



HAL
open science

The peopling of Europe since the Upper Paleolithic: prehistoric encounters between locals and migrants

Oğuzhan Parasayan

► **To cite this version:**

Oğuzhan Parasayan. The peopling of Europe since the Upper Paleolithic: prehistoric encounters between locals and migrants. Genetics. Université Paris Cité, 2022. English. NNT : 2022UNIP5267 . tel-04860593

HAL Id: tel-04860593

<https://theses.hal.science/tel-04860593v1>

Submitted on 1 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Paris Cité

École doctorale Bio Sorbonne Paris Cité (562 BioSPC)
UMR 7592 - Institut Jacques Monod
Équipe Épigénome et Paléogénome

The peopling of Europe since the Upper Paleolithic: Prehistoric encounters between locals and migrants

Par Oğuzhan PARASAYAN

Thèse de doctorat de Génétique

Dirigée par Thierry GRANGE

Présentée et soutenue publiquement le 20 Décembre 2022

Devant un jury composé de :

Garrett HELLENTHAL	Professeur	University College London	Rapporteur
Etienne PATIN	CR-HDR	Institut Pasteur – CNRS	Rapporteur
Jonathan WEITZMAN	Professeur	Université Paris Cité	Président du jury
Diyendo MASSILANI	Professeur	Yale School of Medicine - Yale University	Examineur
Jean-Paul DEMOULE	Professeur Émérite	Université Paris 1 Panthéon – Sorbonne	Examineur
Thierry GRANGE	DR	Université Paris Cité - CNRS - Institut Jacques Monod	Directeur de thèse
Eva-Maria GEIGL	DR	Université Paris Cité - CNRS - Institut Jacques Monod	Membre invité

Abstract

The peopling of Europe since the Upper Paleolithic: Prehistoric encounters between locals and migrants

Patterns of genetic diversity in present-day human populations are reflecting the past demographic fluctuations, migrations and admixture in human (pre)history. Paleogenomics, the generation and analysis of genome-wide data from ancient degraded specimens, allows us to study these demographic events by directly accessing time series comprising their occurrence. European prehistory recurrently witnessed population expansions associated with various climatic changes and/or cultural transitions during which local populations encountered migrating populations resulting in admixture at both the population and individual levels. We focused on three key transition periods in human history, generated ancient genomes of contemporary witnesses and used various state-of-the-art population genomics methods to reconstruct human genome evolution.

First, we analyzed the genomes of two ~37,000-year-old individuals from Crimea and showed them to belong to the first migration wave of anatomically modern humans into Europe carrying ancestry found in present-day Europeans. This expansion corresponds to a period just following a major climatic change known as Heinrich 4 cold period comprising the Campanian Ignimbrite volcanic super-eruption that coincided with human population turnover in Europe and a cultural transition (Mid/Upper Paleolithic). We show the evidence of admixture between earlier inhabitants of Europe and both ancient Crimean individuals. We further reveal the previously unknown genetic links between these Early European individuals and later groups that are important for the understanding of the early peopling of Europe.

Second, we focused on major demographic events in Western Europe during the Holocene that shaped the present-day European genome. We generated genomic data from 116 individuals dating from the Late Mesolithic to the beginning

of the Bronze Age, between 8,200 and 4,200 years ago covering two large-scale populational turnover events in Europe.

We analyzed the genomes of Mesolithic groups in Europe and characterized the population history of a local group in Northwestern Europe. By investigating the earliest Neolithic population that arrived in the Paris Basin during the expansion of farming into Western Europe we show their complex admixture with these pre-existing Mesolithic groups leading to a highly heterogenous population structure. Our results highlight that Neolithic genomes from France are essential to understand the Neolithization of the West. Indeed, our genomic analyses identify in northern France a likely source population for Neolithic Britain and Ireland, which has been sought for by both geneticists and archeologists. We also identify hitherto undetected high hunter-gatherer ancestry levels in individuals from Middle Neolithic contexts that might be meaningful for the social structure of these communities.

The third transition period we studied took place during the 3rd millennium BCE, which was a period of both cultural and genomic transformations associated with migrations of populations from the Pontic-Caspian steppe. Our high-resolution analysis of whole genomes of a family burial in France dating to about 2,500 BCE allowed us to reveal the Europe-wide admixture dynamics during this eventful period. On the territory of present-day France, we observed admixture between local Neolithic farmers and steppe-ancestry carriers coming from the north. We discovered previously undetected admixture pulses coinciding with the cultural transformation events of the 3rd millennium BCE.

In conclusion, our detailed studies of ancient genomes covering a large time interval of human history with major transition periods adds valuable insights to the knowledge of the genetic history of human populations.

Key words: Paleogenomics, ancient DNA, evolutionary genomics, population genomics, peopling, admixture, Upper Paleolithic, Mesolithic, Neolithic, Bell Beaker, Bronze Age

Resumé

Le peuplement de l'Europe depuis le Paléolithique supérieur: des rencontres préhistoriques entre locaux et migrants

La diversité génétique des populations humaines actuelles reflète les fluctuations démographiques, migrations et métissages du passé. La paléogénomique, la production et l'analyse de données génomiques à partir de spécimens anciens dégradés, nous permet d'étudier ces événements démographiques grâce à l'accès direct à des séries temporelles. Au cours de la préhistoire européenne, plusieurs expansions populationnelles associées à des changements climatiques et des transitions culturelles ont conduit à des rencontres et métissages entre populations locales et migrantes. Nous avons ciblé des périodes de transition clé, avons généré des génomes anciens correspondants et utilisé des méthodes de génomique des populations pour reconstruire l'évolution des génomes humains.

Nous avons d'abord analysé les génomes de deux individus de Crimée vieux de ~37,000 ans et montré qu'ils appartenaient à la première vague de migration en Europe des Humains anatomiquement modernes qui étaient porteurs de l'ascendance qui a persisté dans les génomes des Européens actuels. Cette expansion correspond à une période suivant un changement climatique majeur, la période glaciaire Heinrich 4, et l'éruption ignimbrite campanienne du super-volcan Archiflegéen qui a coïncidé avec un renouvellement de populations en Europe et la transition culturelle Paléolithique moyen/supérieur. Nous avons mis en évidence le métissage entre les habitants européens préexistants et les deux individus de Crimée ainsi que des liens insoupçonnés entre les premiers *H. sapiens* européens et les groupes plus tardifs qui éclairent notre compréhension du premier peuplement d'Europe par *H. sapiens*.

Puis, nous avons étudié des événements démographiques majeurs en Europe de l'Ouest durant l'Holocène qui ont façonné le génome des Européens actuels. Nous avons produit des données génomiques à partir de 116 individus datés du Mésolithique

final jusqu'à l'âge du Bronze (6 200-2 200 ans av.n.e.) couvrant deux grands événements de renouvellement des populations en Europe.

Les génomes de groupes mésolithiques européens analysés nous ont permis de caractériser l'histoire populationnelle d'un groupe local en Europe du nord-ouest. Au travers de l'analyse de la population néolithique la plus ancienne arrivant dans le Bassin Parisien, nous montrons des métissages complexes avec les groupes mésolithiques préexistants conduisant à une structure de population très hétérogène. Nos résultats soulignent la position clé des génomes néolithiques du nord de la France pour la compréhension de la néolithisation de l'Ouest européen. En effet, nos analyses génomiques identifient dans le nord de la France une probable population source à l'origine de la néolithisation de la Grande-Bretagne et l'Irlande longtemps recherchée par les généticiens et les archéologues. Dans des individus du Néolithique moyen, nous avons identifié une ascendance chasseur-cueilleur élevée suggérant une structure sociale particulière de ces communautés.

La troisième période de transition analysée se situe au 3ème millénaire av.n.e, une période où des transformations culturelles et génomiques étaient associées avec des migrations de populations de la steppe pontique-caspienne. Notre analyse génomique à haute résolution d'une sépulture familiale en France datant de 2 500 av.n.e. nous a permis de mettre en évidence des dynamiques de métissage ayant eu lieu partout en Europe. Nous avons observé que le métissage au Nord de la France s'est produit entre des agriculteurs néolithiques du Sud et des porteurs de l'ascendance des steppes venant du Nord. Nous avons découvert que ce métissage s'est effectué en deux vagues qui coïncide chacune avec des transformations culturelles au 3ème millénaire av.n.e.

En conclusion, nos études détaillées de génomes anciens couvrant un intervalle de temps large de l'histoire humaine éclairent l'histoire génétique des populations humaines à des périodes clés de transition.

Mots clés: Paléogénomique, ADN ancien, génomique évolutive, génomique des populations, peuplement, métissage, Paléolithique supérieur, Mésolithique, Néolithique, Campaniforme, Age du Bronze

Acknowledgements

I would like to briefly thank everyone who directly or indirectly contributed in this thesis work and in the production of this manuscript

First, I would like to express my deepest gratitude to my supervisors, Eva-Maria Geigl and Thierry Grange, who have guided me both professionally and personally since the very first day. I thank them for constantly sharing my excitement and enthusiasm and for giving me the opportunity to be creative in my research. I also thank them for creating a family atmosphere in the lab, which made this intense work so enjoyable for all of us.

Secondly, I would like to thank all my colleagues in the Epigenomics & Paleogenomics team, with whom I shared many beautiful memories inside and outside the lab. We supported each other at every condition. Very briefly, many thanks to Jeanne for sharing with me her wonderful and kind personality and precious friendship that I can always count on, to Wejdene for always having a warm place in her heart and making me feel like a brother, to Caitlin for sharing her bright personality and always bringing intriguing discussions, to Fabien for showing his knowledge and unending kindness to us.

Then, I would like to thank all the archaeologist and anthropologist colleagues without whom none of these projects would be possible, especially Mike, Jérôme, Corinne and Lamys. for sharing their valuable knowledge with us and allowing me to have an incredible excavation experience that brought me deeply closer to our research.

I would also like to thank Etienne Patin and Garrett Hellenthal for agreeing to be the referees of this thesis and for their valuable discussions that helped improve this manuscript and my research.

Finally, I deeply thank my family and friends from all over the world, who never got tired of listening to me, especially my parents, my brother and my cousin Gökçe, who never stopped believing in me since the beginning, and Jean-Marie, who has always been my greatest support at every stage of this thesis with unlimited patience. Many thanks to my lovely Parisian friends, who have remained as my biggest fans, İris, Lara, Melis, Zeynep B. Zeynep S. and of course Milo. Many thanks to my Magistère people, Vale, Sara, Lisa, Andrea, Berthrand, Guillem, with whom I started this incredible journey. Many thanks to my beloved friends Yağmur, Merve and Özge as well as Beliz and Emrecan for their never-ending lovely support.

Table of Contents

LIST OF ABBREVIATIONS	0
RÉSUMÉ DÉTAILLÉ	1
INTRODUCTION	8
PALEOGENOMICS	8
PREHISTORICAL ENCOUNTERS	11
EURASIAN EXPANSIONS OF THE AMHS	11
EARLY HOLOCENE EXPANSIONS	15
The Pleistocene/Holocene Transition	15
The origins of the Neolithic	17
The Neolithic expansion into Europe	18
The Early Neolithic West of the Rhine river	21
The Late LBK in the Aisne Valley	22
The Mediterranean route of Neolithization	23
Spread of farming into Scandinavia and the British Isles	25
POPULATION DYNAMICS OF THE 3RD MILLENNIUM	27
RATIONALE OF THE THESIS	30
REFERENCES	32
CHAPTER I	39
EARLY UPPER PALEOLITHIC GENOMES OF CRIMEA SHOW MIGRATION AND ADMIXTURE DYNAMICS OF THE FIRST MODERN EUROPEAN ANCESTRIES IN RELATION TO CLIMATIC CRISIS	39
ABSTRACT	40
INTRODUCTION	40
RESULTS	42
DISCUSSION	49
METHODS	51
REFERENCES	57
FIGURES	62
SUPPLEMENTARY INFORMATION	66

CHAPTER II	92
<hr/>	
GENOMICS OF THE EUROPEAN NEOLITHIZATION PROCESS PARALLELED BY FINE-SCALE POPULATION DIFFERENTIATION	92
RESULTS & DISCUSSION	93
MATERIAL & METHODS	118
SUPPLEMENTARY MATERIALS	125
REFERENCES	147
CHAPTER III	151
<hr/>	
LATE NEOLITHIC FAMILY BURIAL REVEALS ADMIXTURE DYNAMICS DURING THE THIRD MILLENNIUM BCE AND THE SHAPING OF THE EUROPEAN GENOME	151
ABSTRACT	152
INTRODUCTION	152
RESULTS	159
DISCUSSION	168
CONCLUSION	175
MATERIAL & METHODS	176
FIGURES	182
REFERENCES	186
SUPPLEMENTARY MATERIALS	190
GENERAL CONCLUSION	226
<hr/>	

List of Abbreviations

ALP	Alföld Linear Pottery
AMHs	Anatomically Modern Humans
ANF	Anatolian Neolithic Farmers
BCE	Before Common Era
BP	Before Present
BVSG	Blicquy/Villeneuve-Saint-Germain
cal	calibrated
CHG	Caucasian Hunter-Gatherers
cM	centiMorgan
EEF	Early European Farmers
EHG	Eastern Hunter-Gatherers
ICC	Impressa-Cardial Complex
IUP	Initial Upper Paleolithic
kya	kilo years ago
LBK	Linearbandkeramik
LBKT	LBK in Transdanubia
LGM	Last Glacial Maximum
MDS	Multi-Dimensional Scaling
PCA	Principal Component Analysis
ROH	Runs of Homozygosity
RRBP	Rubané Récent du Bassin Parisien
SNP	Single Nucleotide Polymorphism
SWA	South West Asia

Résumé Détaillé

L'extraction d'ADN et la génération de données génomiques à partir de spécimens anciens ont considérablement transformé notre compréhension de l'histoire génétique des populations humaines. La paléogénomique, c'est-à-dire la génération et l'analyse de données génomiques à partir de spécimens anciens dégradés, nous permet de reconstruire l'histoire génétique des populations humaines en accédant directement à des séries chronologiques depuis le Paléolithique supérieur. Les archives archéologiques en Europe laissent entrevoir des rencontres préhistoriques récurrentes entre populations locales et migrantes, même si de nombreuses questions demeurent. Les populations sont censées migrer en raison de changements climatiques et environnementaux, de la pression démographique et/ou de développements culturels. Les mouvements de populations seraient accompagnés d'événements de brassage local ou de remplacement de population. Ces événements auraient laissé des traces dans les génomes des descendants qui pourraient être analysées par une approche paléogénomique.

