
47   Gatherers: WHG)[8-10]. The WHG ancestry in the first British farmers was almost entirely due to
48   admixture in continental Europe, indicating that migrants in the Neolithic brought this WHG
49   ancestry with them and local WHG contributed negligibly to the ancestry of later people in
50   Britain. The Neolithic population of Britain was genetically similar to, and almost certainly
51   derived from, contemporaneous populations from continental Europe, especially Iberia and
52   France. This ancestry profile remained stable for about a millennium and a half. From around
53   2450 BCE, the onset of the Chalcolithic period, there was another substantial movement of
54   people into Britain (minimum 90% of ancestry from the new migrants) coinciding with the
55   spread of Bell Beaker traditions from continental Europe which brought a third major
56   component: 'Steppe ancestry' derived originally from people living on the Pontic-Caspian
57   Steppe ~3000 BCE[8]. In the original study[8] reporting this ancestry shift in Britain, no significant
58   average change in the proportion of EEF ancestry was detected from the Chalcolithic/Early
59   Bronze Age (C/EBA; 2450-1550 BCE), through the Middle Bronze Age (MBA; 1550-1150
60   BCE) and Late Bronze Age (LBA; 1150-750 BCE), to the pre-Roman Iron Age (IA; 750 BCE-
61   43 CE). However, that study contained little data after 1300 BCE (Figure 1). Today, however,
62   EEF ancestry is significantly higher on average in southern Britain (defined here as England and
63   Wales although we recognize modern boundaries are arbitrary) than in northern Britain
64   (Scotland), raising the question of when this increase in EEF ancestry occurred[8,11]. The rise in
65   EEF ancestry in England and Wales cannot be explained by migration from northern continental
66   Europe in the early medieval period[12,13], as early medieval migrant groups harboured less EEF
67   ancestry than was present in Bronze Age Britain and hence would have decreased EEF ancestry
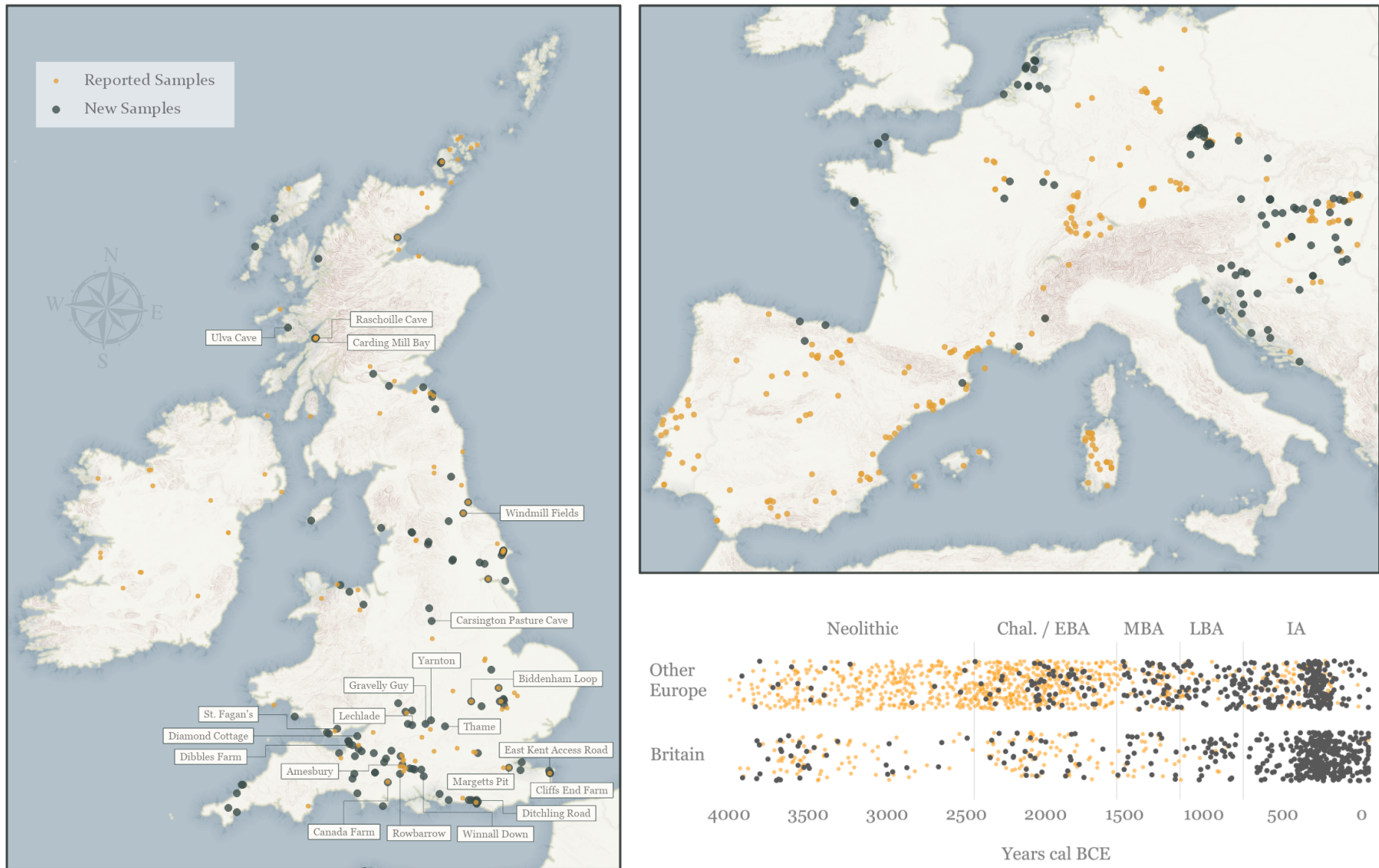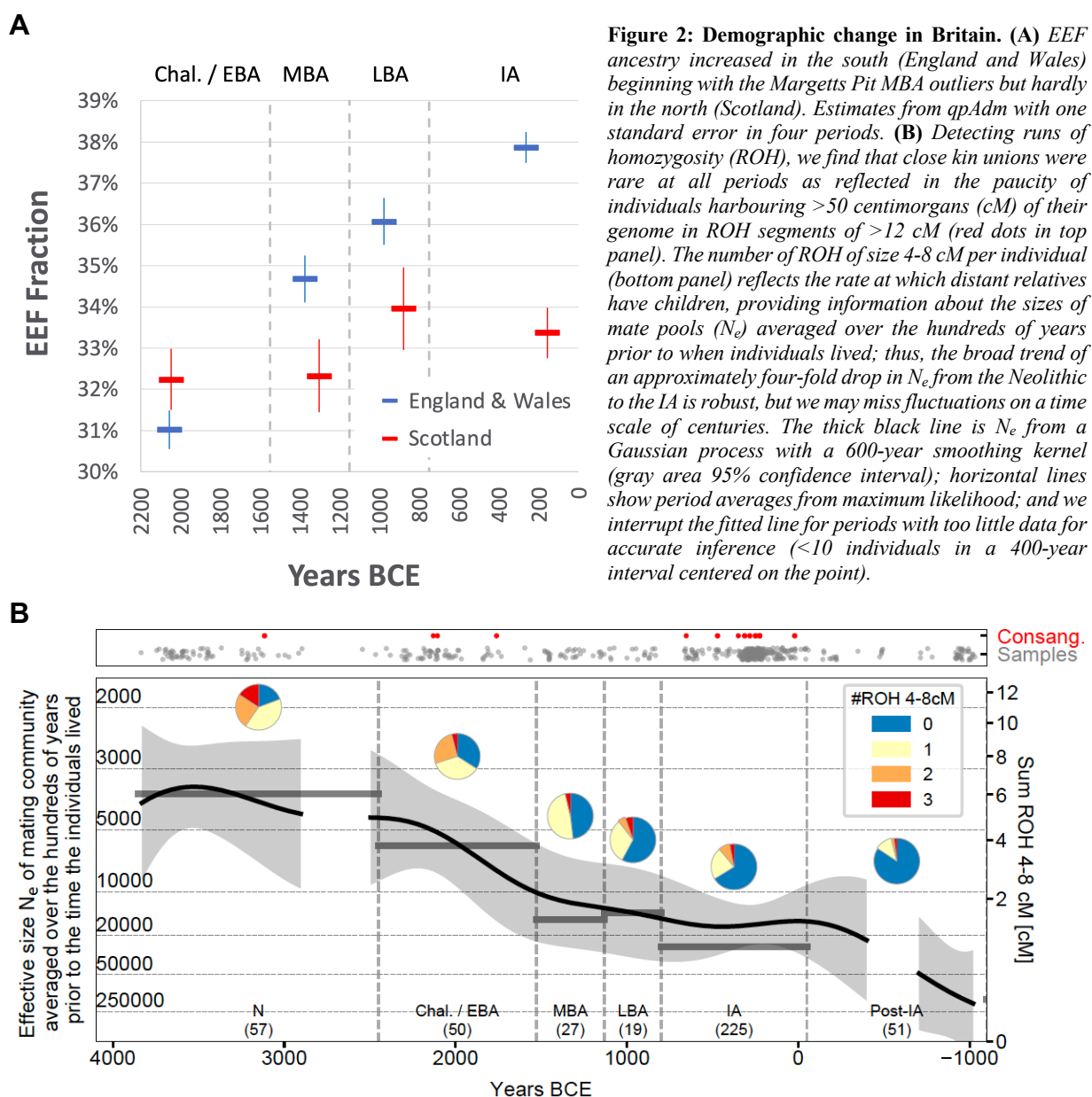68   instead of increasing it as we observe[8].

Figure 1: Ancient DNA Dataset. *Geographic distribution of sites and temporal distribution of individuals in the period 4000 BCE-43 CE. Newly reported data are in black; published data are in orange. In the map of Britain we label sites that harbour individuals who are ancestry outliers relative to others of the same period. In the timeline we denote the archaeological periods according to the British chronology: Neolithic (3950-2450 BCE), Chalcolithic and Early Bronze Age (C/EBA, 2450-1550 BCE), Middle Bronze Age (MBA, 1550-1150 BCE), Late Bronze Age (LBA, 1150-750 BCE), and pre-Roman Iron Age (IA, 750 BCE-43 CE). To aid visualisation, we add jitter on the Y axis and randomly sample dates from their probability distributions given in Online Table 1.*

69   We generated new ancient DNA data from 426 previously unanalysed individuals from Britain,

70   increasing the number of pre-Roman individuals to 608 and multiplying by 28-fold the number

71   from the combined period of the LBA and IA (from 13 to 365) (Fig. 1, Supplementary

72   Information section 1, Online Table 1). We also report data from previously unanalysed ancient

73   individuals mostly dating to the LBA and IA from the Czech Republic (n=160), Hungary

74   (n=54), France (n=52), the Netherlands (n=28), Slovakia (n=25), Croatia (n=21), Slovenia

75   (n=14), Spain (n=10), Serbia (n=8) and Austria (n=3). We increased data quality on 33

76   previously published individuals (Online Table 1). To generate these data, we prepared powder

77   from bones and teeth, extracted DNA[14-16], and generated 1033 sequencing libraries all

78   pretreated with the enzyme uracil-DNA glycosylase to reduce characteristic cytosine-to-thymine

79   errors of ancient DNA (Online Table 2)[14,15,17]. We enriched libraries in solution and sequenced

80   them.[18-20] We co-analysed with previously reported data for a total of 5928 ancient and present-

81   day individuals (Online Table 3). We clustered by time and geography aided by 126 newly

82   reported radiocarbon dates (Online Table 4). We separately labelled individuals that were

83   significantly different in ancestry from the majority cluster from each time period and region

84   (Supplementary Information section 2, Online Table 5). Although we report data from all

85   individuals, we removed a subset from the main analysis: those with evidence of contamination

86   based on finding variation in parts of their genome where they are expected to have DNA from

87   only one parent[21,22], those with a rate of damage in the final nucleotide lower than is typical for

88   authentic ancient DNA[15], those that were first degree relatives of other higher coverage

89   individuals in the dataset[23], or those that had too little data for accurate ancestry inference

90   (<30,000 single nucleotide polymorphisms (SNPs) covered at least once) (Online Table 5).

91   Figure 1 shows a map of analysed individuals. We identified 127 individuals from 50 families

92   as related (within the third degree) to at least one other newly reported individual in the dataset

93   (Online Table 6).

94   We computed symmetry-f_4 statistics[24,25] between all pairs of temporal groupings of individuals

95   in Britain, testing for differences in the rate of allele sharing (genetic drift) with the two major

96   source populations (Steppe and EEF); we computed standard errors using a Block Jackknife.