Grâce aux développements méthodologiques de ces dernières années, tant en biologie moléculaire qu'en bioinformatique, la contribution de la paléogénomique aux questions fondamentales de la biologie évolutive et de l'histoire humaine croît de façon spectaculaire. Cependant, les processus de peuplement associés aux événements mentionnés ci-dessus, et en particulier ceux liés aux transitions des cultures archéologiques, sont encore mal compris pour la plupart des périodes et des zones géographiques. Pour éclairer ces points, nous nous sommes concentrés sur trois périodes de transition clés de l'histoire de l'humanité pour lesquelles nous avons généré des génomes anciens de témoins contemporains et utilisé une diversité de méthodes avancées de génomique des populations pour reconstruire l'évolution du génome humain en visant la compréhension au niveau génétique du mélange entre les populations.

Tout d'abord, il y a environ 38 000 ans, après un refroidissement climatique majeur, l'événement de Heinrich 4, et une éruption volcanique dramatique (la super-éruption de l'ignimbrite Campanienne (CI)) affectant le sud et l'est de l'Europe a conduit à une crise de l'écosystème. À la suite de cette crise, l'Europe a connu un renouvellement de sa population

humaine, qui a introduit pour la première fois une ascendance européenne ayant persisté jusqu'à nos jours. Nous avons produit des séquences génomiques à faible couverture à partir de fragments crâniens, dans lesquels l'ADN est très mal conservé, de deux individus européens du Paléolithique supérieur précoce âgés de 36 200 et 36 800 ans, provenant du site de l'abri sous roche de Buran-Kaya III sur la péninsule de Crimée. Ces individus étaient des témoins de la dynamique des populations post-CI. Pour maximiser la quantité de données comparables entre ces individus et d'autres génomes paléolithiques, nous avons exploité le panel SNP du Human Genome Diversity Project (HGDP). L'analyse combinée de ces génomes avec d'autres données génomiques anciennes en Europe publiées récemment nous a permis de créer un modèle large et actualisé des mouvements de population, des interactions et des remplacements pendant les premières vagues de peuplement de l'Europe par *Homo sapiens* au début du Paléolithique supérieur. Les deux individus de Crimée se sont révélés plus similaires aux populations européennes ultérieures qu'aux premiers habitants de l'Europe associés au Paléolithique supérieur initial (PSI). L'individu BuKa3C, plus ancien, était peu mélangé avec une ascendance similaire à celle trouvée dans le groupe IUP du site de Bacho Kiro et portait un haplogroupe Y associé à l'IUP, alors que l'individu BuKa3A, plus récent, était plus fortement mélangé avec les populations pré-CI. Tous deux partageaient plus d'ascendance avec des individus vivant des milliers d'années plus tard en Europe occidentale et centrale qu'avec tout autre groupe contemporain découvert à ce jour. Puisque les individus de Buran Kaya III portaient une ascendance européenne associée à l'IUP, nous pouvons conclure qu'une partie de cette ascendance a persisté après la crise écologique durant le sévère stade de Heinrich 4 et la super-éruption de l'IC qui a dévasté le sud et l'est de l'Europe au début de cet événement Heinrich 4. Ainsi, des mélanges ont dû se produire entre les populations locales et les premières populations migrantes, au cours desquels les populations migrantes semblent avoir été plus nombreuses que les populations locales, un phénomène que l'on retrouvera avec les expansions à l'Holocène ultérieures. Ce modèle met en évidence que ces différentes populations qui ont successivement colonisé l'Europe ont divergé à partir d'une source similaire, tout en se mélangeant à différents degrés avec des populations AMH pré-CI dont la divergence était plus précoce, et qui sont arrivées plus tôt. Ces interactions entre groupes voisins ont formé les premières ascendances détectables par la suite chez les Européens actuels. En outre, nous avons constaté que les habitants de Buran-Kaya III partageaient le plus de similitudes génomiques avec les groupes du Paléolithique supérieur moyen trouvés en Europe centrale et occidentale plusieurs milliers d'années plus tard, mais aussi avec les groupes du Paléolithique supérieur tardif du Caucase. Cette similitude génétique, associée à un horizon archéologique similaire à celui trouvé sur

plusieurs sites contemporains du Caucase du Nord, suggère que le Caucase pourrait être le point d'entrée de l'ascendance liée à Buran-Kaya III. Cette ascendance est mieux modélisée comme se ramifiant à partir d'une population source menant à une ascendance de type de celle des individus trouvés sur le site de Kostenki après s'être séparée des lignées qui donneraient naissance aux populations représentées par les individus GoyetQ116-1 et BK_1653, trouvées plus tard en Europe centrale et occidentale, et Sunghir3 en Orient. Nous observons également que les seuls groupes qui ne portaient pas d'ascendance européenne pré-CI, Sunghir3 et Kostenki14, datent d'avant 34 000 ans et ne se trouvent qu'en Europe de l'Est, ce qui indique un point d'entrée oriental possible pour ces lignées à partir d'une zone qui pouvait être dépeuplée ou peu habitée par les populations européennes pré-CI à cette époque.

La deuxième période sur laquelle nous nous sommes concentrés est la transition entre le Mésolithique et le Néolithique, consécutive aux expansions de populations en provenance d'Asie du Sud-Ouest qui ont entraîné la diffusion de l'agriculture en Europe au cours de l'Holocène. Pour ce faire, nous avons généré et analysé les données génomiques de 109 individus anciens provenant de 20 sites du Bassin parisien au sens large, comprenant les régions actuelles de Picardie, Champagne et Nord-Pas-de-Calais au nord de la France. Les sites correspondant sont associés à des cultures néolithiques couvrant une période allant de 5100 à 2500 avant notre ère, à l'exception d'un individu mésolithique datant d'environ 6200 avant notre ère. En générant des données génomiques à partir de la plus ancienne population d'agriculteurs arrivée dans le Bassin parisien, dernière étape de l'expansion du Néolithique précoce, nous avons pu discriminer les génomes des populations se déplaçant le long des routes migratoires continentales et méditerranéennes bien qu'elles aient divergé d'un ancêtre commun dans le sud-est de l'Europe environ mille ans auparavant. La néolithisation précoce de l'Europe met en évidence l'émergence d'un processus continu de différenciation progressive du génome des premiers agriculteurs. Cette différenciation est devenue plus apparente dans les populations ultérieures lorsque l'expansion précoce de l'agriculture a pris fin et que la taille des populations des différents groupes a augmenté. La détection de telles structures préhistoriques à haute résolution en Europe nous a permis de localiser les origines possibles des individus ou des groupes dont les ancêtres se déplaçaient et de révéler des interactions qui ne pouvaient pas être détectées par les seules données archéologiques. Nos estimations des dates de métissage pour les individus du Néolithique ancien ont détecté différentes phases principales de mélange des populations se déplaçant par la route de l'Europe centrale associée à la céramique rubanée, dite

LBK, et par la route de la Méditerranée associée à la céramique cardiale, dite ICC. Les individus associés à l'ICC résultent de deux vagues de mélange, l'une correspondant à l'arrivée de cette vague en Europe du sud vers 5900 avant notre ère, et une seconde vague centrée autour de 5500 avant notre ère. Les individus LBK ont également connu deux grandes vagues de mélange, mais à des moments distincts, l'une centrée sur 5600 avant notre ère, la seconde sur 5400 avant notre ère. En utilisant des méthodes basées sur l'inférence d'ascendance locale, nous montrons que la principale source de différenciation entre les agriculteurs néolithiques d'Europe occidentale ne sont pas les différentes ascendances des chasseurs-cueilleurs occidentaux. Lorsque cette approche a été appliquée aux individus LBK du Bassin parisien, nous avons détecté diverses affinités d'ascendance chasseurs-cueilleurs. Notre détection de différences dans la chronologie des principales vagues de mélange avec les populations locales de chasseurs-cueilleurs, ainsi que dans les sources de ces ascendances chasseurs-cueilleurs chez les agriculteurs néolithiques migrant en Europe le long des routes continentales et méditerranéennes indique des différences géographiques et temporelles dans les interactions entre les agriculteurs néolithiques migrants et les chasseurs-cueilleurs mésolithiques locaux. Ces différences ont contribué à la formation phylogéographique du génome européen précoce, même si la diversité de ces modèles de mélange ne semble pas être le seul générateur de différenciation génétique. En effet, la dérive génétique au cours des migrations jouerait aussi un rôle important dans cette différenciation qui, dans de nombreux cas, ne peut être simplement corrélée à l'ascendance chasseur-cueilleur.

D'autre part, en analysant les génomes d'individus de la fin du Néolithique moyen ayant une ascendance chasseur-cueilleur très élevée, nous avons détecté que les ancêtres de ces individus étaient des chasseurs-cueilleurs qui ont reçu un flux génétique récent des groupes néolithiques tout en conservant leur ascendance chasseur-cueilleur. Cela signifie qu'ils provenaient d'une population qui est restée longtemps à l'écart des migrants néolithiques et qu'ils n'ont pas reçu ensuite de flux génétique supplémentaire provenant d'individus néolithiques, c'est-à-dire qu'ils se sont plutôt accouplés au sein de leurs groupes. Les individus et les groupes ayant une forte ascendance de chasseur-cueilleur n'étaient pas identifiables archéologiquement mais semblent avoir été intégrés dans les villages néolithiques, bien qu'ils étaient enterrés ensemble. Les résultats des datations de mélanges indiquent que le mélange entre les chasseurs-cueilleurs et les agriculteurs néolithiques a eu lieu dans la première moitié du 4^{ème} millénaire avant notre ère, ce qui confère une importance particulière à cette période pour les études futures. Nous avons analysé ensuite une population du site du Néolithique

moyen d'Escalles "Mont d'Hubert", Pas-de-Calais, la région des Flandres du Nord de la France près du "Cap blanc Nez" au sud-ouest de Calais, datée de 4230-3879 avant notre ère. Notre analyse a montré que cette population forme un groupe distinct très liée aux populations néolithiques des Îles britanniques plutôt qu'aux autres génomes néolithiques. Leur proximité génétique suggère que la population du Néolithique moyen de la région des Flandres est la population source des agriculteurs néolithiques qui ont colonisé l'Angleterre. Nous avons également analysé les sépultures de deux hypogées à Chouilly " la Grifaine " dans la vallée de la Marne au sud de Reims, datées entre 3500-2900 et 3400-2500 avant J.-C., c'est-à-dire au Néolithique récent. Le cimetière qui comprend ces hypogées et d'autres a été utilisé pendant ~900 ans. Les génomes de l'hypogée F10, légèrement plus ancien, dont les sépultures couvrent la période entre 3500-2900 avant J.-C., se regroupent avec les génomes néolithiques du nord de la France, tandis que les génomes de l'hypogée F11, légèrement plus récente, dont les sépultures sont datées entre 3400-2500 avant J.-C., se regroupent avec les génomes néolithiques du sud de la France et de la péninsule ibérique. Cela suggère que deux populations différentes ont été inhumées à cet endroit, la plus récente ayant une origine différente. Ceci indique que ces populations avaient une identité différente influençant le choix des hypogées utilisées pour enterrer leurs morts. Aucun lien de parenté ayant été détectés entre les individus analysés, nous pouvons inférer que cette identité n'était pas particulièrement liée à l'appartenance à une même famille.

Dans un troisième volet de ma thèse nous avons analysé la période pendant laquelle l'ascendance steppique apparentée aux populations kourganes ou Yamnaya a été introduite en Europe occidentale. Cette ascendance constitue la troisième composante majeure du génome européen actuel et son apparition est associée à l'émergence de nouvelles cultures qui paraissent résulter des rencontres entre populations associées à des ascendances différentes. Nous avons généré des données génomiques à partir des os pétreux de chacun des individus d'une sépulture du Néolithique récent de " Bréviandes les Pointes et les Grèvottes " près de Troyes dans le sud du Bassin parisien avec une couverture de 0,75X à 4,6 X (couverture médiane 1X) et d'un individu de Saint-Martin-la-Garenne (SMGB54), à l'ouest de Paris dans les Yvelines, daté par radiocarbone à 2410-2129 cal BCE (97,8%) et enterré selon les rituels funéraires de la culture campaniforme. Grâce à une analyse à haute résolution, à la reconstitution des phases haplotypiques des génomes et à la reconstruction du génome d'un individu non échantillonné à travers les génomes de ses parents, nous avons mis en lumière les mécanismes et la direction

des processus de mélange en Europe occidentale entre les agriculteurs du Néolithique tardif et les descendants des nomades de la steppe pontique-caspienne. Nous avons identifié deux principales pulsations de mélange entre ascendance néolithique et steppiques. La première datant de l'arrivée des populations de la steppe en Europe de l'Est vers 2900 avant notre ère tandis que la seconde date de leur progression en Europe occidentale vers 2500 avant notre ère. Ces pulsations révèlent des dynamiques d'expansion différentes à travers l'Europe, avec des liens frappants avec des phénomènes culturels différents, à savoir la culture de la céramique cordée et celle du campaniforme. Ce résultat met en évidence les principes fondamentaux du processus de mélange entre les groupes néolithiques locaux et les groupes de migrants d'ascendance steppique avec lesquels l'ascendance steppique s'est introduite dans le pool génétique néolithique européen. La première phase de mélange intense s'est produite à l'arrivée en Europe orientale des migrants originaires des steppes au cours de laquelle on observe l'émergence d'une nouvelle culture, celle de la céramique cordée, qui ne semblait pas pratiquée sous cette forme par aucune des deux populations initiales. Cette population mélangée s'est ensuite déplacée vers l'Ouest tout en se mélangeant à faible fréquence avec les chasseurs cueilleurs rencontrés jusqu'à ce qu'ils arrivent sur un territoire chevauchant le Nord de la France où une deuxième phase de mélange avec les populations néolithiques se produit. Bien que cet événement de mélange se produise au Nord de la France, les populations néolithiques concernées semblent provenir principalement du Sud. On observe cette configuration aussi bien dans la tombe de Bréviandes, contemporaine à ce moment de mélange, pris ainsi sur le fait, mais aussi dans les individus campaniformes plus tardifs, comme celui de Saint-Martin-la-Garenne. Ainsi, cette deuxième phase de mélange est-elle associée à l'émergence d'une nouvelle culture, le campaniforme. Le campaniforme semble résulter de la combinaison de deux traditions, celle des tombes individuelles caractéristiques du Nord de l'Europe et des populations d'ascendance steppique (associées aux cultures « All Over Corded » et « All Over Ornamented »), et celles des gobelets maritimes associées aux populations néolithiques du Sud-Ouest de l'Europe, et donc caractéristiques des deux ascendances génétiques identifiées ici.

En conclusion, au cours de cette thèse, nous avons ciblés différents objets d'étude correspondants à des phases de mélange entre populations d'origine différentes en utilisant des méthodes avancées adaptées à la rareté des restes humains analysés, au degré modéré de préservation de l'ADN, à la quantité et à la qualité des données publiées pour les analyses comparatives. Nous avons échantillonné des individus, témoins contemporains de ces périodes

clés mentionnées ci-dessus, et généré des génomes anciens pour pouvoir aborder des questions scientifiques précises liées à ces mélanges de population. Nous avons mis en évidence des processus complexes de flux génétiques en analysant les schémas spatio-temporels des événements de mélange entre les populations rencontrées lors des périodes de transition du Paléolithique supérieur, du Mésolithique au Néolithique et du Néolithique tardif à l'âge du Bronze, la plupart du temps grâce à la résolution fine permise par l'utilisation de blocs haplotypiques phasés.

Introduction

Paleogenomics

The generation of genome-wide data from ancient specimens with spatiotemporal diversity using high-throughput sequencing technologies gave rise to ancient genomics as a revolutionary field in the last decade. Ancient DNA (aDNA) research has become a powerful tool providing better understanding and reconstruction of the past of different populations and species from the evolutionary biology and (pre) historical point of view. As a multidisciplinary field, the analysis of ancient DNA requires a collaboration between genomics, bioinformatics, archaeological sciences and other humanities allowing us to study many hypotheses that could not be tested by analyzing only genomes of present-day populations or archaeological remains using conventional archaeological and anthropological approaches. The phenomena that can be analyzed through paleogenomics include past population dynamics of humans and other extinct/extant species, archaic introgressions, migratory population movements and the associated spread of cultures, animal and plant domestication, past epidemics and their responsible pathogens etc.

However, there are some challenges that might prevent the recovery of DNA and the production of genomic data from ancient calcified or mummified remains, which can be both technological and ethical. The former includes (i) the extremely low quantity of endogenous DNA preserved in ancient remains, (ii) its dilution in environmental, mostly microbial DNA from the burial environment that constitute the vast majority of the recovered DNA (up to 100%), (iii) as well as postmortem DNA damage leading to very short DNA molecules (around 30-50 bp) and to modified bases. The small size of ancient DNA molecules is due to DNA fragmentation resulting first from postmortem enzymatic degradation, then from depurination through hydrolysis and β -elimination¹. The major DNA damage pattern is an elevated C to T-substitution frequency, in particular at the ends of the DNA fragments, due to deamination of cytosine to uracil resulting in incorrect incorporation of nucleotides during DNA replication

².These base misincorporations often represent a problem since the discrimination between true C/T or G/A heterozygous sites and deaminated sites becomes very challenging, especially in sequencing data with low coverage that are the consequence of low quantities of preserved endogenous DNA in ancient specimens ³.This impacts the research field at several levels: (i) well-preserved remains are scarce and therefore precious resulting in extreme competition between research groups and attempts to monopolize the collection of archaeological remains; (ii) destructive sampling of unique or precious remains prevent the reproduction of the data produced; (iii) restrictions to access archaeological specimens for analysis due to national regulations and religious beliefs; (iv) commercial conventions between some laboratories and biotechnology companies granting exclusive access to custom in-solution probe sets.

Despite these challenges, the available genomic data obtained from ancient specimens, generated via either whole-genome shotgun sequencing or in-solution hybridization capture methods, has dramatically increased since the last decade. Among this massive data production, some groundbreaking studies of ancient genomics has contributed significantly to the study of human evolution and the peopling of Eurasia and the Americas and shed light on the genetic diversity and ancestry defining present-day human populations. Archaic introgression events were detected from Neandertals and Denisovans, a previously unknown Late Pleistocene hominin group contemporary to the Neanderthals whose remains were found in Siberia and the Tibetan plateau, into anatomically modern humans (AMHs) in Eurasia. These introgression events led to regional variations in the genomes of present-day populations of both Neanderthal and Denisova ancestry. In particular, non-Africans were shown to harbor around 2% of Neanderthal ancestry and Oceanians up to 3% of Denisovan ancestry ⁴⁻⁹. Genomes of present-day Europeans were found to be the product of past admixture events during Holocene with varying proportions of three ancestral populations including indigenous Mesolithic hunter-gatherers, who preserved their ancestry since the Last Glacial Maximum (LGM), early Neolithic farmers, who introduced Aegean/Anatolian Neolithic ancestry into Europe starting ~8,500 years ago and steppe pastoralists introducing Yamnaya-like ancestry starting ~5,000 years ago (e.g., Allentoft et al., 2015; Fu et al., 2016; Haak et al., 2015; Lazaridis et al., 2014) ¹⁰⁻¹³. Furthermore, the analysis of our past migration and admixture history also allows us to identify the resulting complex phenotypic traits including immune response, metabolism, pigmentation, diet and behavior that can be under selection as well as the contributions coming from adaptive introgression and admixture in human history ^{5,14-18}. Overall, access to genomes from time series greatly improves our understanding of the evolution of human populations. Depending

on the sampling strategy and the quality of ancient genome data, one can reconstruct population movements, admixture, and the evolution of genetic traits and selected genomic regions. Today, we have a better understanding of demographic changes occurring over large geographical regions and wide time periods compared to what was previously known. Nevertheless, there are many gaps concerning the early dispersal of human populations as well as the more regional population history of later periods at higher resolution. Hence, this thesis will focus on the more recent human past including Eurasian expansions following the out-of-Africa event and population dynamics and admixture histories during the Holocene in Western Europe.

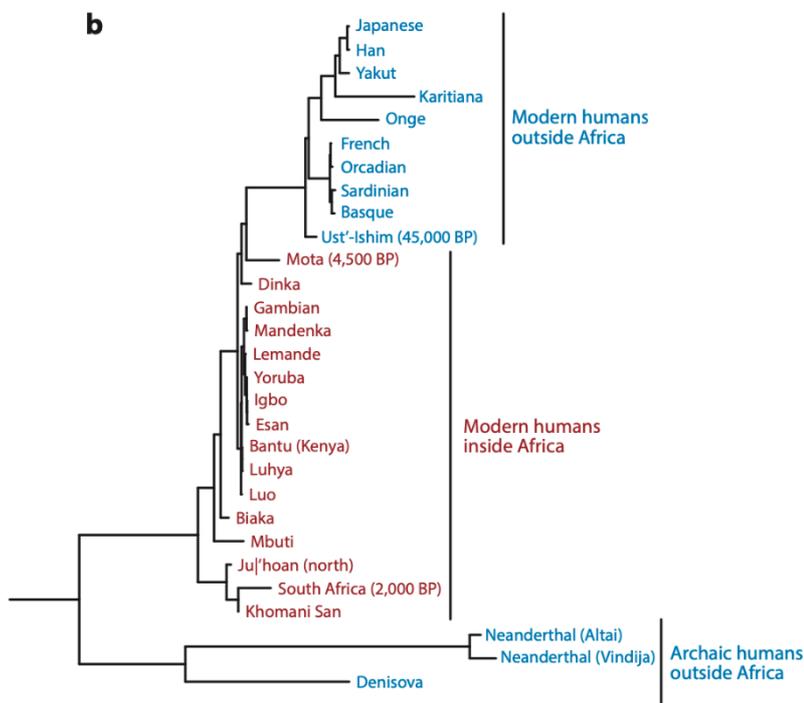


Figure 1: Maximum likelihood tree of present-day human populations from The Simons Genome Diversity Project and ancient anatomically modern and archaic humans (Skoglund and Mathieson 2018)

Prehistorical Encounters

Eurasian Expansions of the AMHs

The model for an African origin of all populations outside of Africa coming from a strong bottleneck during the out-of-Africa event is supported by the much higher genetic diversity detected in African populations than any other population outside of Africa (e.g., Tishkoff et al., 2009; Vigilant et al., 1991)^{20,21}. A further evidence comes from the fact that all non-African populations form a clade in relation to African populations while all world-wide present-day populations and ancient AMHs are more closely related to each other than to Neanderthal and Denisovan archaic humans (Fig. 1) (e.g., Mallick et al., 2016; Posth et al., 2016; Skoglund & Mathieson, 2018)^{7,19,22}. There are different lines of evidence indicating that the out-of-Africa event gave rise to world-wide expansions into Western and Eastern Eurasia and Oceania resulting in diversification of present-day non-African populations. Estimations of the effective population size and the cross-coalescent rates between African and non-African populations indicate a population divergence date varying between 50 to 160 kilo years ago (kya) depending on which reference population from Africa is used⁴. When taking the Yoruba as the African reference population estimations mostly indicate a date around 70 kya (Fig. 2A)^{4,7,23–25}, which might still show differences depending on which coalescent method is used²⁶. The presence of Neandertal-derived ancestry in both present-day and ancient non-African humans might be considered as the strongest evidence supporting the out-of-Africa hypothesis and a single origin of non-African lineages^{6,7,11}. Indeed, this phenomenon seems to be the result of a single episode of admixture outside of Africa since this ancestry is not detected in African populations, except some cases of later gene flow from Eurasia to Africa few millennia ago²⁷. This hypothesis got further support through several different estimations of the date of gene flow that indicates a time range between 50 – 60 kya independently of the level of these ancestry proportions (e.g., Bergström et al., 2020; Fu et al., 2015, 2016; Hajdinjak et al., 2021; Moorjani et al., 2016; Sankararaman et al., 2016).

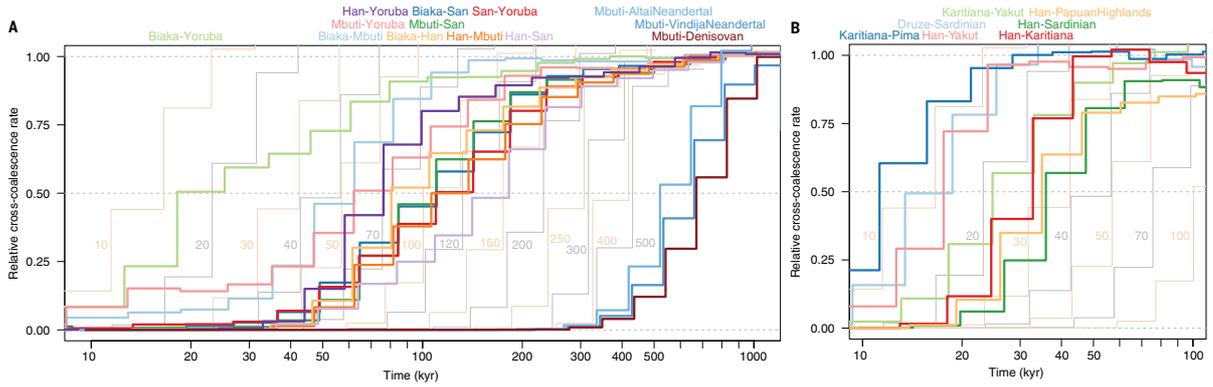


Figure 2: MSMC2 cross-population coalescent rates estimated from four physically phased haplotypes across non-African (Han) and African populations (A) and across different non-African populations (B) (Bergström et al., 2020).

All these estimations about the singular origin of non-African populations contribute to our understanding of the further divergence and diversification of early Eurasian lineages since it is expected that the population split in Eurasia occurred after the Neanderthal-derived gene flow. It is likely that this encounter happened somewhere in the Middle East, which could also be considered as a reasonable location for an expansion into different directions. While estimations from both coalescent-based and demographic modelling indicate a divergence between Western Eurasians, Eastern Eurasians, and Oceanian populations following out-of-Africa and Neanderthal gene flow around 35-50 kya, there are some discrepancies in the date estimations concerning the divergence between Oceanian and Eurasian lineages. While a recent cross coalescent rate estimates across Han – Sardinian and Han – Papuan Highlands genomes gives midpoint dates around 35 kya (Fig. 2B), different studies of coalescent-rate estimates as well as demographic modelling indicate an earlier divergence between Oceanians and the rest of the out-of-Africa metapopulation^{5,23,31}. Since the inferred dates for the first pulse of Denisovan-derived gene flow into Papuan-related populations is estimated to be around 46 kya, the model of an earlier split dates for Oceanians seems a more likely scenario. These discrepancies between estimated dates depending on the models used reveal that we should take all these date estimations with a grain of salt. It should be noted, however, that the estimates of the divergence time between the Western and Eastern Eurasian lineages are consistent with the dates and genetic profiles of the earliest AMHs.