97   We document a significant increase in the degree of allele sharing with EEF populations in

98   England and Wales over the Middle to Late Bronze Age (M-LBA) and into the IA (Extended

99   Data Table 1). To estimate the proportions of EEF, Steppe, and WHG ancestry in any "Target"

100   population[26], we leveraged the fact that the magnitudes of f_4-statistics can be interpreted in

101   terms of mixture fractions. Specifically, *qpAdm* models the f_4-statistics in each "Target"

102   population as a linear combination of the f_4-statistics in populations we use as proxies for the

103   mixing "Sources." *qpAdm* computes statistics of the form $f_4(Target, Source_i; Outgroup_j,$

104    $Outgroup_k)$ that compare Target and Sources to a set of four "Outgroups" that we optimized to

105   provide more precise estimates of ancestry proportions than previous *qpAdm* setups

106   (Supplementary Information section 2). Here, our proxies for the "Sources" are 22 early Balkan

107   Neolithic farmers with minimal hunter-gatherer admixture (EEF); 20 Yamnaya and Poltavka

108   pastoralists (Steppe); and 18 Mesolithic hunter-gatherers from across Western Europe (WHG).

A

Figure 2: Demographic change in Britain. (A) *EEF ancestry increased in the south (England and Wales) beginning with the Margetts Pit MBA outliers but hardly in the north (Scotland). Estimates from qpAdm with one standard error in four periods.* **(B)** *Detecting runs of homozygosity (ROH), we find that close kin unions were rare at all periods as reflected in the paucity of individuals harbouring >50 centimorgans (cM) of their genome in ROH segments of >12 cM (red dots in top panel). The number of ROH of size 4-8 cM per individual (bottom panel) reflects the rate at which distant relatives have children, providing information about the sizes of mate pools (N_e) averaged over the hundreds of years prior to when individuals lived; thus, the broad trend of an approximately four-fold drop in N_e from the Neolithic to the IA is robust, but we may miss fluctuations on a time scale of centuries. The thick black line is N_e from a Gaussian process with a 600-year smoothing kernel (gray area 95% confidence interval); horizontal lines show period averages from maximum likelihood; and we interrupt the fitted line for periods with too little data for accurate inference (<10 individuals in a 400-year interval centered on the point).*

B

109   Applying *qpAdm* to our data, we find that EEF-related ancestry increased in England and Wales

110   from 31.0±0.5% in the C/EBA (*n*=69), to 34.7±0.6% in the MBA (*n*=26), to 36.1±0.6% in the

111   LBA (*n*=23), and stabilized in the IA at 37.9±0.4% (*n*=273) (here and below, we quote one

10

112   standard error). There is no statistically significant change in Scotland (Figure 2 and Extended

113   Data Table 1). EEF ancestry was widespread in southern Britain by the IA, averaging 36.0-

114   38.8% in eight regions of England (Table 1, Extended Data Table 2) (IA Wales sample sizes are

115   too small to provide accurate inference). Within England, EEF ancestry proportions inferred

116   over the mostly later IA individuals do not correlate to latitude (Table 1). We considered the

117   possibility that the rise in EEF ancestry in southern Britain was due to a resurgence of

118   archaeologically less visible populations with more ancestry from people living in Britain

119   during the Neolithic. However, our attempts to model IA populations of England and Wales as a

120   mixture of groups who lived in Neolithic and C/EBA Britain failed at high significance

121   (Extended Data Figure 1). This is due to LBA and IA populations in Britain sharing excess

122   alleles with Neolithic populations from continental Europe not seen in earlier groups from

123   Britain (Supplementary Information section 3). The most plausible explanation is migration

124   from outside Britain into southern Britain in the M-LBA.

125   We modelled each individual from Britain using *qpAdm*, labelling significant ancestry outliers

126   relative to most individuals of their period as candidates for migrants or their recent descendants

127   (Figure 3, Extended Data Figure 2). We discuss each group of outliers from earliest to latest.

Table 1: Regional variation in ancestry in Iron Age Britain

Region	N	Lat-it-ude	Modeling Ancestry With Pre-Bronze Age Sources				Modeling Ancestry with Middle to Late Bronze Age Sources	
			P-value	WHG	EEF	Steppe	P-value	Margetts Pit and Cliffs End Farm-like ancestry
Scotland Orkney	2	59	0.22	14.2 ± 1.1%	34.1 ± 1.2%	51.6 ± 1.6%	0.10	20 ± 9%
Scotland West	4	58	0.12	13.0 ± .8%	32.3 ± 1.0%	54.7 ± 1.2%	0.19	8 ± 7%
Scotland Southeast	12	56	0.67	12.1 ± .6%	33.9 ± .7%	54.0 ± .9%	0.39	16 ± 5%
England North	10	54	0.35	13.4 ± .6%	36.3 ± .8%	50.3 ± 1.0%	0.76	35 ± 5%
England East Yorkshire	47	54	0.61	13.2 ± .4%	37.0 ± .5%	49.8 ± .6%	0.86	44 ± 4%
England Midlands	18	53	0.66	12.6 ± .5%	36.0 ± .6%	51.4 ± .8%	0.77	36 ± 4%
England Southwest	84	53	0.30	13.7 ± .4%	38.7 ± .4%	47.6 ± .6%	0.56	55 ± 5%
England East Anglia	21	52	0.44	13.5 ± .5%	37.0 ± .5%	49.5 ± .7%	0.52	44 ± 4%
England Southcentral	38	52	0.32	13.9 ± .4%	38.8 ± .5%	47.2 ± .6%	0.35	56 ± 5%
England Southeast	3	51	0.13	13.9 ± .5%	38.3 ± .5%	47.8 ± .6%	0.40	52 ± 5%
England Cornwall	16	50	0.40	13.5 ± .5%	36.4 ± .7%	50.1 ± .8%	0.64	39 ± 5%
Wales North	1	53	0.20	12.1 ± 1.6%	34.7 ± 2.0%	53.2 ± 2.5%	0.53	22 ± 14%
Wales South	2	51	0.66	14.2 ± 1.2%	38.6 ± 1.5%	47.2 ± 1.8%	0.57	53 ± 11%

Notes: Regions are ordered first by large grouping (Scotland-England-Wales), then reverse order of latitude. For modeling ancestry with pre-Bronze Age sources the right set is (OldAfrica, WHGA, Balkan_N, OldSteppe); for modeling of ancestry with M-LBA sources it is OldAfrica, OldSteppe, Turkey_N, Netherlands_C.EBA, Poland_Globular_Amphora, Spain.Portugal_4425.to.3800BP, CzechRepublic.Slovakia.Germany_3800.to.2700BP, Sardinia_8100.to.4100BP, CzechRepublic.Slovakia.Germany_4465.to.3800.BP, Sardinia_4100.to.2700BP, Spain.Portugal_6500.to.4425BP. We separate "England East Yorkshire" from "England North" because of East Yorkshire's distinctive cultural context (Arras).

128   First, replicating previous results[8,10], we infer a cluster of Neolithic individuals from western

129   Scotland with high WHG admixture, likely reflecting unions between recent migrants from

130   Europe and descendants of local Mesolithic groups in Britain (Extended Data Figure 2).

131   Second, we infer high variability in EEF ancestry in the C/EBA, before EEF ancestry became
132   relatively homogeneous after ~2000 BCE[8] (Figure 3). This is apparent at Amesbury Down
133   where EEF ancestry in some burials is significantly below the average of 29.9±0.4% (e.g. I2417
134   at 22.2±1.8%), and in others above the group average (e.g. I2416 at 47.7±2.2% and I14200 at
135   45.3±2.2%). The low EEF group are plausibly Beaker-period migrants who mixed with local
136   Neolithic farmers to produce the intermediate EEF ancestry that prevailed by the end of the
137   EBA. Within the Beaker-associated burials, individual I14200 with elevated EEF ancestry is
138   known as the "Amesbury Archer". He was buried in the most well-furnished grave recovered
139   from the Stonehenge mortuary landscape, and his isotopic profile indicates he spent parts of his
140   childhood outside Britain, possibly in the Alps[27]. The fact that the Archer was a migrant but had
141   too little Steppe ancestry to be from the population that drove Steppe ancestry to the high level
142   observed in C/EBA Britain, shows that Bell Beaker-associated migrants to Britain were not
143   from a homogeneous population. The 'Companion' (I2565), a burial found next to the Archer
144   whose isotopic profile like most others at the site is consistent with a local upbringing, was not
145   an ancestry outlier for C/EBA Britain (32.7±3.0% EEF; P=0.47 for consistency with the period
146   average; Figure 3). The Archer and the Companion shared a rare tarsal morphology and similar
147   grave goods that have been hypothesized to reflect a close genetic relationship (Supplementary
148   Information section 4)[28], but our results rule out first- or second-degree genetic relatedness.

149   Third, we observe four outliers with high EEF ancestry in the late MBA and LBA who are
150   candidates for being first generation migrants or the offspring of recent migrants from
151   continental Europe, all of whom were buried in Kent in the southeasternmost part of Britain.
152   The earlier two are from the site of Margetts Pit: 47.8±1.8% in individual I13716 (1391-1129
153   calBCE) and 43.6±1.8% ancestry in I13617 (1214-1052 calBCE). The latter two are from Cliffs
154   End Farm: 43.2±2.0% in individual I14865 (967-811 calBCE) and 43.4±1.8% in individual
155   I14861 (912-808 calBCE). We considered the possibility that we are observing the effect of a
156   short burst of migration in the MBA, explaining the Margetts Pit outliers, followed by co-
157   existence of separate communities with different proportions of EEF ancestry for at least a
158   couple of hundred years, thereby explaining the Cliffs End Farm outliers. However, strontium
159   and oxygen isotope analyses identify multiple individuals of non-local origin at Cliffs End
160   Farm[29] including the genetic outlier I14861, suggesting that the stream of migrants continued
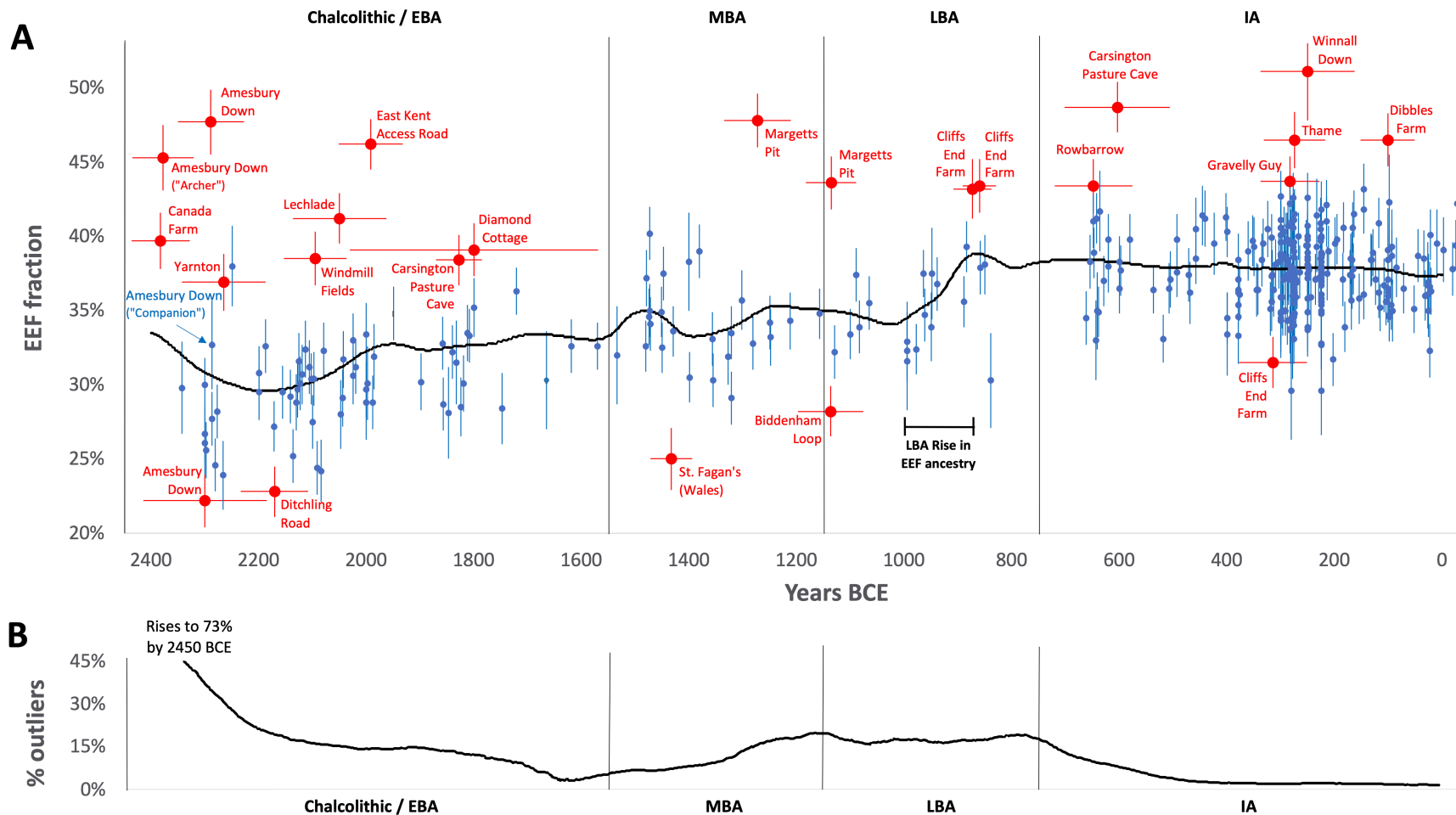161   into the second half of the LBA (Supplementary Information section 5).