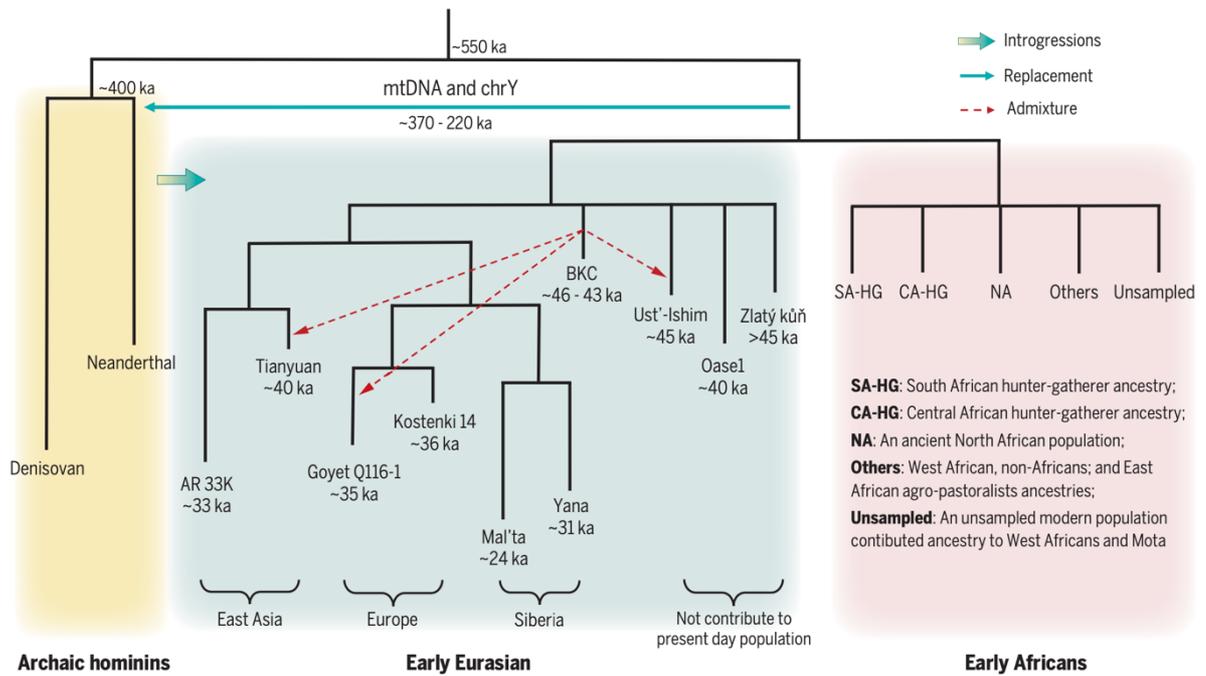


Figure 3: Illustration of the population history of archaic humans and early modern humans of interest (Liu et al., 2021).

The earliest Eurasian AMHs, Ust-Ishim from Siberia (Fu et al., 2014), early Bacho Kiro individuals from Bulgaria ²⁹, and Zlaty-kun from Czechia ³³, dating to ~45 kya, could be considered as the most basal group to more recent individuals representing populations in both western and eastern Eurasia. Indeed, these Initial Upper Paleolithic (IUP) specimens share similar genetic characteristics and they have apparently not contributed directly to the present-day populations. In contrast, individuals dating to later periods between 40 – 35 kya, Tianyuan from China ³⁴, Goyet-Q116-1 from Belgium ¹¹ and Kostenki 14 from western Russia ^{11,35} have contributed genetic ancestry that can be detected in present-day local populations (Fig. 3). which could be considered as the signatures of the diversification between Western and Eastern Eurasia. Surprisingly, if the geographic origins of Goyet Q116-1 and Tianyuan clearly identify them as representatives of the Western and Eastern expansion routes postdating the bifurcation of these expansions, they share some detectable drift despite their large geographical distance ¹¹. It was proposed that this shared genetic drift results from gene flow(s) from a population related to IUP Bacho Kiro individuals²⁹. If the IUP lineage was still present when the second wave arrived that gave rise to modern populations, the earlier populations could have admixed with the newcomers in different places, thus resulting in shared genetic drift between two populations that were separated previously. Nevertheless, there are still many aspects in the

population history of early AMHs that are unclear and require much more samples and analyses to fill the gaps in order to better understand the genetic diversity in the Upper Paleolithic.

What is better known so far is the population history around the LGM, a ~7,500-year-long period of extreme cold climatic conditions due to the formation of ice-sheets starting from 26,5 kya. During this period the northern parts of Eurasia and Northern America were inhabitable. These environmental conditions exerted a significant environmental constraint on human populations³⁶. Consequently, the genetic diversity of the populations following this period was strongly reduced compared to the relatively high heterozygosity levels detected in pre-LGM individuals demonstrating the occurrence of LGM bottlenecks³⁷. Such bottlenecks are also supported by demographic modelling of Mesolithic hunter-gatherers³⁸, as well as observed shifts in mitochondrial haplogroup frequencies during the LGM²². During the post-LGM period, different hunter-gatherer lineages appeared showing east-west and north-south genetic clines likely connected to post-LGM expansions of European hunter-gatherers that survived in southern refugia. Regional differences were detected in Scandinavia, Iberia, Northwestern Europe as well as in the Caucasus and in Southwest Asia (SWA) in general^{11,13,38-42}. Some of the earliest members of the post-LGM period, which are associated with the Magdalenian culture and dating back to 19-15 kya in Spain (El Miron) and Belgium (Goyet Q2), preserved the ancestry closely related to the pre-LGM Goyet Q116-1-like population that was detected in later Mesolithic individuals in Spain^{11,43}. Another group of the earliest post-LGM members includes individuals associated with the Epi-Gravettian culture dating back to 17-14 kya in Northern Italy. This group possesses a new ancestry that was undetected during the pre-LGM and found in most of the later Western European Mesolithic individuals⁴⁴. However, the genomic data dating back to the early post-LGM period are still scarce so far and the subsequent cold intervals during the following millennia must have had certain consequences for the surviving human populations that remain to be understood.

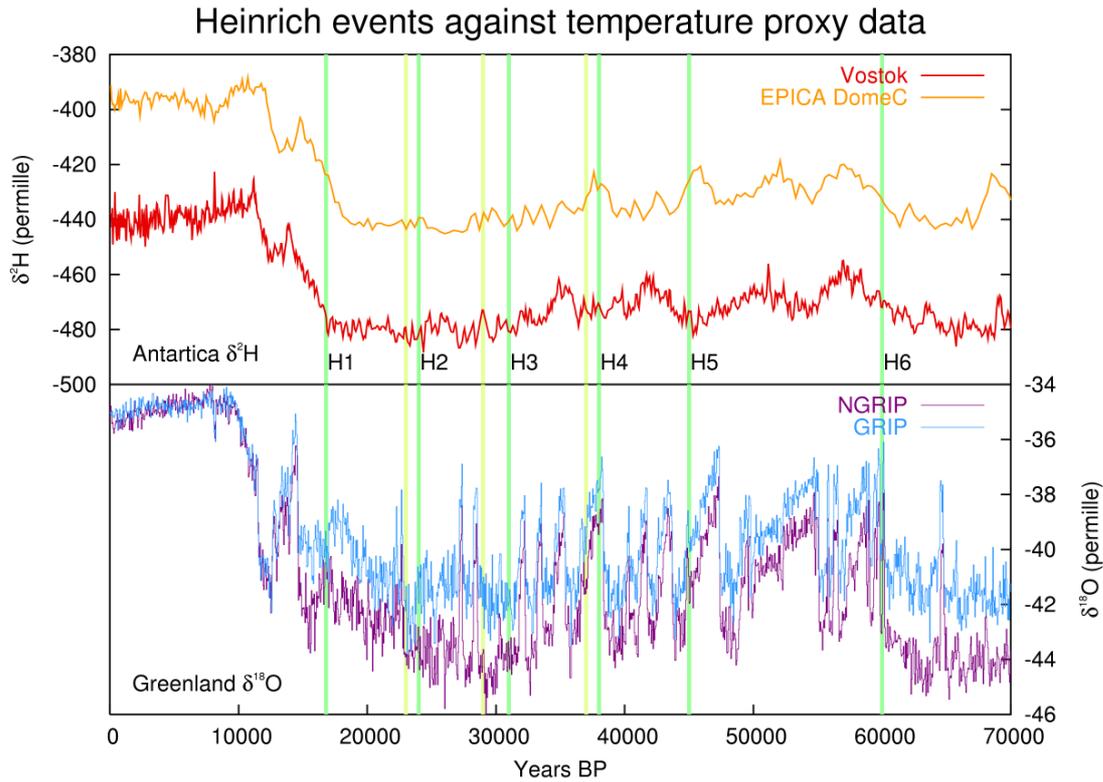


Figure 4: Reconstructed temperatures based on isotope measurements showing prehistoric Heinrich events (H1-6). Different colors indicate isotopic values per mil versus time (Hydrogen on top, Oxygen on bottom) coming from datasets obtained from different geographical locations (Heinrich event, The Azimuth Project)

Early Holocene Expansions

The Pleistocene/Holocene Transition

During the Pleistocene the climate was characterized by strong temperature oscillations leading to a sequence of alternating warm and cold periods. The latter became more severe with time. The Last Glacial Maximum, which was the most severe one, ended about 14,500 years ago with global warming. This post-glacial warming period was interrupted by a cold spill of 1,000 years, the Younger Dryas, that lasted from 12,700 to 11,700 BP. The Younger Dryas period was followed by a period of relatively stable and warmer climate lasting up to now, the Holocene (Fig.4). These climatic changes with their associated dramatic changes in environmental conditions must have had a major impact on human behavior. Indeed, the

environmental transition to the Holocene coincides in SWA with the Neolithic transition from a foraging lifestyle to a sedentary farming lifestyle, which was the beginning of the formation in SWA of complex societies that are the ancestors of our present-day societies.

The main hypothesis concerning the driving force to farming practices is the population growth model proposing that the change of the subsistence strategy occurred under population pressure in order to provide sufficient food to the larger number of people. This hypothesis finds support in the observation that life style changes had some negative consequences on the actors. Indeed, there is evidence that the health of agriculturalists was generally worse than the one of foragers since farming represented heavy physical labor and might have led to a less healthy, balanced diet ^{45,46}. It remains to be seen whether this transition was only caused by the changing and fluctuating environmental conditions at the end of the Late Pleistocene. It is conceivable that the idea of farming arose earlier since there is evidence of early plant exploitation dating back to 23 kya in the Levant that was not continued ^{47,48}. The idea itself, however, might have been preserved. Indeed, the Early Natufian Culture that is nearly contemporary with warm interval of the Bølling-Allerød phase lasted approximately from 14,500 to 13,000 BP (Fig.4) shows the earliest example of sedentary life style before the Holocene in the southern Levant ⁴⁹. This culture was, however, likely disturbed by the extremely arid period of the Younger Dryas

(Fig 4.)

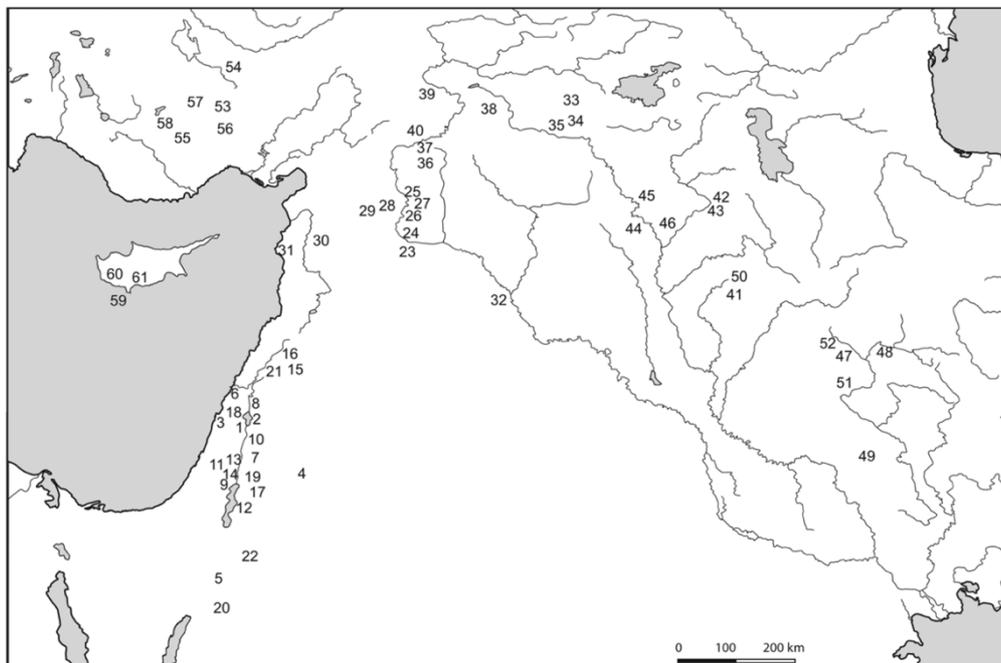


Figure 5 :Distribution of main Late Epipaleolithic and Neolithic sites in the Near East. 1, Ohalo II; 2, Ein Gev IV; 3, Neve David; 4, Kharahneh IV; 5, Beidha; 6, Hayonim; 7, Wadi al-Hammeh 27; 8, Ain Mallaha; 9, Jericho; 10, Iraq ed Dubb; 11, Hatoula; 12, Dhra; 13 (Zeder 2011)

The origins of the Neolithic

The Neolithic transition from foraging to food producing subsistence strategies with cultivated crops and domesticated animal took place in the Fertile Crescent spanning a large region from the Southern to the Northern Levant and from Northern Mesopotamia to the Zagros Mountains (Fig. 5). However, this transition was not a homogenous process in terms of both plant and animal domestication and human populations. The earliest spread of agriculture in SWA was a cultural diffusion with the transmission of ideas and farming technologies. This was demonstrated by the strong genetic differentiation between pre-pottery (aceramic) Neolithic farmers dating to the ~12th millennium BP in the Western and Eastern Fertile Crescent. The genetic continuity between Late Epipaleolithic hunter-gatherers and local Neolithic farmers indicates a local transition from foraging to farming practices ^{51,52}. In Central Anatolia, where aceramic Neolithic period started in the ~11th millennium BP, a demographic pattern similar to the one in the Fertile Crescent was detected since local Central Anatolian forager populations adopted the farming practices possibly via cultural diffusion ^{40,53,54}. Based on demographic modelling inferences, it was proposed that the Central Anatolian ancestry found in both Epipaleolithic and aceramic Neolithic individuals was a product of an admixture between Southeastern European and Levantine hunter-gatherers that took place in the final phase of the Late Pleistocene ³⁸. Agriculture spread to the West including the Aegean and Marmara regions up to Northern Greece during the 9th millennium BP ⁴⁶. Although later Anatolian Neolithic farmers of ceramic Neolithic sites such as Barcin in the Marmara region and Çatalhöyük in Central Anatolia are assumed to originate from the westward expansion of earlier Central Anatolian farmers, the genomes of later ceramic farmers show slight differences compared to both aceramic and Epipaleolithic individuals, which is so far explained by potential gene flow coming from the Fertile Crescent populations ^{40,54}. However, the genetic structure of the Aegean Mesolithic population is yet to be understood due to the lack of genetic data which might change the perception about the transition to farming in western Anatolia. What is clearer so far is that the ancient population found in Northwestern Anatolia and the Marmara and Aegean regions is the most plausible source of the first farmers of Europe, based on both genetic and archaeological data ^{38,46,52,53,55–57}.