Figure 3: By-individual analysis of the England and Wales time transect. (A) *Estimates of EEF ancestry and one standard error for all individuals that fit a three-way admixture model (EEF + WHG + Yamnaya) at p>0.01 using qpAdm and date to 2450 BCE-43 CE (we plot the mean of the posterior interval of the ^{14}C date or the mean of the archaeological context range). Most individuals are shown in blue while significant outliers at the ancestry tails are in red with a horizontal error bar to show one standard error for the date (outliers are identified either as p<0.005 based on a qpWave test from the main cluster from their period and |Z|>3 for a difference in EEF proportion, or p<0.1 and |Z|>3.5) (Online Table 5). Outliers like the four MBA and LBA individuals at Margetts Pit and Cliffs End Farm fit as the source for the increase in EEF ancestry in the IA. The black line shows an estimate of population-wide EEF ancestry proportion at each time obtained by weighting each individual's EEF estimate by the inverse square of their standard error as well as the probability that their date falls at each point in time (based on the date mean and error given in Online Table 5; we filter out individuals with date errors >120 years). The incorporation of increased EEF ancestry into the majority of samples occurred ~1000-875 BCE. **(B)** Proportion of outliers measured over 300-year sliding windows centered on each point; to estimate this we randomly sample dates of all individuals 100 times assuming normality and their mean and standard deviation in Online Table 5 (we remove individuals with EEF errors >0.022 and date standard errors >120 years). Major periods of migration into Britain are evident as periods when elevated proportions of individuals are outliers: between 2450-1800 BCE (17% outliers) and 1300-750 BCE (17% outliers again). Our ability to temporally resolve the period of outliers is poorer than our ability to resolve the period of incorporation of increased EEF ancestry into the broader population; thus, for example, the earliest outlier with elevated EEF ancestry in the M-LBA at Margetts Pit I13716 could have died as late as 1129 BCE based on the uncertainty of her ^{14}C date.*

162   Fourth, the fraction of individuals whose ancestry is significantly different from the main group

163   is 17% over the first part of the C/EBA (2450-1800 BCE), much lower at 4% from the end of

164   the EBA through the beginning of the MBA (1800-1300 BCE), high again at 17% between the

165   end of the MBA through the LBA (1300-750 BCE), and low again at 3% through the IA (Figure

166   3). This is consistent with two major periods of migration into southern Britain in the

167   Chalcolithic and then again in the M-LBA. We considered the possibility that our failure to

168   observe a high rate of outliers in the IA compared with the immediately preceding period was

169   because ancestry had homogenized to an extent between northern and southern Europe by this

170   time, which could make outliers more difficult to detect. However, the average EEF ancestry in

171   Britain in the IA was 37.9±0.4%, which is substantially different from much of contemporary

172   Western and Central Europe—52.6±0.6% in Iberia, 49.8±0.4% in Austria, Hungary, and

173   Slovenia, 45.4±0.5% in the Czech Republic, Slovakia and Germany, 45.6±0.5% in France and

174   Switzerland, and 34.4±1.2% in the Netherlands (Figure 4A)—which would have made the

175   majority of migrants from these regions detectable given the small standard errors in most

176   individual ancestry estimates of <2% (Online Table 5). Our sampling from western France and

177   Belgium is poor, and it is possible that EEF ancestry proportions there were similar to Britain,

178   so we cannot rule out migration from this region in the IA. Nevertheless, the lack of a change in

179   EEF ancestry proportion in the IA is consistent with reduced migration from continental Europe

180   and suggests relatively closed and self-sustaining social communities in Britain during the IA[30].

181   Demographic change in Britain over our period of study is also evident from another aspect of

182   the data: the rate of observations of runs of homozygosity (ROH), which can occur when a

183   person's parents are relatively closely related. The larger the pool of people from which

184   individuals draw their mates, the less likely it is for parents to be closely related, and thus we

185   can average the number of 4-8 centimorgan (cM) ROH segments to estimate the effective size

186   of the pool of people within which people were mating in the ~600 year period prior to the time

187   when the analysed individuals lived[31]. We find that the size of the mating pool increased by

188   roughly four-fold from the Neolithic to the IA, but we caution that this inference should not be

189   interpreted as an estimate of population size changes over this period as mating pool sizes are

190   also affected by changing social customs. First, if the distance over which people ranged to find

191   their mates was higher in some cultural contexts than in others it would cause the mating pool

192   sizes to be different even if there was no difference in actual population size; for example,

193   mating pool size may be less than the island-wide population size if members of communities

194   mixed little with their neighbours as is plausible in the British IA where there was high regional

195 variation in material culture[30], or larger if individuals mated not only with people outside their

196 local communities but also outside Britain. Second, we have a major gap in sampling at the end

197 of the Neolithic (roughly 3000-2450 BCE) and thin sampling at the beginning of the IA, which

198 means that demographic processes in these periods may be missed. Third, due to the method

199 effectively averaging over centuries, this analysis may also fail to detect major population

200 declines over the space of a few decades, if these were followed by rapid population recovery.

201 Future work that complements measurement of changing degree of parental relatedness over

202 time with estimates of changing degree of relatedness of individuals across communities[32]

203 should make it possible to examine how much of this signal of increasing size over time was

204 driven by true growth in population size, and how much by increasing rates of gene exchange

205 across geographically dispersed communities.

206 We co-analysed our ancient DNA time transect in Britain alongside European transects (Figure

207 4A, Online Tables 5 and 7). Average EEF ancestry increased in this period in North-Central

208 Europe and the Netherlands, just as in Britain. The first individuals from North-Central Europe

209 (Czech Republic/Slovakia/Germany) with greatly increased EEF ancestry are associated with

210 artefacts traditionally classified as part of the LBA Knoviz culture, a component of the broader

211 Urnfield cultural complex (1300-800 BCE) that spread across much of Central Europe. This is

212 especially striking as the Knoviz individuals are genetically consistent with being from a very

213 similar population as the Margetts Pit and Cliffs End Farm outliers (Supplementary Information

214 section 6). Later individuals in North-Central Europe have similar EEF proportions, consistent

215 with continuity through the LBA-IA with earlier Urnfield-associated groups. By contrast, in

216 MBA and LBA France/Switzerland and South-Central Europe (Austria/Hungary/Slovenia) there

217 was little change in average EEF ancestry, while EEF ancestry decreased in MBA and LBA

218 Iberia (Spain/Portugal). There are also two exceptions to this broad European pattern of

219 ancestry convergence in Europe—Scotland in the far north, and Sardinia in the far south[33,34]—

220 both of which have extreme proportions of EEF ancestry (Scotland low and Sardinia high)

221 relative to the European average across this period (Online Table 7).

222 Placing Britain within the broader context of Europe is also illuminating with respect to the

223 frequency trajectories of genetic variants of phenotypic importance. This study multiplies by

224 almost eight-fold the number of IA individuals with genome-wide data from Western and

225 Central Europe (from 80 to 624; Online Table 5), making it possible for the first time to

226 accurately track the frequency change of genetic variants into the IA (Online Table 8). In

227 addition to showing how variants associated with light skin pigmentation, such as *SLC45A2,*

228 became more common in the IA throughout Europe[19], we obtain an unexpected result for the

229 derived allele at the polymorphism *MCM6-LCT* rs4988235 correlated with lactase persistence

230 into adulthood[19] (Extended Data Figure 3). Previous work showed that the frequency of this

231 allele in IA Iberia was only a small fraction of its present-day incidence[35], which we confirm by

232 showing that its frequency was ~9% in the IA compared to ~40% today. Similarly, in a pool of

233 individuals from Central Europe (Austria, Hungary, Slovenia, Czech Republic, Slovakia and

234 Germany) its frequency was ~7% in the IA compared to ~48% today. However, in Britain, most

235 of the rise in frequency had occurred by the IA (50% compared to the current 73%), suggesting

236 that intense selection acted earlier and the main rise in frequency occurred about a millennium

237 earlier than in Central Europe (Figure 4B, Extended Data Figure 3). We find no evidence that

238 the main rise in frequency of the lactase persistence allele occurred in continental Europe and

239 that its frequency rise in Britain was due to M-LBA migration. Thus, the Margetts Pit and Cliffs

240 End Farm outliers who are plausible members (or recent offspring) of the migrating population

241 did not carry the allele, and most of its frequency rise in Britain occurred after the M-LBA

242 (Figure 4B, Online Table 8). This raises questions about whether dairy products were consumed

243 in a qualitatively different way or were economically more important in LBA-IA Britain than in

244 Central Europe.

245 The ancestry change in Britain we document during the M-LBA was more subtle than the one

246 that occurred during the Neolithic and Beaker-period migrations. In England and Wales, fpr

247 example, allele frequency differentiation between the Neolithic and C/EBA was F_{ST}~0.02, but

248 between the C/EBA and the IA it was an order of magnitude smaller at F_{ST}~0.002 (Extended

249 Data Table 1). The pre-LBA population in Britain also made a substantial genetic contribution

250 to the IA population—its genetic signature was not even close to being completely replaced—

251 which contrasts with the two earlier major Holocene ancestry shifts[8,10]. Direct evidence for a

252 substantial contribution from the earlier population specifically on the male line comes from Y

253 chromosome haplogroup R1b-P312 L21/M529 (R1b1a1a2a1a2c1), which was present at 89±5%

254 in sampled individuals from C/EBA Britain and has always been nearly absent on the continent

255 (Online Table 9), and today occurs much more frequently in Britain and Ireland (frequency of

256 14-71% depending on region[36]) than in continental Europe (Extended Data Figure 4).
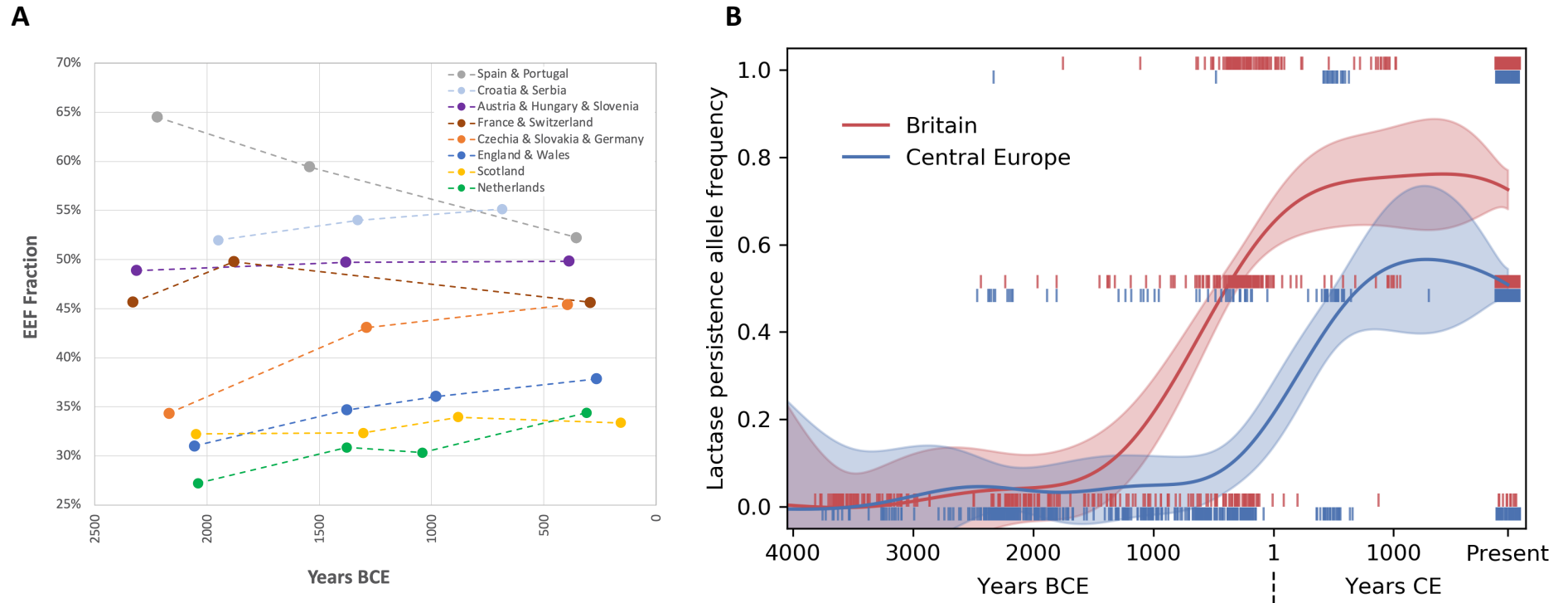
Figure 4: Genetic change in Britain in the context of Europe-wide trends. (A) *North-south ancestry convergence. We show eight ancient DNA time transects for up to four periods, using the average of dates of individuals in periods defined for each region as in Online Table 5 (we use the N/C/EBA/MBA/LBA/IA division into archaeological periods for Britain, but avoid using it for most of the rest of Europe because chronological boundaries differ). Dotted lines show which points are regionally grouped and should not be interpreted as implying a smooth change over time.* ***(B)*** *The allele conferring lactase persistence made its major rise about a millennium earlier in Britain than in Central Europe suggesting different selection regimes and possibly cultural differences in the use of dairy products in the two regions in the IA. We visualise the frequency trajectory of the lactase persistence allele at SNP rs4988235 by using the GaussianProcessRegressor function from the Scikit-learn library in Python with parameter alpha=0.1 and 1*RationalQuadratic kernel with parameter length_scale_bounds=(1, 1000). We use the GLIMPSE[37] software to impute diploid genotype posterior probabilities (GP), restricting to samples with max(GP)>0.9 for this SNP. The analysis includes 459 ancient individuals from Britain and 468 from Central Europe (Czech Republic, Slovakia, Croatia, Hungary, Austria, Germany and Slovenia); to represent modern Britain we use a pool of 190 CEU and GBR individuals from the 1000 Genomes Project[38], and to represent modern Central Europe we use 288 individuals from the modern Czech Republic[39]. Each vertical bar represents the derived allele frequency for each individual with values [0, 0.5, 1]; we use jitter on the x-axis for each vertical bar, and we show in shading the inferred 95% confidence interval for the allele frequency at each time point.*