The Neolithic expansion into Europe

These early farmers expanded into Europe through demic diffusion, during which they inhabited new areas and introduced their farming technology and animal and crop domesticates, and interacted with pre-existing forager populations to varying extents in different parts of Europe ^{39,41,58–64}.

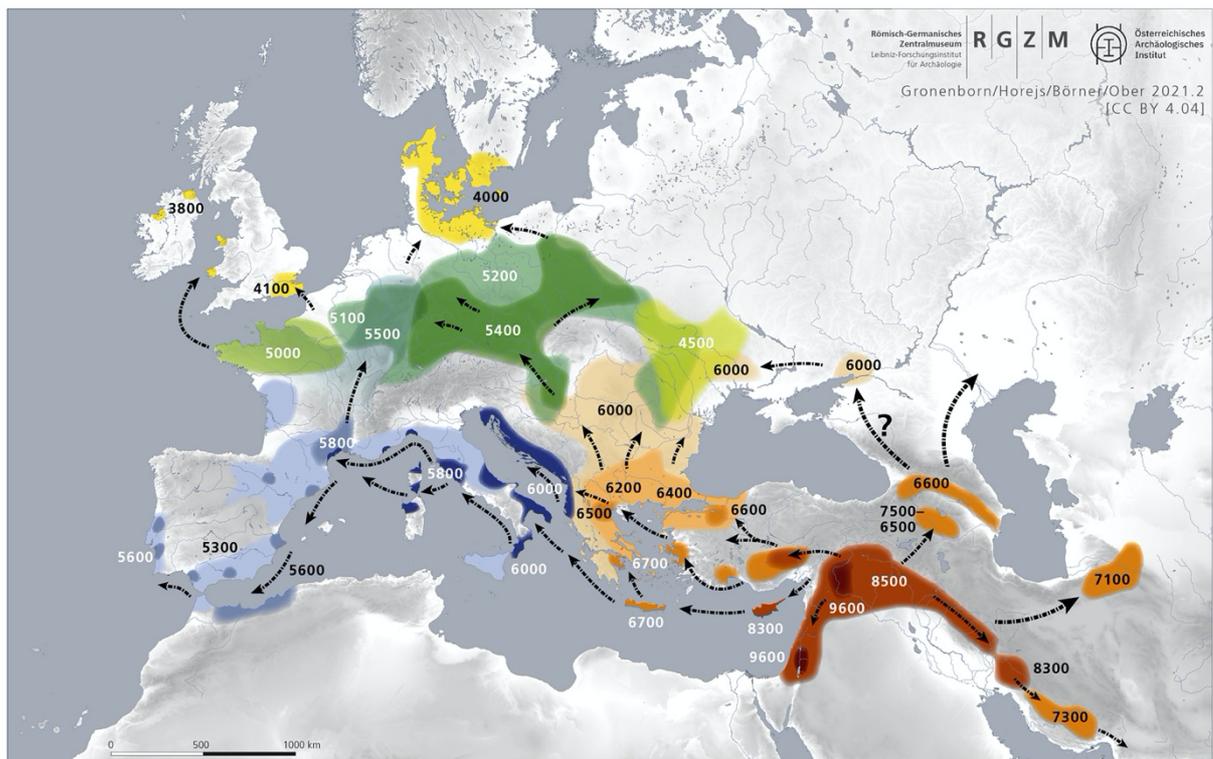


Figure 6: Condensed and simplified map of the spread of farming in western Eurasia. Gronenborn/Horejs/Börner/Ober 2021.2 (RGZM/ÖAI)

Southeastern Europe is a core zone for the understanding of the Neolithization of Europe as it is connected to both the earliest expansion of farming from the Northern Aegean area and the further spread of farming into the West and North. This region hosted Mesolithic forager populations with their unique archaeological culture and genetic structure recorded in many sites along the Danube gorges of the Iron Gates in present-day Serbia and Romania, which showed short lived interactions between foragers and farmers based on both material and genomic evidence ⁶⁵. Southeastern European Mesolithic foragers have their distinct genetic ancestry with a West-to-East cline, while they show stronger genetic affinity to the Mesolithic individuals in Western Europe compared to the ones in Eastern Europe ⁶⁶. This ancestry

possibly appeared during the LGM bottlenecks as suggested by the previously discussed demographic models ³⁸. Moreover, the genomes of these foragers show higher affinity to the ancestry that is found in Anatolian farmers relative to contemporary foragers in the West. This could be explained if they were one of the ancestral source populations of Anatolian foragers, as proposed, or through more recent interactions between these populations in Southeastern Europe and Northwestern Anatolia ^{38,66}. However, some later individuals in the site of Lepenski Vir dating to ~8,200 BP, the time period of the post-early Neolithic expansion, show a mixture of both farmer and forager ancestries that is a strong evidence of a local interaction of two distinct populations ⁶⁶. Admixture between Neolithic farmers and European hunter-gatherers was nevertheless rare since most of this early expansion was driven by farmers having Northwestern Anatolian/Aegean farmer ancestry and the pre-existing Mesolithic ancestry was not found in the majority of farmers. Another significant feature of the earliest farming communities in the Balkans that hints to links with their Aegean ancestors is the pottery. Some well-known Early Neolithic material cultures include the pre-Karanovo and Karanovo I in Bulgaria, the Starcevo in mostly Serbia, the Criş in Romania and the Korös in Hungary, which show a strong similarity of patterns from Thrace to the Carpathian Basin despite local variations ⁴⁶.

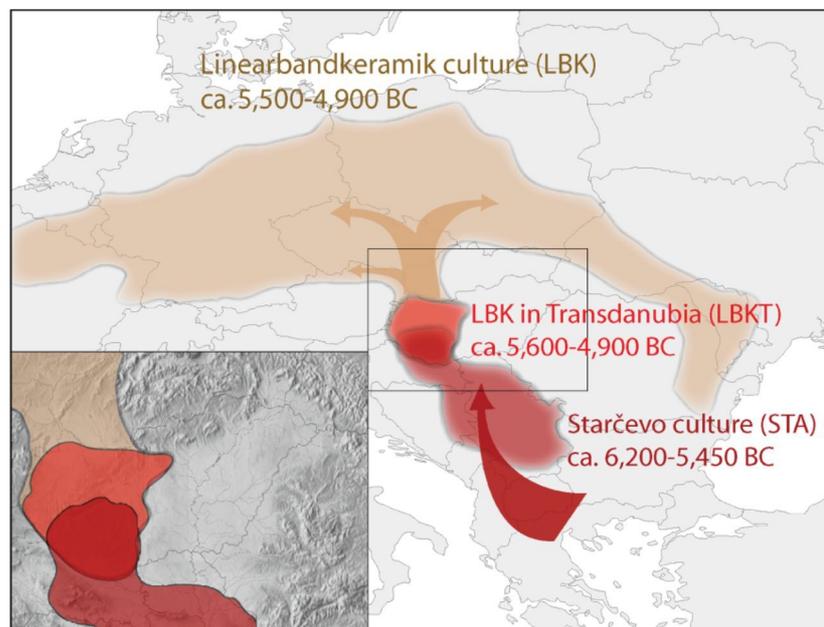


Figure 7 : Geographical distribution of the Starčevo, LBK cultures and the Continental route of the European Neolithization. The shaded areas of the maps show the distribution of the Starčevo culture (STA) and the LBK in Transdanubia (LBKT) and Central Europe (LBK) (Szécsényi-Nagy A. et al.2015)

The expansion of farming in Southeastern Europe that started around 8200 BP and took place through the Balkans reached the south of the Carpathian Basin and Transylvania by 8,000 BP. After this rapid expansion, however, the farmers did not continue to move to further locations for around 400 -500 years. During this time period, the Starčevo culture of the earliest farmers in the Balkans occupied a territory including present-day Serbia, northern Croatia and western Hungary. During the last phase of the Starčevo culture, a new decoration pattern with incised bands originated in Transdanubia that is referred to as Linearbandkeramik (Linear Band Pottery) Culture, commonly abbreviated as LBK, which became the main Early Neolithic culture (Middle Neolithic in Hungary) of continental Europe. This culture expanded west, north and eastwards, up to the Paris Basin in France, the Baltic Sea in Poland and to Western Ukraine, respectively (Fig. 7) ^{46,68,69}. Local subcultures of the LBK Culture developed in western Hungary, called LBK in Transdanubia (LBKT), and in eastern Hungary, called Alföld Linear Pottery (ALP), as well as the Bükk cultures spread in the opposite directions ⁷⁰.

Genetic analyses of individuals from Neolithic Hungary showed low differentiation between the Early Neolithic individuals associated with the Starčevo on one hand and the Körös cultures on the other hand, both having close relatedness to Anatolian farmers. However, one of the most striking findings from these studies was some Neolithic individuals in the Early Neolithic Körös site of eastern Hungary, who were carriers of a hunter-gatherer genome, showing a strong evidence of interaction between hunter-gatherers and Neolithic farmers and of possible acculturation of the former ones ^{61,71}. Furthermore, in western Hungarian sites associated with the LBKT culture, the estimated hunter-gatherer ancestry is quite limited compared to contemporary eastern Hungarian sites associated with the ALP culture, where hunter-gatherer ancestry is found to be higher than 10%. The genomic difference is in agreement with cultural differences and different levels of interactions in both earlier Körös and later ALP sites ⁶¹.

Contrary to Eastern Hungary, the LBK expansion through the Danube valley into Central Europe does not provide much evidence for interactions between farming and foraging populations. The areas occupied by early LBK populations, which were mostly highly fertile fields for agriculture called loess soils, were not populated by Late Mesolithic groups ^{46,72}. Even though both Late Mesolithic and LBK-associated groups lived contemporarily, the archaeological evidence of interaction in Central Europe, as restricted as it is, is in accordance with genetic data since LBK-associated individuals from present-day Germany and Austria show limited evidence of admixture ^{61,63,73}. The spread of farming in continental Europe shows a pattern similar to the one in the Balkans, and it is described as an overwhelmingly

demographic expansion by migrating long-distances and colonizing new settlement areas. However, the appearance of the LBK material culture represents a break with the ancestral cultures from the Aegean and Anatolian regions although it preserved the uniformity of its pottery style and of the architecture and the similarity between colonized areas along its expansion ⁴⁶. This pattern could be explained by a founder effect. This view finds support in the genetic similarity between different LBK-associated groups ⁶¹. Both genetic and archaeological data provide evidence for a scenario in which LBK-associated populations migrating to the West originated from LBKT-associated groups in Hungary and mixed with hunter-gatherers since their admixture proportions slightly increased with a range of 0-5% ⁶¹. However, the details of these early interactions remain unclear in the absence of a high-resolution genetic study focusing on this specific question.

The Early Neolithic West of the Rhine river

On the western end of the Continental route, the earliest farmers that were associated with the LBK culture crossed the Rhine river around 5,300 BCE and settled first in Alsace and moving from there further westward. Both the properties of the areas inhabited and the material culture including architecture, lithic tools and ceramics resembled the ones of the Central European LBK sites in the East of the Rhine with local decorative characteristics ⁷⁴. The Alsace region can be further divided into two groups based on cultural LBK differences. The differences found in southern Alsace (Haute-Alsace) and northern Alsace (Basse-Alsace) might be explained by different LBK origins between southern and northern Alsace or the presence and input of indigenous pre-Neolithic groups ⁷⁵. The genetic data so far published include three LBK-associated individuals, two from northern Alsace and one from southern Alsace ⁷⁵. The two northern Alsatian individuals show the genetic characteristics of eastern LBK populations with small proportions of hunter-gatherer ancestry while the southern Alsatian individuals have much higher hunter-gatherer ancestry, a proportion that was not detected in any other LBK individuals. These results might provide some genetic support to the archaeologically defined differences between southern and northern Alsace, which is further supported by the cultural connection between Northern Alsace and Baden-Württemberg ⁷⁵. The genetic data are not conclusive enough, however, given the low number of genomes sequenced.

The final phase of the LBK culture dates back to ~5,100-5,000 calBCE until 4,900 calBCE in the Paris Basin and it is traditionally referred to as RRBP – “le Rubané Récent du Bassin

Parisien” 74,76. Based on some shared properties of lithic technologies, there is a relation attested between LBK sites of the Northeast Paris Basin and Belgium 74. However, there is also archaeological evidence connecting the Paris Basin to southern Alsace 75.