To gain insight into the possible sources of the ancestry that spread across southern Britain by the end of the LBA, we fit the pooled IA individuals from England and Wales in *qpAdm* as a mixture of a group related to the main C/EBA cluster, and a second source. We tested 65 second sources—63 from continental Europe and 2 from Britain (the Margetts Pit outlier pool, and the Cliffs End Farm outlier pool)—and found that 20 fit at p>0.05; we then pooled the genetically similar Margetts Pit and Cliffs End Farm individuals and performed further *qpAdm* testing with more stringent *qpAdm* setups, leaving eight second sources that consistently fit well with modest standard errors (Table 2, Supplementary Information section 6). The Margetts Pit and Cliffs End Farm pool fit as contributing 49.4±3.0% of the ancestry of IA southern Britons, providing our best estimate of the degree of genetic turnover. The seven continental populations that fit as sources are estimated to contribute 24-69%; thus, even without results from representatives of the putative source population living in Britain itself, we infer large genetic turnovers. Although only 20% of the continental candidate populations we tested are from France, 86% of the fitting populations are: the fitting populations are four from Occitanie in southern France (600-200 BCE), two from Grand Est in northeastern France (800-200 BCE), and one from Spain (a ~600 BCE group). These fitting second sources all post-date the ancestry change in Britain by hundreds of years and hence they cannot be the true sources; however, they are plausibly descended from earlier local populations. An origin in France is also suggested by the fact that all of the high EEF outliers in Britain in the M-LBA, and all of the 1000-875 BCE individuals that track the ramp-up of EEF ancestry from MBA to IA levels, are from Kent in far southeastern Britain (Extended Data Figure 5). Southwestern individual I12624 from Blackberry Field, Potterne in Wiltshire (950-750 BCE) had a typical EEF proportion for the IA of southern Britain, suggesting the spread of descendants of the migrant stream throughout southern Britain may have already begun by this time, although we caution that this is the only datapoint we have in the second half of the LBA from outside Kent, and so more sampling is needed to understand the geographic and temporal course of the spread of this ancestry. Our data point to the new ancestry being ubiquitous in southern Britain by the beginning of the IA.

Table 2: Populations fitting as proxies for the new ancestry in IA southern Britain

Surrogate for source of the new ancestry	N	Mean	p-value	Ancestry
Margetts Pit and Cliffs End Farm MLBA	4	1036 BCE	0.07	49.4 ± 3.0%
Spain IA Tartessian	2	629 BCE	0.16	23.7 ± 1.2%
France GrandEst IA1 (shotgun data)	5	620 BCE	1.00	48.9 ± 3.7%
France Occitanie IA2 (high EEF subgroup, shotgun data)	1	450 BCE	0.85	25.8 ± 1.7%
France Occitanie IA2 (high WHG subgroup, shotgun data)	1	450 BCE	0.39	33.5 ± 4.1%
France Occitanie IA2 (shotgun data)	2	400 BCE	0.25	53.3 ± 5.4%
France Occitanie IA2 (low Steppe subgroup, shotgun data)	2	363 BCE	0.33	36.5 ± 2.6%
France GrandEst IA2	12	250 BCE	0.09	68.5 ± 3.3%

284 We can fully explain the significant regional variation in ancestry in IA England and Wales

285 (Table 1 and Extended Data Table 2) as driven by different proportions of ancestry from the

286 population to which the Margetts Pit and Cliffs End Farm individuals belonged, obtaining

287 estimates ranging from 35±5% in northern England to 56±5% in south-central England (Table

288 1). Archaeological data show that the IA was a period when material culture was increasingly

289 regional in character[30]; as we show here, this was accompanied by subtle genetic structure. We

290 highlight the case of East Yorkshire, where most individuals we analysed are associated with

291 'Arras Culture' contexts comprising square-ditched barrows and occasional chariot burials[40,41];

292 similarities between Arras funerary traditions and those of IA societies in the Paris Basin and

293 the Ardennes/Champagne regions of France and Belgium have led to suggestions that East

294 Yorkshire was influenced by direct migration from continental Europe in the IA[42]. Our estimate

295 of the Margetts Pit/Cliffs End Farm ancestry source for East Yorkshire burials is 44±4% (Table

296 1) which is typical for middle latitudes of Britain at this time (East Anglia is similar). However,

297 the East Yorkshire burials are distinctive in another way: regional differentiation in IA Britain,

298 as measured by F_{ST}, is higher between East Yorkshire and other groups than between any other

299 pairs of IA populations in England and Wales (Extended Data Table 2). Without ancient DNA

300 from the putative IA sources for the Arras culture in continental Europe, we cannot determine if

301 the genetic differentiation we observe of people in Arras burial contexts is attributable to the

302 subsequent isolation in the IA of the population from the rest of southern Britain, or later

303 streams of migration specifically affecting East Yorkshire that we have limited power to detect.

304 The period from 1500-1150 BCE has long been recognized as a time when cultural connections

305 between Britain and regions of continental Europe intensified, and when societies on both sides

306 of the Channel shared cultural features including domestic pottery, metalwork and ritual

307 depositional practices[43-47]. From around 750 BCE there is more limited archaeological evidence

308 of contact between Britain and the continent, with little that would suggest the significant

309 movement of people[43]. Both the genetic and archaeological data concur in showing that, by the

310 beginning of the IA, there is little evidence of demographically significant migration into

311 Britain. Our findings do not establish whether the population movements we infer were a cause

312 or consequence of M-LBA exchange networks, but they do suggest that interactions between

313 local populations of Britain and new migrants bringing ideas from continental Europe could

314 have been a vector for some of the cultural change we see in M-LBA England and Wales. Much

315 of northern and central France is currently almost entirely unsampled, and thus we cannot at

316 present test if the gene flow between the two regions in this period was largely unidirectional.

Population movements are often a significant driver of cultural change, including in the languages people speak. While periods of intense migration such as the one we infer here do not always result in language shifts[35], genetic evidence of significant migration is important because it documents demographic processes that are plausible conduits for language spread[48,49]. Several researchers have interpreted linguistic data as providing evidence for early Celtic languages spreading into Britain from France at the end of the Bronze Age or in the Early Iron Age[50,51]. Our identification of substantial migration into Britain from sources that best fit populations in France provides an independent line of evidence in support of this, and points to the M-LBA as a prime candidate for the period of this language spread. While the lack of evidence for M-LBA EEF ancestry change in Scotland (Figure 2A) could be interpreted as a concern for the case that Celtic language spread into Britain at this time, in fact a later arrival of Celtic languages in Scotland is entirely consistent with evidence that non-Celtic and Celtic languages coexisted there into the first millennium CE[52]. Our finding of a decrease of EEF ancestry in Iberia, where the proportion was relatively high in the EBA, and a roughly simultaneous increase in Britain where the proportion was relatively low in the EBA (Figure 4a), could, in theory, reflect a Celtic-speaking group of people with intermediate EEF ancestry spreading into both regions, although such a simple model cannot explain all the north-south ancestry convergence in Europe (Supplementary Information section 7); the true gene flows were more complex. Nevertheless, the fact that the Margetts Pit and Cliffs End Farm outliers are genetically very similar to our Knoviz culture sample (Supplementary Information section 6) is striking in light of the fact that some scholars have hypothesized Central European Urnfield groups to have links to Celtic language spread[53]. In contrast, our failure to find evidence of large-scale migration into Britain from continental Europe in the IA suggests that, if Celtic language spread was driven by large-scale movement of people, it is unlikely to have occurred at this time. The adoption in IA Britain of cultural practices originating in continental Europe—particularly those linked to the La Tène tradition[54]—were also evidently independent of large-scale population movements, although there certainly were smaller movements of people, attested by individual IA outliers with high EEF ancestry such as those at Thame or Winnall Down (Figure 3).

An important direction for future work is to generate new ancient DNA data from continental contexts especially in central and western France—and also Ireland—to test the alternative scenarios of population history consistent with the finding in this study, and to develop theories integrating the genetic findings within archaeological frameworks.

Methods

Ancient DNA laboratory work

All ancient human skeletons analysed in this study were sampled with written permission of the stewards of the samples and every sample is represented by at least one co-author. Researchers who wish to obtain further information about specific samples should write to the corresponding authors and/or the authors who provided the archaeological contextualization for those samples in Supplementary Material section 1. In dedicated clean rooms at Harvard Medical School, the University of Vienna, the Natural History Museum in London, and the University of Huddersfield, as well as during sampling trips, we obtained powder from ancient bones and teeth using methods including fine sandblasting, drilling and milling[55,56]. We extracted DNA using a variety of methods[57-59], and prepared double- or single-stranded libraries treated with the enzyme Uracil DNA Glycosylase (UDG) to reduce characteristic errors associated with ancient DNA degradation[15-17,60]. We enriched these sequences manually or in multiplex using automated liquid handlers for sequences overlapping the mitochondrial genome[20,61] as well as about 1.24 million single nucleotide polymorphisms ("1240K capture")[18]. We pooled enriched libraries which we had marked with unique 7-base pair internal barcodes and/or 7- to 8-base pair indices and sequenced on Illumina NextSeq500 or HiSeqX10 instruments using paired-end reads of either 76 base pairs or 101 base pairs in length (Online Table 2).

Bioinformatic analysis

After trimming barcodes and adapters[6], we merged read pairs with at least 15 base pairs of overlap allowing no more than one mismatch if base quality was at least 20, or up to three mismatches if base qualities were <20; we chose the nucleotide of the higher quality in case of a conflict while setting the local base quality to the minimum of the two (for these steps we used a custom toolkit at https://github.com/DReichLab/ADNA-Tools). We aligned merged sequences to the mitochondrial genome RSRS[62] or the human genome hg19 (GRCh37, https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13/), and aligned these sequences using the *samse* command[63] of *BWA* version 0.7.15-r1140 with parameters -n 0.01, -o 2, and -l 16500. After identifying PCR duplicates by tagging all aligned sequences with the same start and stop positions and orientation and in some cases in-line barcodes using Picard MacDuplicates (http://broadinstitute.Github.io/picard/), and restricting to sequences that spanned at least 30 base pairs, we selected a single copy of each such sequence that had the

383   highest base quality score. For subsequent analysis, we trimmed the last 2 bases of each

384   sequence for UDG-treated libraries and the last 5 for non-UDG-treated libraries to reduce the

385   effects of characteristic errors associated with ancient DNA degradation. We built

386   mitochondrial consensus sequences, determined haplogroups using *HaploGrep2*[64] and

387   Phylotree version 17, and estimated the match rate to the consensus sequence using

388   contamMix v.1.0-12[21] when coverage was at least two-fold. To represent the nuclear data, we

389   randomly sampled a single sequence covering each of the 1.24 million SNP targets, and

390   estimated coverage based on the subset of these targeted SNPs on the autosomes. We used

391   *ANGSD* to estimate contamination based on polymorphism on the X chromosome in males

392   with at least 200 SNPs covered twice (males should be non-polymorphic if their data are

393   uncontaminated)[22]. We automatically determined Y chromosome haplogroups using both

394   targeted SNPs and off-target sequences aligning to the Y chromosome based on comparisons

395   to the Y chromosome phylogenetic tree from Yfull version 8.09 (https://www.yfull.com/),

396   providing two alternative notations for Y chromosome haplogroups: the first using a label

397   based on the terminal mutation, and the second describing all the associated branches of the

398   Y chromosome tree based on the notation of the International Society of Genetic Genealogy

399   (ISOGG) database version 15.73. (http://www.isogg.org).