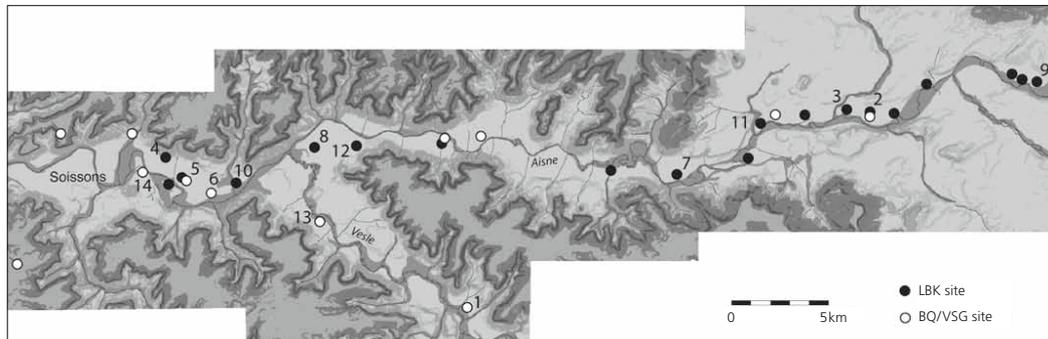


Figure 8: Distribution of Linear Pottery (LBK) and Blicquy/Villeneuve-Saint-Germain (BVSG) sites in the Aisne valley study zone (map after Chartier 2010, with modifications). Sites mentioned in the text: 1 Bazoches-sur-Vesle; 2 Berry-au-Bac “le Chemin de la Pêcheur”; 3 Berry-au-Bac “le Vieux Tordoir”; 4 Bucy-le-Long “la Fosselle”; 5 Bucy-le-Long “la Fosse Tounise”; 6 Bucy-le-Long “le Fond du Petit Marais”; 7 Cuiry-lès-Chaudardes; 8 Chassemy “le Grand Horle”; 9 Menneville “Derrière le Village”; 10 Missy-sur-Aisne “le Culot”; 11 Pontavert “le Port aux Marbres”; 12 Presles-et-Boves; 13 Vasseny; 14 Villeneuve-Saint-Germain. (Ilett 2012)

The Late LBK in the Aisne Valley

A highly notable example of the late LBK are the many linearly distributed LBK sites in the Aisne Valley in the Eastern Paris Basin, which were thoroughly excavated and studied for decades (Fig. 8). Although the sites have some unique properties such as the distribution along the Aisne river, and the presence of both short-lived and long-lived sites with varying sizes covering a time span of 200 years, they represent the characteristics of the typical West-Central European LBK culture 76

At the final stage of this Late LBK period, a new local Early Neolithic culture emerged in Northern France and Belgium, especially in the Paris Basin, referred to as Blicquy/Villeneuve-Saint-Germain (BVSG). The BVSG is assumed to follow the late LBK lasting from 4,900 to 4,700 BCE 77. However, there are indications of shifts in the settlement patterns based on house

plans and ceramic decorations and a striking absence of continuity between the usage of LBK and BVSG sites despite the proximity of the sites to each other (Fig. 8) ^{76,78}. Since the Paris Basin represents the end of the Continental Neolithic route, understanding the final phase of LBK in this area is essential to complete our view of the European Neolithization including its migration history, the Neolithic-Mesolithic admixture as well as inter and intra-regional Neolithic admixture history in Western Europe and the general history of interactions that can be deduced from it.

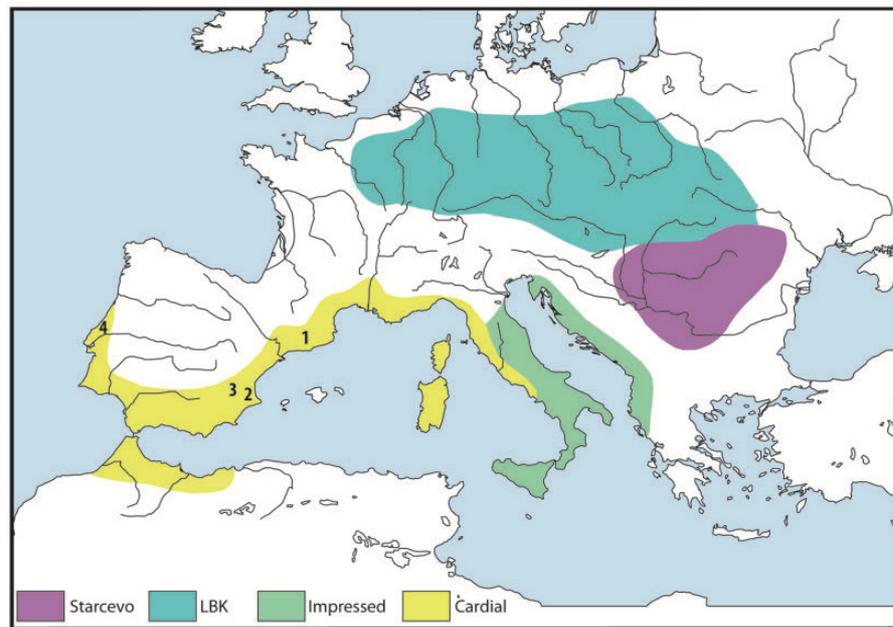


Figure 9: Main cultural horizons associated with the earliest Neolithic of Central and Western Europe approximately 6,000–5,500 cal BCE. 1, Cova Bonica; 2, Cova de la Sarsa; 3, Cova de l'Or; 4, Galeria da Cisterna-Almonda (Olalde et al., 2015)

The Mediterranean route of Neolithization

Prior to the LBK expansion into Central Europe, farming spread into Western Europe between 8,000 – 7,500 years ago via the Northern Mediterranean coast. The early members of this maritime expansion with origins in Northwestern Greece were associated with the Impressed Ware (Impresso-Cardial Complex - ICC) culture and inhabited the Adriatic coast and mainland Italy (Fig. 9) ⁴⁶. The two different expansions, the Continental and the Mediterranean expansion, have their roots the same metapopulation in Southeastern Europe since the genomes of the Early Neolithic individuals from Croatia are not separable from

contemporary individuals in the Balkans, both preserving their ancestry from Northwestern Anatolia ⁶⁶. Contrary to Neolithic Central Europe, the Neolithic populations spreading along the Mediterranean coast occupied mostly caves, rock shelters, and open camps sites. Some of these sites have evidence of occupation by both Late Mesolithic and Early Neolithic groups ⁷⁹. The Early Neolithic individuals from Central Italy dating to between 6,000-5,500 calBCE show similar patterns to Balkan and Central European Early Neolithic population, all having little to no hunter-gatherer ancestry. However, later Early Neolithic individuals dating to 5,500-5,200 calBCE have slightly elevated hunter-gatherer ancestry, while the latest member of this group has an ancestry proportion of around 15%, which separates this individual from other Early Neolithic individuals in Italy ⁵⁸. In Southern France and Iberia, where Early Neolithic sites are associated with the Cardial culture (Fig. 9), a newly established cultural complex dating to 5,500-5,400 calBCE, the genetic profile is overall different from the one of the Early Neolithic groups in Italy and in Central Europe. There are genome-wide data available for only four Early Neolithic Individuals dating to earlier than 5,000 BCE in Southern France ⁶³. Strikingly, their genomes reveal very recent hunter-gatherer admixture, around 4 generations ago, and HG ancestry proportions varying between 20-50%. This genomic profile makes this group of individuals unique among other contemporary individuals, and is in agreement with the archaeological findings about Mesolithic and Neolithic interactions in this part of Europe ⁶³. However, the amount of genomic data is insufficient to disentangle the hypotheses on the genesis of the Cardial culture, and in particular the question whether the input from the Mesolithic groups was responsible for changes in cultural practices ⁸⁰. Moreover, the pattern observed in Southern France is not reflecting the Cardial groups in the Iberian Peninsula, a region for which much more genetic data have been produced. Early Neolithic Iberian populations dating to 5,200-5,000 calBCE have around 10% hunter-gatherer ancestry, which is much less than in the French Cardial group yet higher than in Central European Early Neolithic populations ^{43,61,62,81}. Therefore, the modalities of the expansion of farming via the Mediterranean route remain elusive since only very little genetic information from Early Neolithic sites associated with the Impresso-Cardial cultures is available.

Spread of farming into Scandinavia and the British Isles

The spread of farming into Southern Scandinavia and the British Isles was delayed by ~1,000 years, as it only took place independently around 4,000 BCE (Fig. 6). In Scandinavia, the Neolithic transition was the result of the expansion of farmer populations associated with the Funnel Beaker (TRB - Trichterbecher) Culture, which emerged from the Early Neolithic culture in Northern Germany and Southern Scandinavia starting from 4,100 calBCE⁴⁶. These farmers do not show any signal of admixture with local Scandinavian hunter-gatherers. However, a pottery-producing group of foragers associated with the Pitted Ware Culture (PWC), which had a subsistence strategy based on marine hunting and fishing, was contemporary to the later TRB population and received gene flow from TRB farmers⁸².

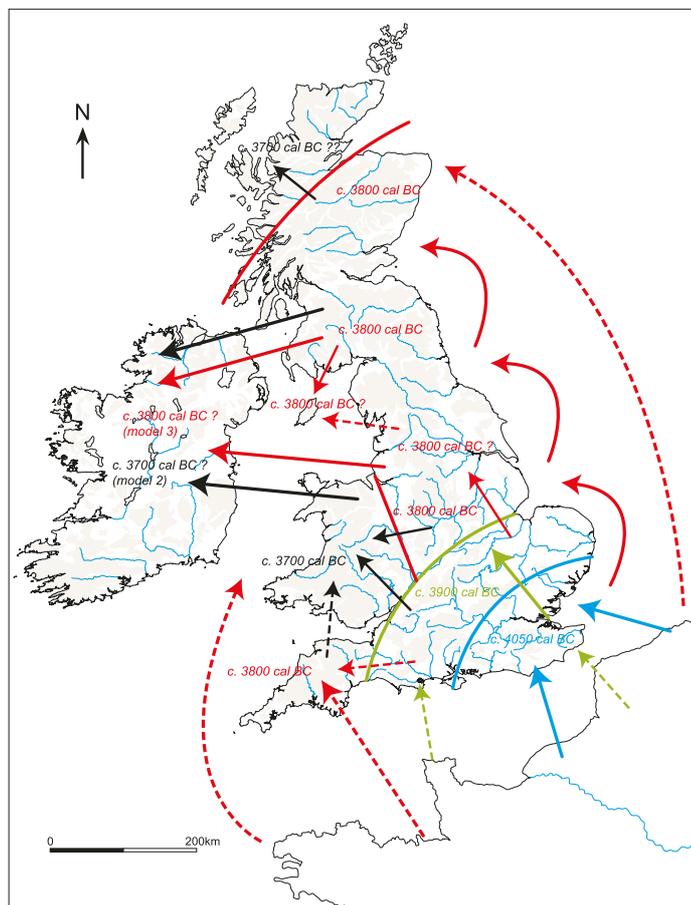


Figure 10: Fig. 15.8. Interpretive map of suggested dates, source areas and directions in the spread of Neolithic things and practices across Britain and Ireland. The colors denote contacts at different dates (A. W. R. Whittle et al., 2011).

The Neolithization of Britain and Ireland occurred almost at the same time as in Southern Scandinavia. It is not clear what was the driving force for northward migrations or why there was such a delay compared to the more southern regions of Europe. Concerning the arrival of the first farmers to the British Isles there are many unknowns regarding their origin. Based on the archaeological evidence different potential routes were described (Fig. 10). The “Atlantic, Breton Neolithic” model argues in favor of an origin in Brittany whose farmers spread to the Northern British Isles. The “Trans-Channel East strand” suggests the Northeastern coast of France as the source of Neolithic Britain via the southeastern corner of the island (also called the ‘Carinated Bowl Neolithic’) ⁸⁴. The “Trans-Channel West strand” proposes a population on the Northwestern coast of France as the source of Neolithic Britain via the southwestern corner of the island ⁸⁴. One way to identify the potential source of Early Neolithic Britain is to compare the genetic affinity of Neolithic British individuals to Iberian and Central European Early Neolithic groups. Consequently, it was proposed that the Neolithic population in Britain was likely to descend from an Iberian population that had arrived there via the Mediterranean route rather than the Continental route ⁵⁹. These results were also confirmed by another study including Neolithic individuals from France and which connected the populations in Northern and Southern France to the previously proposed connections between Iberia and Britain ⁶³. However, these studies were based on the genetic affinity between British and other western Neolithic groups but were unable to detect a contemporary population that was more strongly associated with Neolithic populations from Britain and Ireland. Furthermore, all the statistical tests of the genetic data that were performed included Central European LBK groups, which carry less hunter-gatherer ancestry compared to British and other Western Neolithic individuals. This introduced a potential bias since the Central LBK groups do not represent the end of the Continental expansion route, contrary to Iberia. It would be more relevant to include individuals from the Paris Basin LBK, but such data have not been produced yet.

Population Dynamics of the 3rd Millennium

The 3rd millennium BCE was a period of both cultural and genomic transformations in Europe, which led for the first time to the appearance of most of the present-day European genetic structure. These transformations have their origins in the westward expansions of steppe-ancestry populations related to those associated with the Yamnaya culture from the Pontic-Caspian steppes (Fig. 11)^{10,12}. In Eastern Europe, this westward expansion coincided with another expansion by people from a Late Neolithic culture called Globular Amphora that occurred in the opposite direction around 3,000-2,800 BCE, which possibly led to the emergence of a new culture and genetic structure found in the groups associated with the Corded Ware culture (CWC) (Fig. 12)⁸⁵. However, these two interacting groups were genetically (and culturally) highly differentiated. While the Yamnaya pastoralists were genetically a mix of well-structured Caucasian and Eastern hunter-gatherer-related groups ⁸⁶, the Late Neolithic people, who encountered this Yamnaya-related ancestry, were a mix of Aegean farmer and Western hunter-gatherer groups ⁸⁷. The individuals analyzed that belonged to this newly emerged CWC group in Central Eastern Europe carried small portions of Late Neolithic ancestry but most of their genomes resembled the Yamnaya profile (Fig. 11) (Haak et al., 2015; Linderholm et al., 2020; Papac et al., 2021; Saag et al., 2021). In less than two centuries, CWC-associated people expanded to the borders of the Volga and Rhine rivers in the East and the West, respectively. Although there are some analogies to the Early Neolithic expansions in terms of directionality and possibly interactions between locals and migrants, the Corded Ware transmission was more complex due to its high speed and the way of forming a new culture ⁸⁵.

cultural complex and steppe ancestry^{60,93}. In Iberia, the overall steppe ancestry found in BBC-associated individuals is much lower compared to Britain and the Netherlands^{93,94}. Further studies comprising these key periods around ~2,500 calBCE in France (Fig. 11, Fig. 12) are necessary in order to better understand the mechanisms behind the spread of steppe-ancestry in Western Europe.

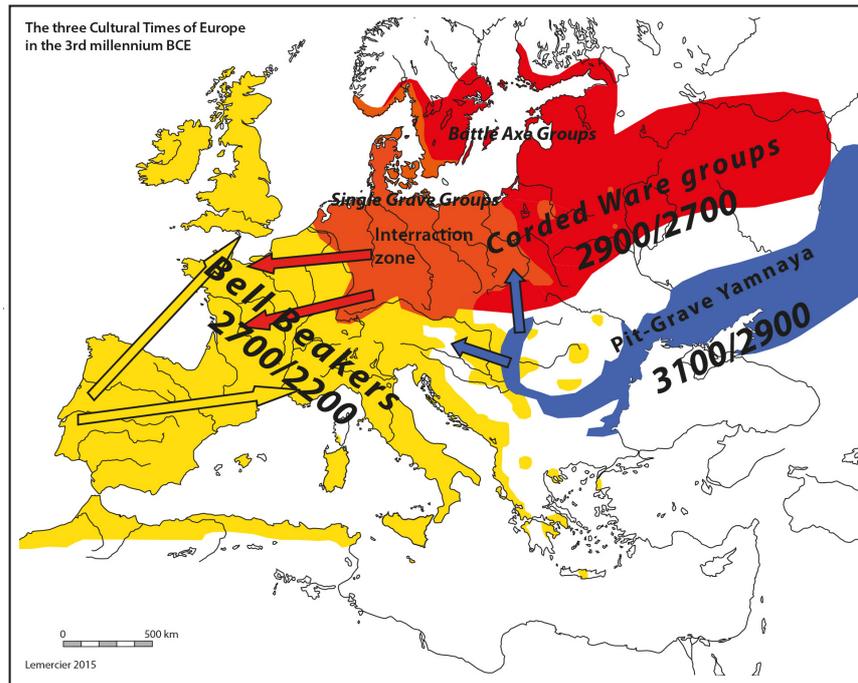


Figure 12: Yamnaya – Corded Ware – Bell Beaker distribution in Europe (reproduced from Lemercier 2018, Fig. 7) (Heyd, 2019)

Rationale of the Thesis

The briefly reviewed genetic history of human populations since the Upper Paleolithic in Europe shows recurring prehistorical encounters between local and migrating populations. People migrated due to climatic and environmental changes, demographic pressure, and/or cultural developments. Population movements were accompanied by local admixture (e.g. Cassidy et al., 2020; Fu et al., 2015; Martiniano et al., 2017) or population replacement events (e.g. Brace et al., 2019, Olalde et al., 2018). However, the mechanisms of these populational events are still elusive for most time periods and geographic areas. Therefore, we aimed at revealing and understanding at the genetic level admixture between populations by focusing on three key transition periods in European prehistory. First, the time $\sim 38,000$ years ago following major climatic changes and an ecosystem crisis beginning with a dramatic volcanic eruption (the Campanian Ignimbrite (CI) super-eruption at the beginning of the Heinrich Event 4 (Fig. 4)). In the aftermath of this crisis Europe experienced human population turnover that introduced the present-day European like ancestry for the first time. Second, the transition from the Mesolithic to the Neolithic lifestyle following the population expansions from Southwest Asia that resulted in the spread of farming into Europe during the Holocene (Fig.6). The last period we focused on is the 3rd millennium BCE, which was a period of both cultural and genomic transformations associated with migrations of populations from the Pontic-Caspian steppe forming the present-day European genome.

Within the framework of this thesis project, we sampled individuals, who were contemporary witnesses of these key periods, and generated ancient genomes. We performed various state-of-the-art population genomics methods to reconstruct the population history of the corresponding human groups. First, we analyzed the genomes of two early $\sim 36,000$ and $37,000$ -years-old European individuals from the rock shelter site of Buran-Kaya III in the Crimean Peninsula that were witnesses of the post-CI population dynamics. Then, we generated genomic data from 116 individuals dating from the Late Mesolithic to the beginning of the Bronze Age, between 8200 and 4200 years ago covering two large-scale populational turnover events in Europe. To provide genomic insights into the Neolithization of western Europe, we focused on the Paris Basin including the archaeologically well-studied Aisne Valley (Fig. 8) as the final phase of the continental expansion of farmer populations, where archaeological evidence for interactions between farmers and foragers were found (Allard,

2007). Lastly, we analyzed a collective burial dating back to 2500 BCE, a time when Yamnaya-related steppe ancestry was introduced in western Europe (Fig. 10) contributing the third major component of the present-day European genome and triggering the emergence of new cultures as a consequence of the encounters of populations associated with this new ancestry (Fig. 11).

References

1. Briggs AW, Stenzel U, Johnson PLF, et al. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Natl Acad Sci.* 2007;104(37):14616-14621. doi:10.1073/pnas.0704665104
2. Dabney J, Meyer M, Pääbo, Svante. Ancient DNA damage. *Cold Spring Harbor perspectives in biology.* Published online 2013. doi:https://doi.org/10.1101/cshperspect.a012567
3. Briggs AW, Stenzel U, Meyer M, Krause J, Kircher M, Pääbo S. Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Res.* 2010;38(6):e87-e87. doi:10.1093/nar/gkp1163
4. Bergström A, McCarthy SA, Hui R, et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science.* 2020;367(6484):eaay5012. doi:10.1126/science.aay5012
5. Choin J, Mendoza-Revilla J, Arauna LR, et al. Genomic insights into population history and biological adaptation in Oceania. *Nature.* 2021;592(7855):583-589. doi:10.1038/s41586-021-03236-5
6. Green RE, Krause J, Briggs AW, et al. A Draft Sequence of the Neandertal Genome. *Science.* 2010;328(5979):710-722. doi:10.1126/science.1188021
7. Mallick S, Li H, Lipson M, et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature.* 2016;538(7624):201-206. doi:10.1038/nature18964
8. Reich D, Green RE, Kircher M, et al. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature.* 2010;468(7327):1053-1060. doi:10.1038/nature09710
9. Reich D, Patterson N, Kircher M, et al. Denisova Admixture and the First Modern Human Dispersals into Southeast Asia and Oceania. *Am J Hum Genet.* 2011;89(4):516-528. doi:10.1016/j.ajhg.2011.09.005
10. Allentoft ME, Sikora M, Sjögren KG, et al. Population genomics of Bronze Age Eurasia. *Nature.* 2015;522(7555):167-172. doi:10.1038/nature14507
11. Fu Q, Posth C, Hajdinjak M, et al. The genetic history of Ice Age Europe. *Nature.* 2016;534(7606):200-205. doi:10.1038/nature17993
12. Haak W, Lazaridis I, Patterson N, et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature.* 2015;522(7555):207-211. doi:10.1038/nature14317
13. Lazaridis I, Patterson N, Mittnik A, et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature.* 2014;513(7518):409-413. doi:10.1038/nature13673
14. Childebayeva A, Rohrlach AB, Barquera R, et al. Population Genetics and Signatures of Selection in Early Neolithic European Farmers. *Mol Biol Evol.* 2022;39(6):msac108. doi:10.1093/molbev/msac108
15. Dannemann M, Prüfer K, Kelso J. Functional implications of Neandertal introgression in modern humans. *Genome Biol.* 2017;18(1):61. doi:10.1186/s13059-017-1181-7

16. Mathieson I, Lazaridis I, Rohland N, et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*. 2015;528(7583):499-503. doi:10.1038/nature16152
17. Racimo F, Marnetto D, Huerta-Sánchez E. Signatures of Archaic Adaptive Introgression in Present-Day Human Populations. *Mol Biol Evol*. 2017;34(2):296-317. doi:10.1093/molbev/msw216
18. Sankararaman S, Mallick S, Patterson N, Reich D. The Combined Landscape of Denisovan and Neanderthal Ancestry in Present-Day Humans. *Curr Biol*. 2016;26(9):1241-1247. doi:10.1016/j.cub.2016.03.037
19. Skoglund P, Mathieson I. Ancient Genomics of Modern Humans: The First Decade. *Annu Rev Genomics Hum Genet*. 2018;19(1):381-404. doi:10.1146/annurev-genom-083117-021749
20. Tishkoff SA, Reed FA, Friedlaender FR, et al. The Genetic Structure and History of Africans and African Americans. *Science*. 2009;324(5930):1035-1044. doi:10.1126/science.1172257
21. Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC. African Populations and the Evolution of Human Mitochondrial DNA. *Science*. 1991;253(5027):1503-1507. doi:10.1126/science.1840702
22. Posth C, Renaud G, Mittnik A, et al. Pleistocene Mitochondrial Genomes Suggest a Single Major Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. *Curr Biol*. 2016;26(6):827-833. doi:10.1016/j.cub.2016.01.037
23. Pagani L, Lawson DJ, Jagoda E, et al. Genomic analyses inform on migration events during the peopling of Eurasia. *Nature*. 2016;538(7624):238-242. doi:10.1038/nature19792
24. Schiffels S, Durbin R. Inferring human population size and separation history from multiple genome sequences. *Nat Genet*. 2014;46(8):919-925. doi:10.1038/ng.3015
25. Wang K, Mathieson I, O'Connell J, Schiffels S. Tracking human population structure through time from whole genome sequences. *PLOS Genet*. 2020;16(3):e1008552. doi:10.1371/journal.pgen.1008552
26. Terhorst J, Kamm JA, Song YS. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nat Genet*. 2017;49(2):303-309. doi:10.1038/ng.3748
27. Chen L, Wolf AB, Fu W, Li L, Akey JM. Identifying and Interpreting Apparent Neanderthal Ancestry in African Individuals. *Cell*. 2020;180(4):677-687.e16. doi:10.1016/j.cell.2020.01.012
28. Fu Q, Hajdinjak M, Moldovan OT, et al. An early modern human from Romania with a recent Neanderthal ancestor. *Nature*. 2015;524(7564):216-219. doi:10.1038/nature14558
29. Hajdinjak M, Mafessoni F, Skov L, et al. Initial Upper Palaeolithic humans in Europe had recent Neanderthal ancestry. *Nature*. 2021;592(7853):253-257. doi:10.1038/s41586-021-03335-3
30. Moorjani P, Sankararaman S, Fu Q, Przeworski M, Patterson N, Reich D. A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. *Proc Natl Acad Sci*. 2016;113(20):5652-5657. doi:10.1073/pnas.1514696113

31. Malaspinas AS, Westaway MC, Muller C, et al. A genomic history of Aboriginal Australia. *Nature*. 2016;538(7624):207-214. doi:10.1038/nature18299
32. Liu Y, Mao X, Krause J, Fu Q. Insights into human history from the first decade of ancient human genomics. *Science*. 2021;373(6562):1479-1484. doi:10.1126/science.abi8202
33. Prüfer K, Posth C, Yu H, et al. A genome sequence from a modern human skull over 45,000 years old from Zlatý kůň in Czechia. *Nat Ecol Evol*. 2021;5(6):820-825. doi:10.1038/s41559-021-01443-x
34. Yang MA, Gao X, Theunert C, et al. 40,000-Year-Old Individual from Asia Provides Insight into Early Population Structure in Eurasia. *Curr Biol*. 2017;27(20):3202-3208.e9. doi:10.1016/j.cub.2017.09.030
35. Seguin-Orlando A, Korneliussen TS, Sikora M, et al. Genomic structure in Europeans dating back at least 36,200 years. *Science*. 2014;346(6213):1113-1118. doi:10.1126/science.aaa0114
36. Clark PU, Dyke AS, Shakun JD, et al. The Last Glacial Maximum. *Science*. 2009;325(5941):710-714. doi:10.1126/science.1172873
37. Svensson E, Günther T, Hoischen A, et al. Genome of Peștera Muierii skull shows high diversity and low mutational load in pre-glacial Europe. *Curr Biol*. 2021;31(14):2973-2983.e9. doi:10.1016/j.cub.2021.04.045
38. Marchi N, Winkelbach L, Schulz I, et al. The genomic origins of the world's first farmers. *Cell*. 2022;185(11):1842-1859.e18. doi:10.1016/j.cell.2022.04.008
39. Cassidy LM, Maoldúin RÓ, Kador T, et al. A dynastic elite in monumental Neolithic society. *Nature*. 2020;582(7812):384-388. doi:10.1038/s41586-020-2378-6
40. Feldman M, Fernández-Domínguez E, Reynolds L, et al. Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nat Commun*. 2019;10(1):1218. doi:10.1038/s41467-019-09209-7
41. González-Fortes G, Jones ER, Lightfoot E, et al. Paleogenomic Evidence for Multi-generational Mixing between Neolithic Farmers and Mesolithic Hunter-Gatherers in the Lower Danube Basin. *Curr Biol*. 2017;27(12):1801-1810.e10. doi:10.1016/j.cub.2017.05.023
42. Günther T, Malmström H, Svensson EM, et al. Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation. *PLOS Biol*. 2018;16(1):e2003703. doi:10.1371/journal.pbio.2003703
43. Villalba-Mouco V, van de Loosdrecht MS, Posth C, et al. Survival of Late Pleistocene Hunter-Gatherer Ancestry in the Iberian Peninsula. *Curr Biol*. 2019;29(7):1169-1177.e7. doi:10.1016/j.cub.2019.02.006
44. Bortolini E, Pagani L, Oxilia G, et al. Early Alpine occupation backdates westward human migration in Late Glacial Europe. *Curr Biol*. 2021;31(11):2484-2493.e7. doi:10.1016/j.cub.2021.03.078
45. Bocquet-Appel JP. The Agricultural Demographic Transition During and After the Agriculture Inventions. *Curr Anthropol*. 2011;52(S4):S497-S510. doi:10.1086/659243
46. Shennan S. *The First Farmers of Europe: An Evolutionary Perspective*. Cambridge University Press; 2018. doi:10.1017/9781108386029