400

401   **Determination of ancient DNA authenticity**

402   We determined ancient DNA authenticity based on five criteria. First, we required that the

403   lower bound of the 95% confidence interval for contamination from *ANGSD* (if we were able

404   to compute it) was <1%. Second, we required that the upper bound of the 95% confidence

405   interval for match rate to mitochondrial consensus sequence (if we were able to compute it)

406   was >95%. Third, we required that the average rate of cytosine-to-thymine errors at the

407   terminal nucleotide for all sequences passing filters was >3% for double-stranded partially

408   UDG-treated libraries[15] and >10% for single-stranded USER-treated libraries and double-

409   stranded non-UDG-treated libraries (the latter libraries are all from previously published data

410   that we reanalysed here)[65]. Fourth, we required the ratio of sequences mapping to the Y

411   chromosome to the sum of sequences mapping to the X and Y chromosome for the 1240K

412   data to be less than 3% (consistent with a female) or >35% (consistent with a male). Fifth, to

413   report an individual we required the number of SNPs covered at least once to be at least 5,000

414   (for most actual population genetic analyses, we required at least 30,000). For some

415   individuals with evidence of contamination, we analysed only sequences with terminal

416   damage to enrich for genuine ancient DNA using *pmdtools*, allowing us to study more

417   individuals[66]. We do not include in our main analyses data from 71 individuals that failed our

418   authenticity criteria (marked as "QUESTIONABLE" in Online Table 1); however, we

419   publish the data as part of this study as a resource.

420

421   **Approach to chronological uncertainty**

422   We restricted individuals for which we newly report data to those whose date estimate (mean

423   of the posterior distribution from radiocarbon carbon dating, or midpoint of the

424   archaeological context date) is older than 43 CE based on information we had available as of

425   July 1 2021. For the great majority of individuals, assignments to chronological periods did

426   not change subsequently. However, there were 23 exceptions, and we study these as part of

427   their original analysis groupings (Supplementary Information section 8).

428

429   **Population genetic analyses**

430   We used Principal Component Analysis as implemented in *smartpca* to visualise gradients of

431   ancestry, using the option *lsqproject* to project ancient individuals onto the patterns of genetic

432   variation learned from modern individuals[67]. We computed f_4-statistics and F_{ST} and carried out

433   *qpWave* and *qpAdm* analyses in ADMIXTOOLS[25]. We inferred relatives up to the third to fifth

434   degree using a previously described method[23].

435

436   **Allele frequency estimates of variants with functional importance**

437   We clustered individuals into the temporal groups specified in Online Table 5. To estimate

438   the allele frequency of a given SNP in a particular group for Online Table 8, we used

439   sequence counts at each SNP position in each individual and used a maximum likelihood

440   approach[19]. We obtained confidence intervals using the Agresti-Coull method implemented

441   in the *binom.confint* function of the R-package *binom*. The imputation-based methodology

442   for studying the trajectory of the lactase persistence allele is described in the Figure 4 legend.

443

444   **Accelerator Mass Spectrometry (AMS) radiocarbon dating**

445   We carried out AMS dating at a variety of laboratories; we refer readers to the individual labs

446   for the experimental procedures. We calibrated all dates using OxCal 4.4.2[68] and IntCal20[69].

447

448   **Reporting summary**

449   Further information on research design is available in the Nature Research Reporting Summary

450   linked to this paper.

Data availability

The raw data are available as aligned sequences (bam files) through the European Nucleotide Archive under accession number [to be made available on publication]. Less processed versions of the data (fastq files) are available from the corresponding authors on request, whereas more processed versions of the data (the genotype data obtained by random allele sampling used in analysis) are available at https://reich.hms.harvard.edu/datasets. Any other relevant data are available from the corresponding authors upon reasonable request.

Extended Data Table 1: Ancestry change over time in Britain

	Sample size	P-value	qpAdm results (3-way model) WHG	EEF	Steppe	WHG error	EEF error	time	England.and.Wales_N	England.and.Wales_C.EBA	England.and.Wales_MBA	England.and.Wales_LBA	England.and.Wales_IA	England.and.Wales_PostIA	England.and.Wales_Modern	Scotland_N	Scotland_C.EBA	Scotland_MBA	Scotland_LBA	Scotland_IA	Scotland_PostIA	Scotland_Modern	Ireland_N	Ireland_C.EBA	Ireland_PostIA	Ireland_Modern	Channel.Islands_8100.to.5700BP	Channel.Islands_5700.to.4450BP	Channel.Islands_IA
England.and.Wales_N	37	0.7597	20.8%	76.7%	2.6%	0.5%	0.5%	0.6%	■	0.02	0.0176	0.0171	0.0161	0.0219	0.0226	0.0013	0.0192	0.0188	0.0188	0.0197	0.0206	0.0239	0.0046	0.0275	0.0233	0.0225	0.0177	0.0073	0.0153
England.and.Wales_C.EBA	69	0.3840	12.6%	31.0%	56.4%	0.4%	0.5%	0.6%	-65.7	■	0.0007	0.0012	0.0017	0.0084	0.0107	0.0204	0.0013	0.0002	0.0013	0.0019	0.006	0.0109	0.0259	0.0112	0.0091	0.0085	0.0357	0.0173	0.0055
England.and.Wales_MBA	26	0.0918	13.5%	34.7%	51.8%	0.5%	0.6%	0.7%	-58.2	-7.3	■	0.0004	0.0008	0.0066	0.0088	0.0181	0.0011	0.0009	0.0013	0.0016	0.004	0.009	0.0227	0.0099	0.0064	0.0071	0.0333	0.0151	0.0043
England.and.Wales_LBA	23	0.4609	13.6%	36.1%	50.4%	0.5%	0.6%	0.7%	-52.3	-9.9	2.9	■	0.0006	0.0056	0.007	0.0179	0.0028	0.0012	0.0017	0.0022	0.0037	0.0077	0.0209	0.0089	0.0065	0.0052	0.0319	0.0141	0.0037
England.and.Wales_IA	273	0.3637	13.6%	37.9%	48.5%	0.3%	0.4%	0.5%	-63.9	-19.4	7	2.3	■	0.0053	0.0073	0.0175	0.0027	0.0011	0.0016	0.0018	0.0035	0.0076	0.0204	0.0099	0.0064	0.0049	0.0306	0.0136	0.0032
England.and.Wales_PostIA	38	0.0002	15.0%	36.6%	48.3%	0.4%	0.5%	0.6%	61	-11	-2.5	1	5.8	■	0.003	0.0239	0.0085	0.0051	0.0074	0.0076	0.0014	0.0037	0.0188	0.0069	4E-05	0.0024	0.0333	0.017	0.0049
England.and.Wales_Modern	62	0.6315	14.1%	40.0%	45.9%	0.4%	0.4%	0.6%	-61.3	-19.5	-8.8	-4	-3.5	8.5	■	0.0243	0.0107	0.0071	0.0094	0.0097	0.0034	0.0016	0.0184	0.0083	0.0029	0.0021	0.034	0.0175	0.0072
Scotland_N	44	0.6642	23.1%	74.3%	2.5%	0.4%	0.5%	0.6%	2.7	-65.1	-55.5	-51.3	-64.4	-61.3	-61.6	■	0.0184	0.0186	0.0182	0.0197	0.0227	0.026	0.0079	0.0296	0.0243	0.0248	0.0196	0.0084	0.0164
Scotland_C.EBA	10	0.1517	13.5%	32.2%	54.3%	0.6%	0.7%	1.0%	52	-3	1.6	4.3	6.4	3.5	7.8	-50.6	■	0.0011	0.002	0.0022	0.0064	0.0107	0.0243	0.0099	0.0079	0.0098	0.0338	0.0194	0.0067
Scotland_MBA	5	0.5635	14.0%	32.3%	53.7%	0.8%	0.9%	1.1%	45.2	-1.7	2	4.1	6.2	3.9	7.4	-44.8	0.5	■	0.0009	0.0013	0.0032	0.0074	0.0216	0.0078	0.007	0.0061	0.032	0.0132	0.0036
Scotland_LBA	4	0.8346	12.4%	34.0%	53.7%	0.8%	1.0%	1.2%	39.8	-4	-0.1	1.3	3.2	1	4.2	-40.4	-1.1	1.7	■	0.0002	0.0047	0.0098	0.0239	0.0101	0.0084	0.0074	0.0357	0.0152	0.007
Scotland_IA	18	0.1850	12.7%	33.4%	54.0%	0.6%	0.6%	0.8%	56.1	-3.8	1.7	4.1	8.4	4.3	10.2	-56	0.2	1.1	-1.4	■	0.0047	0.0095	0.0251	0.0108	0.0083	0.0069	0.035	0.0178	0.0044
Scotland_PostIA	10	0.4713	12.9%	36.4%	50.7%	0.6%	0.7%	0.9%	50.4	-7.4	-1.5	1.2	3.7	0.3	5.1	48.3	-2.5	-3	-0.6	-2.9	■	0.0034	0.0189	0.0068	0.0021	0.0015	0.0331	0.0162	0.0037
Scotland_Modern	78	0.7341	14.3%	37.5%	48.2%	0.4%	0.4%	0.6%	62.1	-12.9	-3.5	0.2	5.1	-1.2	7.9	-62.4	-4.2	-4.5	-1.5	-5.5	1	■	0.0201	0.0089	0.0032	0.001	0.0352	0.0179	0.0078
Ireland_N	51	0.6505	21.6%	77.9%	0.5%	0.4%	0.5%	0.5%	-0.5	-69.3	-59	-54.9	-69.3	-65.8	-65.9	3.3	51.4	45.4	40.9	57.2	52	67.2	■	0.0238	0.0189	0.019	0.0183	0.0081	0.0158
Ireland_C.EBA	3	0.4166	13.6%	30.5%	55.9%	0.9%	1.2%	1.5%	37.9	1.5	4.7	6.4	8	5.9	9	-38	-3.3	-2.8	-4.3	-3.9	-5.4	-6.6	-38.8	■	0.0056	0.0068	0.0408	0.0256	0.0094
Ireland_PostIA	3	0.0109	14.0%	34.9%	51.1%	0.9%	1.1%	1.3%	37.6	-3.8	-0.3	1.5	3.1	1.1	4.1	-37.5	1.4	1.8	0	1.3	-0.8	-1.5	38.6	-3.9	■	0.0027	0.0336	0.0166	0.0049
Ireland_Modern	30	0.6461	12.9%	36.8%	50.3%	0.4%	0.5%	0.7%	57.6	-8.7	0	3.2	7.3	1.3	10.6	-56.8	1.8	1.7	0.5	3.6	-1.2	-3.7	-61.1	-5.5	-0.5	■	0.0346	0.0161	0.005
Channel.Islands_8100.to.5700BP	3	0.7577	16.1%	82.3%	1.6%	1.3%	1.4%	1.6%	3.5	36.4	33.7	31.8	32.7	33.2	31.8	4.4	33.8	30.8	28.6	33.9	32	33	3.3	29.8	29.3	30.3	■	0.0126	0.0266
Channel.Islands_5700.to.4450BP	3	0.4611	31.0%	67.1%	1.9%	1.2%	1.3%	1.4%	-7.9	28.1	24.7	23.7	23.8	24.4	22.7	-7	24.1	23.3	20.8	24.9	23.4	24.4	-8.3	23	20.5	21.1	-8.4	■	0.0099
Channel.Islands_IA	4	0.8603	15.4%	43.9%	40.7%	0.9%	1.2%	1.4%	-28.3	11.3	7.5	6	5.3	6.7	4.2	-27.3	7.3	7.8	6.5	8.5	6.7	6.4	-29.3	9.3	5.9	5.9	22.4	13.8	■

Note: We pool all individuals from each period and region removing those failing *qpAdm* modeling at p<0.01 according to Online Table 5). In the left columns are *qpAdm* estimates of ancestry for each group. Below diagonal are Z-scores from f_4*(Row population, Column population; Turkey_N, OldSteppe)* (highlighted in red if |Z|>3). Above diagonal are inbreeding-corrected F_{ST} values (highlighted in yellow if F_{ST}>0.005).

Extended Data Table 2: Fine structure in Iron Age Britain.

|	N	P-value for qpAdm (3-way model)	WHG (3-way model)	EEF (3-way model)	Steppe (3-way model)	WHG err. (3-way model)	EEF err. (3-way model)	Steppe err. (3-way model)	Scotland West	Scotland Southeast	Scotland Orkney	England Midlands	England North	England Cornwall	England East Anglia	England East Yorkshire	England Southeast	England Southwest	England Southcentral	Wales North	Wales South
			qpAdm results (3-way model)						*Tests for difference in ancestry between row & column (below diagonal f₄-statistic Z-score, above-diagonal F_ST)*												
Scotland West	4	0.12	13.0%	32.3%	54.7%	0.8%	1.0%	1.2%	0	0.0007	0.0006	0.0032	0.0035	0.0052	0.0035	0.0046	0.0034	0.004	0.0034	n/a	0.0038
Scotland Southeast	12	0.67	12.1%	33.9%	54.0%	0.6%	0.7%	0.9%	0.3	0.0	0.001	0.0012	0.0008	0.0028	0.0017	0.003	0.0014	0.0015	0.0019	n/a	0.0018
Scotland Orkney	2	0.22	14.2%	34.1%	51.6%	1.1%	1.2%	1.6%	0.7	1.1	0	0.0018	0.0013	0.0037	0.0007	0.0029	0.0014	0.0021	0.0021	n/a	0.0074
England Midlands	18	0.66	12.6%	36.0%	51.4%	0.5%	0.6%	0.8%	2.8	3.4	0.9	0.0	0.0001	0.0022	0.001	0.0028	0.0008	0.0009	0.0013	n/a	0.0016
England North	10	0.35	13.4%	36.3%	50.3%	0.6%	0.8%	1.0%	2.4	2.6	0.9	0.1	0	0.0027	0.0005	0.0016	0.0002	0.0007	0.0009	n/a	0.0019
England Cornwall	16	0.40	13.5%	36.4%	50.1%	0.5%	0.7%	0.8%	3.0	3.8	1.1	0.9	0.9	0.0	0.0025	0.0041	0.002	0.0021	0.0024	n/a	0.0024
England East Anglia	21	0.44	13.5%	37.0%	49.5%	0.5%	0.5%	0.7%	3.7	4.8	1.7	0.9	1.1	0.1	0	0.002	0.0007	0.0011	0.0013	n/a	0.0012
England East Yorkshire	47	0.61	13.2%	37.0%	49.8%	0.4%	0.5%	0.6%	4.1	5.4	2.1	1.5	1.7	-0.6	-0.5	0.0	0.0022	0.0026	0.0023	n/a	0.0028
England Southeast	36	0.13	13.9%	38.3%	47.8%	0.5%	0.5%	0.6%	5.4	7.2	2.8	-3.8	-3.2	-2.5	-3.4	-3.2	0	0.0008	0.0005	n/a	0.0008
England Southwest	84	0.30	13.7%	38.7%	47.6%	0.4%	0.4%	0.6%	5.6	8.4	3.3	-4.5	-4.3	-3.3	-3.7	-3.4	0.2	0.0	0.0009	n/a	0.0013
England Southcentral	38	0.32	13.9%	38.8%	47.2%	0.4%	0.5%	0.6%	5.6	7.5	3.3	-4.6	-3.6	-2.7	-3.0	-3.3	0.0	-0.2	0	n/a	0.0013
Wales North	1	0.20	12.1%	34.7%	53.2%	1.6%	2.0%	2.5%	0.8	1.1	2.0	1.9	2.0	2.5	2.9	3.1	3.6	3.6	3.5	0.0	n/a
Wales South	2	0.66	14.2%	38.6%	47.2%	1.2%	1.5%	1.8%	-2.7	-3.1	-1.5	-1.6	-1.3	-1.2	-1.0	-0.9	0.0	0.4	0.3	-1.9	0

Note: This is an expanded version of Table 1 including not just ancestry estimates for each group but also pairwise population comparisons. We pool all individuals from each period and region removing those failing *qpAdm* modeling at p<0.01 according to Online Table 5). In the left columns are *qpAdm* estimates of ancestry for each group. Below diagonal are Z-scores from $f_4(Row\ population,\ Column\ population;\ Turkey\_N,\ OldSteppe)$ (highlighted in red if |Z|>3). Above diagonal are inbreeding-corrected F_{ST} values (highlighted in yellow if F_{ST}>0.0025).

A

| Modeled population	N	Neolithic and C/EBA Groups Used in Modeling	
		England/Wales	Scotland
England/Wales MBA	26	0.34	0.046
England/Wales LBA	23	0.023	0.0074
England/Wales IA	273	$<10^{-6}$	$<10^{-6}$
Scotland MBA	5	0.88	0.028
Scotland LBA	4	0.25	0.77
Scotland IA	18	0.0091	0.0028

B
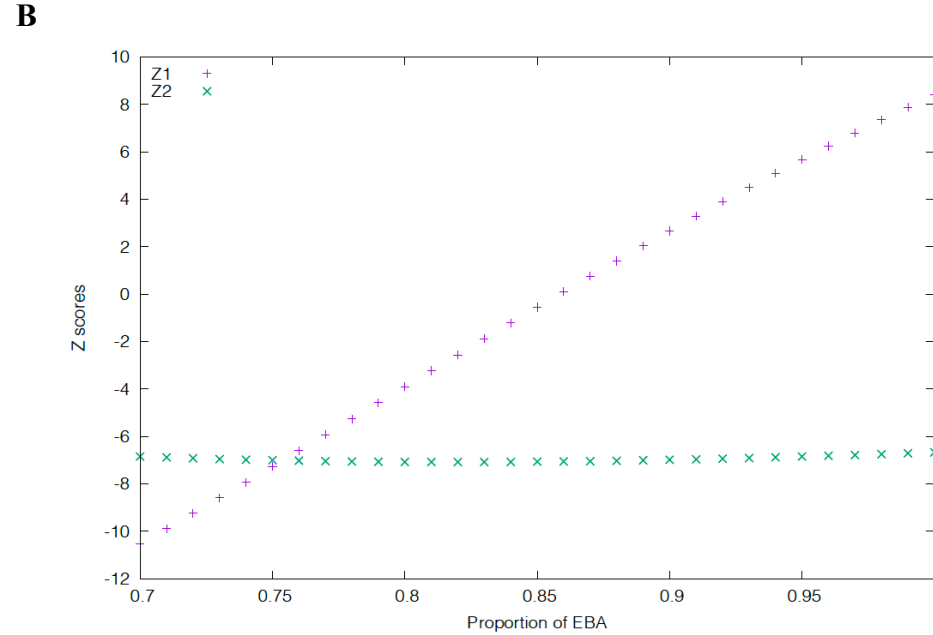
Extended Data Figure 1: Post-MBA Britain was not a mix of earlier British populations. (A) *qpAdm P-values for modeling British groups as a mix of Neolithic and Chalcolithic/EBA populations from England and Wales, Scotland, or Ireland (outgroups OldAfrica, OldSteppe, Turkey_N, CzechRepublic.Slovakia.Germany_3800.to.2700BP, Netherlands_C.EBA, Poland_Globular_Amphora, Spain.Portugal_4425.to.3800BP, CzechRepublic.Slovakia.Germany_4465.to.3800.BP, Sardinia_4100.to.2700BP, Sardinia_8100.to.4100BP, Spain.Portugal_6500.to.4425BP). We highlight p<0.05 (yellow) or p<0.005 (red). Both sources and target populations in this analysis remove outlier individuals ("Filter 2" in Online Table 5); we obtain qualitatively similar results when outlier individuals are not removed (not shown). In England and Wales, the main MBA group, and to a marginal extent the main LBA grouping, can be modelled as a mixture of the Neolithic and C/EBA populations; we can reject such a models for the IA (we can more weakly reject this model even for Scotland, whether using England/Wales or Scotland sources, suggesting some of the same stream of migration also affected Scotland albeit probably in a more subtle way). (B) To obtain insight into the source of the new ancestry in Britain in the IA, we computed f_4(England.and.Wales_IA, α(England.and.Wales_N) + (1-α)(England.Wales_C.EBA); R1, R2) for several different (R1, R2) pairs. If England.and.Wales_IA is a simple mixture of England.and.Wales_N and England.and.Wales_C.EBA without additional ancestry, then for some mixture proportion the statistic will be consistent with zero for all (R1, R2 pairs). When (R1, R2) = (OldAfrica, OldSteppe) feasible Z-scores (Z1 in the plot) are observed when α~0.85, showing that ~85% ancestry from England.and.Wales_C.EBA ancestry is needed to contribute the proportion of Steppe ancestry in England.and.Wales_IA. However, when (R1, R2) is (Balkan_N, Sardinian_8100.to.4100BP), we get infeasible Z-scores (Z2) of <-6 across the range where the other Z-score (Z1) is remotely feasible. Thus, Iron Age people from England and Wales must have ancestry from an additional population deeply related to Sardinian Early Neolithic groups.*
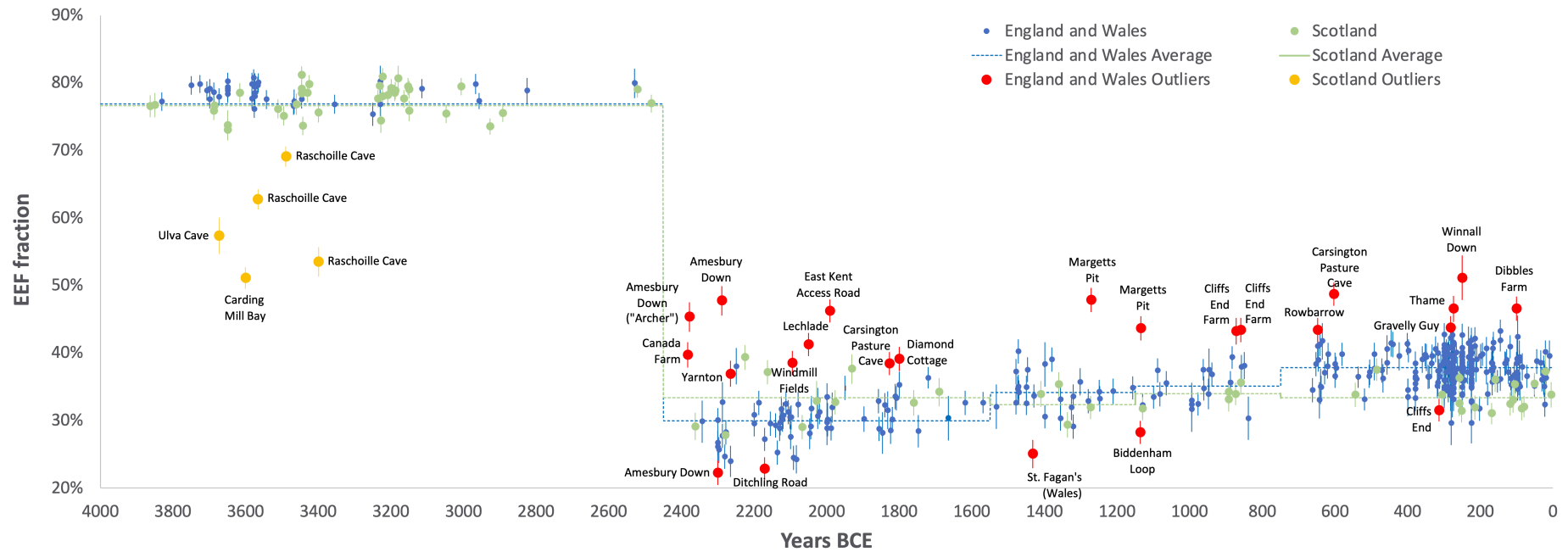
Extended Data Figure 2: By-individual analysis of the British time transect. *Version of Figure 3 with the time transect extended into the Neolithic, and showing the individuals from Scotland. Estimates of EEF ancestry and one standard error are shown for all individuals in the British time transect that pass basic quality control, that fit to a three-way admixture model (EEF + WHG + Yamnaya) at p>0.01 using qpAdm, and for the Neolithic period that fit a two-way admixture model (EEF + WHG) at p>0.01. Blue and green show individuals from southern and northern Britain that fit the average for the main cluster of their time, while red and orange show outliers at the ancestry tails (identified either as p<0.005 based on a qpWave test from the main cluster of individuals from their period and |Z|>3 for a difference in their EEF ancestry proportion from the period, or alternatively p<0.1 and |Z|>3.5). The averages for the main clusters in both southern and northern Britain in each archaeological period (Neolithic, C/EBA, MBA, LBA and IA) are shown in dashed lines.*
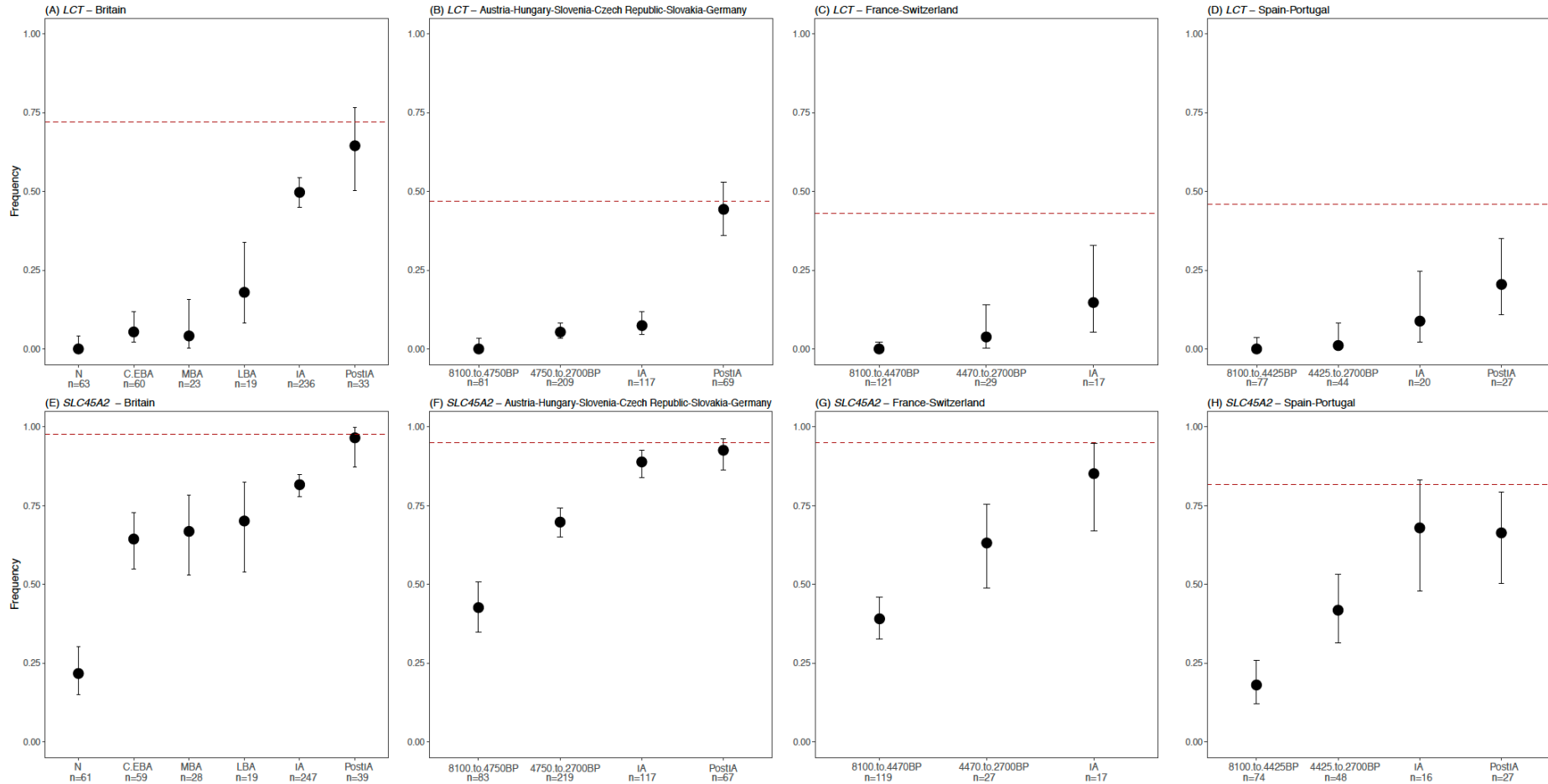
Extended Data Figure 3: Frequency change over time at two phenotypically important alleles. *Present-day frequency is shown by the red line. (A-D/Top) Lactase persistence allele at rs4988235. (E-H/Bottom) Light skin pigmentation allele at rs16891982. In Britain the rise in frequency of the lactase persistence allele occurred earlier than in Central Europe. This analysis is based on direct observation of alleles; imputation results are consistent (Figure 4B). Online Table 8 gives full numerical results for 107 phenotypically important alleles.*
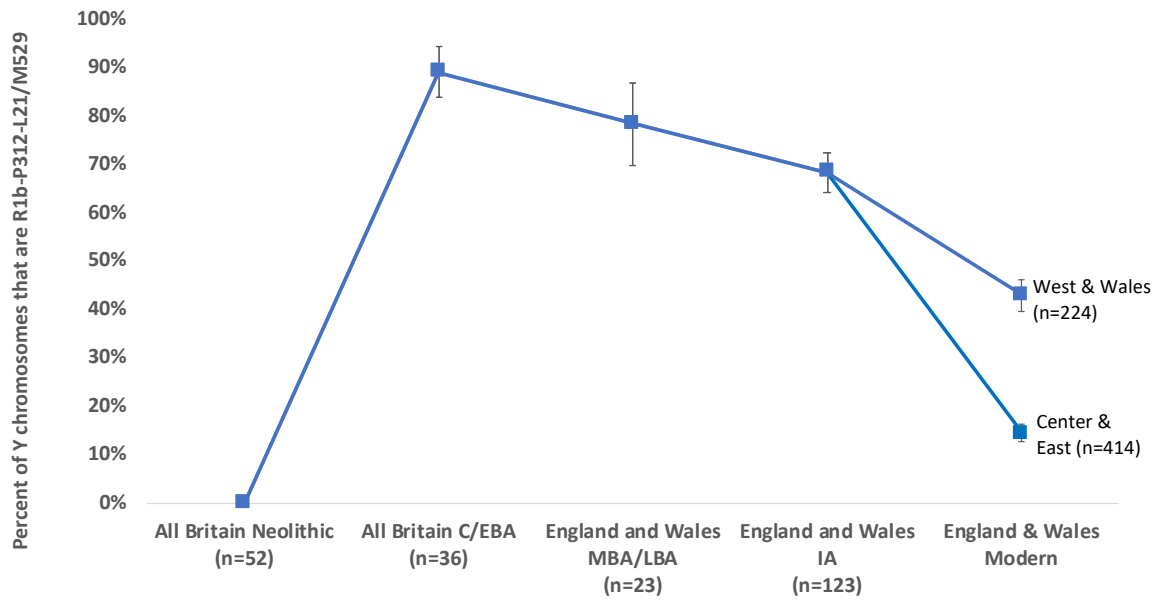
Extended Data Figure 4: Y chromosome haplogroup frequencies change over time. *We show the estimated frequency of the characteristically British Y chromosome haplogroup R1b-P312 L21/M529 in all individuals for which we are able to make a determination and which are not first-degree relatives of a higher coverage individual in the dataset. The frequency increases significantly from ~0% in the whole island Neolithic, to 89±4% in the whole island C/EBA. It declines albeit non-significantly to 79±9% in the MBA and LBA (from this time on restricting to England and Wales because of the autosomal evidence of a change in ancestry in the south but not the north). It further declines to 68±4% in the IA, a significant reduction relative to the C/EBA (P=0.014 by a two-sided chi-square contingency test). There is a further reduction from this time to the present, where the proportion is 43±3% in Wales and the west of England (P=5x10^{-6} for a reduction relative to the IA), and 14±2% in the center and east of England (P=3x10^{-32} for a reduction relative to the IA), potentially due to later migrations bearing a different distribution of Y chromosome haplogroup frequencies.*
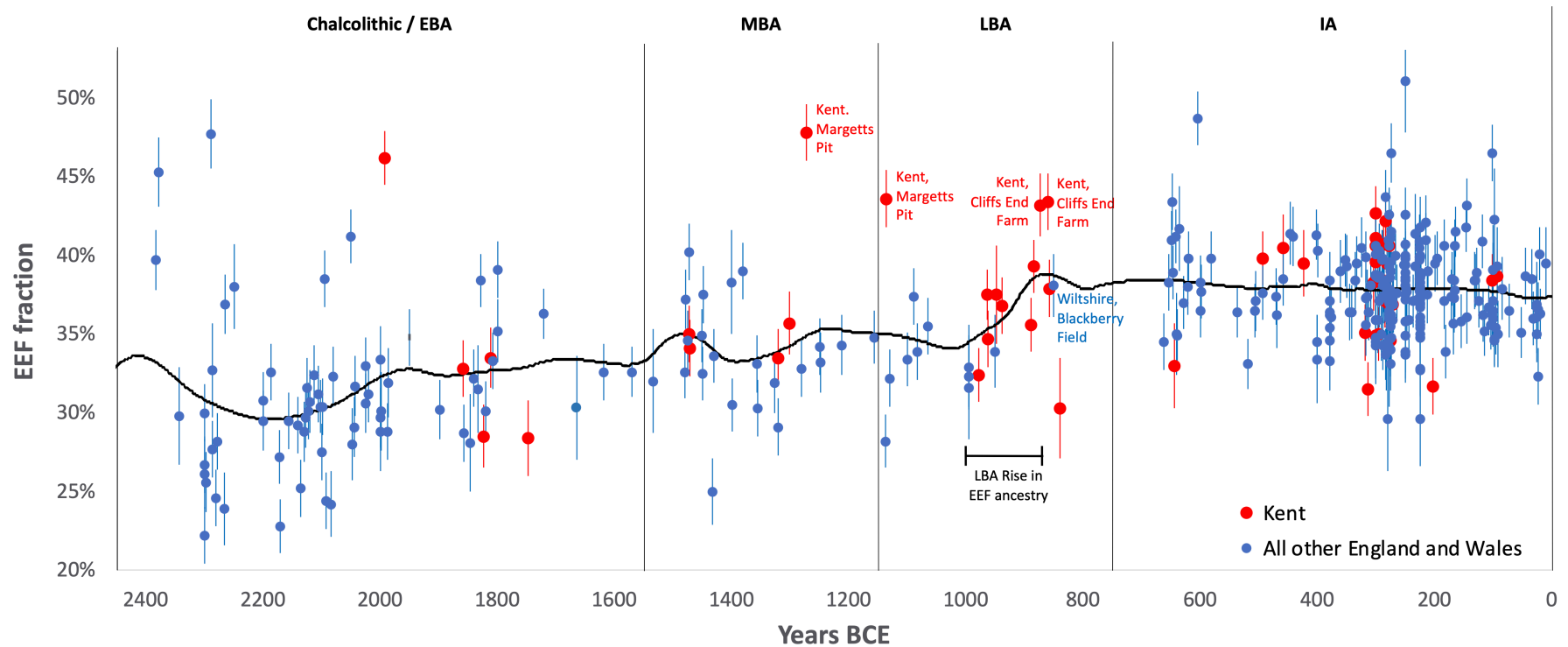
Extended Data Figure 5: Version of Figure 3A contrasting Kent to the rest of southern Britain. *We show the period 2450-1 BCE. All the high EEF outliers at the M-LBA are from Kent—the part of the island closest to France—and in addition all the individuals from 1000-875 BCE from the group of samples showing the ramp-up from MBA to IA levels of EEF ancestry are from Kent (5 from Cliffs End Farm and 3 from East Kent Access Road). This suggests the possibility that this small region was the gateway for migration to Britain at the M-LBA. Further sampling from the rest of Britain at the M-LBA is critical in order to understand the dynamics of how this ancestry spread more broadly, but the fact that only sample from the LBA that is not from Kent, I12624 from Blackberry Field in Potterne in Wiltshire at 950-750 BCE, already has IA levels of EEF ancestry, suggests that this ancestry began spreading more broadly by the date of this individual.*

References

508   1   Booth, T. J. A stranger in a strange land: a perspective on archaeological responses
509   to the palaeogenetic revolution from an archaeologist working amongst
510   palaeogeneticists. *World Archaeology* **51**, 586-601,
511   doi:10.1080/00438243.2019.1627240 (2019).
512   2   Anthony, D. W. Migration in archeology: The baby and the bathwater. *American*
513   *Anthropologist* **92**, 895-914 (1990).
514   3   Altschul, J. H. *et al.* Opinion: To understand how migrations affect human securities,
515   look to the past. *Proc Natl Acad Sci U S A* **117**, 20342-20345,
516   doi:10.1073/pnas.2015146117 (2020).
517   4   Kossinna, G. *Die Herkunft der Germanen: Zur Methode der Siedlungsarchäologie.*
518   (Kabitzsch, 1911).
519   5   Vander Linden, M. Population history in third-millennium-BC Europe: assessing the
520   contribution of genetics. *World Archaeology*, doi:10.1080/00438243.2016.1209124
521   (2016).
522   6   Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European
523   languages in Europe. *Nature* **522**, 207-211, doi:10.1038/nature14317 (2015).
524   7   Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522**, 167-+,
525   doi:10.1038/nature14507 (2015).
526   8   Olalde, I. *et al.* The Beaker phenomenon and the genomic transformation of
527   northwest Europe. *Nature* **555**, 190-196, doi:10.1038/nature25738 (2018).
528   9   Cassidy, L. M. *et al.* Neolithic and Bronze Age migration to Ireland and establishment
529   of the insular Atlantic genome. *Proc Natl Acad Sci U S A* **113**, 368-373,
530   doi:10.1073/pnas.1518445113 (2016).
531   10   Brace, S. *et al.* Ancient genomes indicate population replacement in Early Neolithic
532   Britain. *Nat Ecol Evol* **3**, 765-771, doi:10.1038/s41559-019-0871-9 (2019).
533   11   Galinsky, K. J., Loh, P. R., Mallick, S., Patterson, N. J. & Price, A. L. Population
534   structure of UK Biobank and ancient Eurasians reveals adaptation at genes
535   influencing blood pressure. *Am J Hum Genet* **99**, 1130-1139,
536   doi:10.1016/j.ajhg.2016.09.014 (2016).
537   12   Leslie, S. *et al.* The fine-scale genetic structure of the British population. *Nature* **519**,
538   309-314, doi:10.1038/nature14230 (2015).
539   13   Schiffels, S. *et al.* Iron Age and Anglo-Saxon genomes from East England reveal
540   British migration history. *Nat Commun* **7**, 10408, doi:10.1038/ncomms10408 (2016).
541   14   Briggs, A. W. *et al.* Removal of deaminated cytosines and detection of in vivo
542   methylation in ancient DNA. *Nucleic Acids Res* **38**, e87, doi:10.1093/nar/gkp1163
543   (2010).
544   15   Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil-DNA-
545   glycosylase treatment for screening of ancient DNA. *Philosophical transactions of the*
546   *Royal Society of London. Series B, Biological sciences* **370**, 20130624,
547   doi:10.1098/rstb.2013.0624 (2015).
548   16   Gansauge, M.-T., Aximu-Petri, A., Nagel, S. & Meyer, M. Manual and automated
549   preparation of single-stranded DNA libraries for the sequencing of DNA from ancient
550   biological remains and other sources of highly degraded DNA. *Nature Protocols* **15**,
551   2279-2300, doi:10.1038/s41596-020-0338-0 (2020).
552   17   Gansauge, M. T. *et al.* Single-stranded DNA library preparation from highly degraded
553   DNA using T4 DNA ligase. *Nucleic Acids Res* **45**, e79, doi:10.1093/nar/gkx033
554   (2017).
555   18   Fu, Q. *et al.* An early modern human from Romania with a recent Neanderthal
556   ancestor. *Nature* **524**, 216-219, doi:10.1038/nature14558 (2015).
557   19   Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians.
558   *Nature* **528**, 499-503, doi:10.1038/nature16152 (2015).
559   20   Fu, Q. *et al.* DNA analysis of an early modern human from Tianyuan Cave, China.
560   *Proc Natl Acad Sci U S A* **110**, 2223-2227, doi:10.1073/pnas.1221359110 (2013).

561   21   Fu, Q. *et al.* A revised timescale for human evolution based on ancient mitochondrial
562   genomes. *Curr Biol* **23**, 553-559, doi:10.1016/j.cub.2013.02.044 (2013).
563   22   Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of Next
564   Generation Sequencing Data. *BMC Bioinformatics* **15**, 356, doi:10.1186/s12859-014-
565   0356-4 (2014).
566   23   Kennett, D. J. *et al.* Archaeogenomic evidence reveals prehistoric matrilineal
567   dynasty. *Nat Commun* **8**, 14115, doi:10.1038/ncomms14115 (2017).
568   24   Reich, D., Thangaraj, K., Patterson, N., Price, A. L. & Singh, L. Reconstructing Indian
569   population history. *Nature* **461**, 489-494, doi:10.1038/nature08365 (2009).
570   25   Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065-1093,
571   doi:10.1534/genetics.112.145037 (2012).
572   26   Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for
573   present-day Europeans. *Nature* **513**, 409-413, doi:10.1038/nature13673 (2014).
574   27   Evans, J. A., Chenery, C. A. & Montgomery, J. A summary of strontium and oxygen
575   isotope variation in archaeological human tooth enamel excavated from Britain.
576   *Journal of Analytical Atomic Spectrometry* **27**, 754-764, doi:10.1039/C2JA10362A
577   (2012).
578   28   Fitzpatrick, A. P. *The Amesbury Archer and the Boscombe Bowmen: Early Bell
579   Beaker burials at Boscombe Down, Amesbury, Wiltshire, Great Britain: Excavations
580   at Boscombe Down*. Vol. 1 (Wessex Archaeology, 2011).
581   29   Millard, A. R. in *Cliffs End Farm, Isle of Thanet, Kent: A Mortuary and Ritual Site of
582   the Bronze Age, Iron Age and Anglo-Saxon Period With Evidence for Long-Distance
583   Maritime Mobility*   (eds J. I. McKinley *et al.*) 135-146 (Wessex Archaeology, 2014).
584   30   Champion, T. C., Haselgrove, C., Armit, I., Creighton, J. & Gwilt, A. *Understanding
585   the British Iron Age: An Agenda for Action. A Report for the Iron Age Research
586   Seminar and the Council of the Prehistoric Society.* (Trust for Wessex Archaeology,
587   2001).
588   31   Ringbauer, H., Novembre, J. & Steinrücken, M. Human Parental Relatedness
589   through Time - Detecting Runs of Homozygosity in Ancient DNA. *bioRxiv*,
590   2020.2005.2031.126912, doi:10.1101/2020.05.31.126912 (2020).
591   32   Fernandes, D. M. *et al.* A genetic history of the pre-contact Caribbean. *Nature* **590**,
592   103-110, doi:10.1038/s41586-020-03053-2 (2021).
593   33   Fernandes, D. M. *et al.* The spread of steppe and Iranian-related ancestry in the
594   islands of the western Mediterranean. *Nat Ecol Evol* **4**, 334-345, doi:10.1038/s41559-
595   020-1102-0 (2020).
596   34   Marcus, J. H. *et al.* Genetic history from the Middle Neolithic to present on the
597   Mediterranean island of Sardinia. *Nat Commun* **11**, 939, doi:10.1038/s41467-020-
598   14523-6 (2020).
599   35   Olalde, I. *et al.* The genomic history of the Iberian Peninsula over the past 8000
600   years. *Science* **363**, 1230-1234, doi:10.1126/science.aav4040 (2019).
601   36   Busby, G. B. J. *et al.* The peopling of Europe and the cautionary tale of Y
602   chromosome lineage R-M269. *Proceedings of the Royal Society B: Biological
603   Sciences* **279**, 884-892, doi:10.1098/rspb.2011.1044 (2012).
604   37   Rubinacci, S., Ribeiro, D. M., Hofmeister, R. & Delaneau, O. Efficient phasing and
605   imputation of low-coverage sequencing data using large reference panels. *bioRxiv*,
606   2020.2004.2014.040329, doi:10.1101/2020.04.14.040329 (2020).
607   38   Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74,
608   doi:10.1038/nature15393 (2015).
609   39   Hubacek, J. A. *et al.* Frequency of adult type-associated lactase persistence LCT-
610   13910C/T genotypes in the Czech/Slav and Czech Roma/Gypsy populations. *Genet
611   Mol Biol* **40**, 450-452, doi:10.1590/1678-4685-GMB-2016-0071 (2017).
612   40   Stead, I. M. *The Arras Culture*. (Yorkshire Philosophical Society, 1979).
613   41   Halkon, P. *The Arras Culture of Eastern Yorkshire: Celebrating the Iron Age*.
614   (Oxbow Books, 2020).

615 42 Jay, M., Montgomery, J., Nehlich, O., Towers, J. & Evans, J. British Iron Age chariot
616 burials of the Arras culture: a multi-isotope approach to investigating mobility levels
617 and subsistence practices. *World Archaeology* **45**, 473-491,
618 doi:10.1080/00438243.2013.820647 (2013).
619 43 Cunliffe, B. *Britain Begins*. (Oxford University Press, 2013).
620 44 Koch, J. T., Cunliffe, B.W. *Celtic from the West 2: Rethinking the Bronze Age and the*
621 *Arrival of Indo-European in Atlantic Europe*. (Oxbow Books, 2013).
622 45 Needham, S. & Bowman, S. Flesh-hooks, technological complexity and the Atlantic
623 Bronze Age feasting complex. *European Journal of Archaeology* **8**, 93-136,
624 doi:10.1177/1461957105066936 (2005).
625 46 Marcigny, C., Bourgeois, J. & Talon, M. in *Rythmes et contours de la geographie*
626 *culturelle sur le littoral de la Manche entre le IIIe et le debut du Ier millenaire (chapter*
627 *title: Movement, exchange and identity in Europe in the 2nd and 1st millennia BC:*
628 *beyond frontiers* (eds A. Lehoerff & M. Talon) 63-78 (Oxbow Books, 2017).
629 47 Marcigny, C. in *Les Anglais in Normandie (chapter title: Les relations transmanche*
630 *durant l'age du Bronze entre 2300 et 800 avant notre ère)* 47–54 (2011).
631 48 McConvell, P. in *Migration History in World History (chapter title: The archaeo-*
632 *linguistics of migration)* 153-186 (Brill, 2010).
633 49 Bellwood, P. S. & Renfrew, C. *Examining the Farming/Language Dispersal*
634 *Hypothesis*. (McDonald Institute for Archaeological Research, 2002).
635 50 Sims-Williams, P. An Alternative to 'Celtic from the East' and 'Celtic from the West'.
636 *Cambridge Archaeological Journal* **30**, 511-529, doi:10.1017/S0959774320000098
637 (2020).
638 51 Mallory, J. P. *The Origins of the Irish*. (Thames & Hudson Inc., 2013).
639 52 Rodway, S. The Ogham inscriptions of Scotland and Brittonic Pictish. *Journal of*
640 *Celtic Linguistics* **21**, 173-234, doi:10.16922/jcl.21.6 (2020).
641 53 Herm, G. *The Celts: The people who came out of the darkness*. (Weidenfeld and
642 Nicolson, 1976).
643 54 Guggisberg, M. in *Oxford Handbook of the European Iron Age* (ed K. Rebay-
644 Sailsbury C. Haselgrove, P. Wells) (Oxford University Press, 2018).
645 55 Pinhasi, R., Fernandes, D. M., Sirak, K. & Cheronet, O. Isolating the human cochlea
646 to generate bone powder for ancient DNA analysis. *Nature Protocols* **14**, 1194-1205,
647 doi:10.1038/s41596-019-0137-7 (2019).
648 56 Sirak, K. A. *et al.* A minimally-invasive method for sampling human petrous bones
649 from the cranial base for ancient DNA analysis. *Biotechniques* **62**, 283-289,
650 doi:10.2144/000114558 (2017).
651 57 Dabney, J. *et al.* Complete mitochondrial genome sequence of a Middle Pleistocene
652 cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci U S A*
653 **110**, 15758-15763, doi:10.1073/pnas.1314445110 (2013).
654 58 Korlevic, P. *et al.* Reducing microbial and human contamination in DNA extractions
655 from ancient bones and teeth. *Biotechniques* **59**, 87-93, doi:10.2144/000114320
656 (2015).
657 59 Rohland, N., Glocke, I., Aximu-Petri, A. & Meyer, M. Extraction of highly degraded
658 DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nat*
659 *Protoc* **13**, 2447-2461, doi:10.1038/s41596-018-0050-5 (2018).
660 60 Briggs, A. & Heyn, P. in *Methods in Molecular Biology* Vol. 840 143-154 (Springer,
661 2012).
662 61 Maricic, T., Whitten, M. & Paabo, S. Multiplexed DNA sequence capture of
663 mitochondrial genomes using PCR products. *PLoS One* **5**, e14004,
664 doi:10.1371/journal.pone.0014004 (2010).
665 62 Behar, D. M. *et al.* A "Copernican" reassessment of the human mitochondrial DNA
666 tree from its root. *Am J Hum Genet* **90**, 675-684, doi:10.1016/j.ajhg.2012.03.002
667 (2012).
668 63 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler
669 transform. *Bioinformatics* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).

670 64 Weissensteiner, H. *et al.* HaploGrep 2: mitochondrial haplogroup classification in the
671 era of high-throughput sequencing. *Nucleic Acids Res* **44**, W58-63,
672 doi:10.1093/nar/gkw233 (2016).
673 65 Sawyer, S., Krause, J., Guschanski, K., Savolainen, V. & Paabo, S. Temporal
674 patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA.
675 *PLoS One* **7**, e34131, doi:10.1371/journal.pone.0034131 (2012).
676 66 Skoglund, P. *et al.* Separating endogenous ancient DNA from modern day
677 contamination in a Siberian Neandertal. *Proc Natl Acad Sci U S A* **111**, 2229-2234,
678 doi:10.1073/pnas.1318934111 (2014).
679 67 Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS*
680 *genetics* **2**, e190, doi:10.1371/journal.pgen.0020190 (2006).
681 68 Bronk Ramsey, C. Bayesian Analysis of Radiocarbon Dates. *Radiocarbon* **51**, 337-
682 360, doi:10.1017/S0033822200033865 (2009).
683 69 Reimer, P. J. *et al.* The IntCal20 northern hemisphere radiocarbon age calibration
684 curve (0–55 cal kBP). *Radiocarbon* **62**, 725-757, doi:10.1017/RDC.2020.41 (2020).
685