47. Piperno DR, Weiss E, Holst I, Nadel D. Processing of wild cereal grains in the Upper Palaeolithic revealed by starch grain analysis. *Nature*. 2004;430(7000):670-673. doi:10.1038/nature02734
48. Snir A, Nadel D, Groman-Yaroslavski I, et al. The Origin of Cultivation and Proto-Weeds, Long Before Neolithic Farming. *PLOS ONE*. 2015;10(7):e0131422. doi:10.1371/journal.pone.0131422
49. Bar-Yosef O. The Natufian culture in the Levant, threshold to the origins of agriculture. *Evol Anthropol Issues News Rev*. 1998;6(5):159-177. doi:10.1002/(SICI)1520-6505(1998)6:5<159::AID-EVAN4>3.0.CO;2-7
50. Zeder MA. The Origins of Agriculture in the Near East. *Curr Anthropol*. 2011;52(S4):S221-S235. doi:10.1086/659307
51. Broushaki F, Thomas MG, Link V, et al. Early Neolithic genomes from the eastern Fertile Crescent. *Science*. 2016;353(6298):499-503. doi:10.1126/science.aaf7943
52. Lazaridis I, Nadel D, Rollefson G, et al. Genomic insights into the origin of farming in the ancient Near East. *Nature*. 2016;536(7617):419-424. doi:10.1038/nature19310
53. Kılınç GM, Omrak A, Özer F, et al. The Demographic Development of the First Farmers in Anatolia. *Curr Biol*. 2016;26(19):2659-2666. doi:10.1016/j.cub.2016.07.057
54. Yaka R, Mapelli I, Kaptan D, et al. Variable kinship patterns in Neolithic Anatolia revealed by ancient genomes. *Curr Biol*. 2021;31(11):2455-2468.e18. doi:10.1016/j.cub.2021.03.050
55. Hofmanová Z, Kreutzer S, Hellenthal G, et al. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc Natl Acad Sci*. 2016;113(25):6886-6891. doi:10.1073/pnas.1523951113
56. Özdoğan M. Archaeological Evidence on the Westward Expansion of Farming Communities from Eastern Anatolia to the Aegean and the Balkans. *Curr Anthropol*. 2011;52(S4):S415-S430. doi:10.1086/658895
57. Özdoğan, Mehmet. The Earliest Farmers of Europe. Where did They Come from? In: Báčvarov K, Nikolov V, Gleser R, eds. *Southeast Europe and Anatolia in Prehistory: Essays in Honor of Vassil Nikolov on His 65th Anniversary*. Universitätsforschungen zur prähistorischen Archäologie. Verlag Dr. Rudolf Habelt GmbH; 2016. <https://books.google.fr/books?id=qaCHnQAACAAJ>
58. Antonio ML, Gao Z, Moots HM, et al. Ancient Rome: A genetic crossroads of Europe and the Mediterranean. *Science*. 2019;366(6466):708-714. doi:10.1126/science.aay6826
59. Brace S. Ancient genomes indicate population replacement in Early Neolithic Britain. *Nat Ecol Evol*. 2019;3. doi:10.1038/s41559-019-0871-9
60. Brunel S, Bennett EA, Cardin L, et al. Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proc Natl Acad Sci*. 2020;117(23):12791-12798. doi:10.1073/pnas.1918034117
61. Lipson M, Szécsényi-Nagy A, Mallick S, et al. Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature*. 2017;551(7680):368-372. doi:10.1038/nature24476
62. Olalde I, Schroeder H, Sandoval-Velasco M, et al. A Common Genetic Origin for Early Farmers from Mediterranean Cardial and Central European LBK Cultures. *Mol Biol Evol*. 2015;32(12):3132-3142. doi:10.1093/molbev/msv181

63. Rivollat M, Jeong C, Schiffels S, et al. Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers. *Sci Adv.* 2020;6(22):eaa5344. doi:10.1126/sciadv.aaz5344
64. Skoglund P. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc Natl Acad Sci USA.* 2014;111. doi:10.1073/pnas.1318934111
65. Borić D. Adaptations and Transformations of the Danube Gorges Foragers (c. 13.000–5500 BC): An Overview. In: Krauss R, Deutsches Archäologisches Institut, eds. *Beginnings - New Research in the Appearance of the Neolithic Between Northwest Anatolia and the Carpathian Basin: Papers of the International Workshop, 8th-9th April 2009, Istanbul, Organized by Dan Ciobotaru, Barbara Horejs and Raiko Krauss.* Forschungscluster 1, Von der Sesshaftigkeit zur komplexen Gesellschaft. VML, Verlag Marie Leidorf; 2011. <https://books.google.fr/books?id=otEZMwEACAAJ>
66. Mathieson I, Alpaslan-Roodenberg S, Posth C, et al. The genomic history of southeastern Europe. *Nature.* 2018;555(7695):197-203. doi:10.1038/nature25778
67. Szécsényi-Nagy A, Brandt G, Haak W, et al. Tracing the genetic origin of Europe's first farmers reveals insights into their social organization. *Proc R Soc B Biol Sci.* 2015;282(1805):20150339. doi:10.1098/rspb.2015.0339
68. Bogucki P. The Spread of Early Farming in Europe. *Am Sci.* 1996;84(3):242-253.
69. Oross K, Bánff E. Three successive waves of Neolithisation: LBK development in Transdanubia. *Documenta Praehistorica.* 36(0). doi:10.4312/dp.36.11
70. Domboróczki L. *Settlement Structures of the Alföld Linear Pottery Culture (ALPC) in Heves County (North-Eastern Hungary): Development Models and Historical Reconstructions on Micro, Meso and Macro Levels.* Polska Akademia Umiejętności, Kraków; 2009.
71. Gamba C, Jones ER, Teasdale MD, et al. Genome flux and stasis in a five millennium transect of European prehistory. *Nat Commun.* 2014;5(1):5257. doi:10.1038/ncomms6257
72. Bakels CC. *The Western European Loess Belt: Agrarian History, 5300 BC - AD 1000.* Springer; 2009. https://books.google.fr/books?id=x8_ZjwEACAAJ
73. Mathieson I. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature.* 2015;528. doi:10.1038/nature16152
74. Allard P. The Mesolithic-Neolithic transition in the Paris Basin: a review. In: Whittle A, Cummings V, eds. *Going Over: The Mesolithic-Neolithic Transition in North-West Europe.* British Academy; 2007:0. doi:10.5871/bacad/9780197264140.003.0011
75. Bickle P, Whittle A, Anders A, et al. *The First Farmers of Central Europe.* Oxbow Books; 2013. Accessed October 17, 2022. <http://www.jstor.org/stable/j.ctt14bs18v>
76. Ilett M. Linear Pottery and Blicquy/Villeneuve-Saint-Germain settlement in the Aisne valley and its environs: an overview. In: Wolfram, Sabine, Stäuble, Harald, eds. *Siedlungsstruktur Und Kulturwandel in Der Bandkeramik: Beiträge Der Internationalen Tagung „Neue Fragen Zur Bandkeramik Oder Alles Beim Alten?!“ Leipzig 23. Bis 24. September 2010.* ; 2012.
77. Dubouloz, Jérôme. Datation absolue du premier Néolithique du Bassin parisien : complément et relecture des données RRBP et VSG. *Bulletin de la Société préhistorique française.* Published online 2003:671-689.

78. Hamon C, Gomart L. Social Rules and Household Interactions Within the LBK: Long-Standing Debates, New Perspectives. 2021;7(1):690-704. doi:10.1515/opar-2020-0158
79. Malone C. The Italian Neolithic: A Synthesis of Research. *J World Prehistory*. 2003;17(3):235-312.
80. Guilaine J, Manen C. 2021 From Mesolithic to Early Neolithic in the western Mediterranean. In: Whittle A, Cummings V, eds. *Going Over: The Mesolithic-Neolithic Transition in North-West Europe*. British Academy; 2007:0. doi:10.5871/bacad/9780197264140.003.0003
81. Valdiosera C, Günther T, Vera-Rodríguez JC, et al. Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia. *Proc Natl Acad Sci*. 2018;115(13):3428-3433. doi:10.1073/pnas.1717762115
82. Mittnik A, Wang CC, Pfrengle S, et al. The genetic prehistory of the Baltic Sea region. *Nat Commun*. 2018;9(1):442. doi:10.1038/s41467-018-02825-9
83. Whittle AWR, Healy F, Bayliss A, Allen MJ. *Gathering Time: Dating the Early Neolithic Enclosures of Southern Britain and Ireland*. Oxbow Books; 2011. <https://books.google.fr/books?id=OV5ApwAACAAJ>
84. Sheridan A. The Neolithization of Britain and Ireland: The 'Big Picture.' In: Finlayson, Bill, Warren, Graeme, eds. *Landscapes in Transition*. Oxbow Books; 2010.
85. Heyd V, Kulcsár, Gabriela, Preda-Balanica, Bianca Elena. *Yamnaya Interactions: Proceedings of the International Workshop Held in Helsinki, 25-26 April 2019*. Vol 44. Archaeolingua; 2021.
86. Jones ER, Gonzalez-Fortes G, Connell S, et al. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat Commun*. 2015;6(1):8912. doi:10.1038/ncomms9912
87. Schroeder H, Margaryan A, Szmyt M, et al. Unraveling ancestry, kinship, and violence in a Late Neolithic mass grave. *Proc Natl Acad Sci*. 2019;116(22):10705-10710. doi:10.1073/pnas.1820210116
88. Haak W, Lazaridis I, Patterson N, et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*. 2015;522(7555):207-211. doi:10.1038/nature14317
89. Linderholm A, Kılınc GM, Szczepanek A, et al. Corded Ware cultural complexity uncovered using genomic and isotopic analysis from south-eastern Poland. *Sci Rep*. 2020;10(1):6885. doi:10.1038/s41598-020-63138-w
90. Papac L, Ernée M, Dobeš M, et al. Dynamic changes in genomic and social structures in third millennium BCE central Europe. *Sci Adv*. 2021;7(35):eabi6941. doi:10.1126/sciadv.abi6941
91. Saag L, Vasilyev SV, Varul L, et al. Genetic ancestry changes in Stone to Bronze Age transition in the East European plain. *Sci Adv*. 2021;7(4):eabd6535. doi:10.1126/sciadv.abd6535
92. Narasimhan VM, Patterson N, Moorjani P, et al. The formation of human populations in South and Central Asia. *Science*. 2019;365(6457):eaat7487. doi:10.1126/science.aat7487
93. Olalde I, Brace S, Allentoft ME, et al. The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature*. 2018;555(7695):190-196. doi:10.1038/nature25738

94. Villalba-Mouco V, Oliart C, Rihuete-Herrada C, et al. Genomic transformation and social organization during the Copper Age–Bronze Age transition in southern Iberia. *Sci Adv.* 2021;7(47):eabi7038. doi:10.1126/sciadv.abi7038
95. Heyd V. Yamnaya – Corded Wares – Bell Beakers: or how to conceptualize events of 5000 years ago that shaped modern Europe. In: *Yamnaya – Corded Wares – Bell Beakers, or How to Conceptualize Events of 5000 Years Ago That Shaped Modern Europe. In: Valchev. T. (Ed.), Studia in Honorem Plae Pliev. Регионален Исторически Музей - Ямбол/Yambol. Вестни На Ямболския Музей /Година VI/9, 125-136. Regional Historical Museum of Yambol; 2019.*
96. Martiniano R, Cassidy LM, Ó'Maoldúin R, et al. The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLOS Genet.* 2017;13(7):e1006852. doi:10.1371/journal.pgen.1006852

Chapter I

Early Upper Paleolithic genomes of Crimea show migration and admixture dynamics of the first modern European ancestries in relation to climatic crisis

E. Andrew Bennett^{1,5,6}, Oguzhan Parasayan^{1,6}, Sandrine Prat², Stéphane Péan³, Laurent Crépin³, Alexandr Yanevich⁴, Thierry Grange^{1*}, Eva-Maria Geigl^{1*}

¹Institut Jacques Monod, CNRS, Université de Paris, 75013 Paris, France.

²UMR 7194 (HNHP), MNHN/CNRS/UPVD, Alliance Sorbonne Université, Musée de l'Homme, Palais de Chaillot, 17 Place du Trocadéro, 75116, Paris, France.

³UMR 7194 (HNHP), MNHN/CNRS/UPVD, Muséum national d'Histoire naturelle, Alliance Sorbonne Université, Institut de Paléontologie Humaine, 1 rue René Panhard, 75013, Paris, France.

⁴Institute of Archaeology, National Academy of Sciences of Ukraine, Heroiv Stalingrada 12, 04210 Kyiv, Ukraine.

⁵Present address: Key Laboratory of Vertebrate Evolution and Human Origins, Institute of Vertebrate Paleontology and Paleoanthropology, Center for Excellence in Life and Paleoenvironment, Chinese Academy of Sciences, Beijing 100044, China

⁶These authors contributed equally

*Correspondence to: eva-maria.geigl@ijm.fr or thierry.grange@ijm.fr

Abstract

We describe the genomes of two c. 37,000-year-old individuals from the site of Buran-Kaya III in Crimea as belonging to the first migration wave of humans carrying ancestry found in present-day Europeans. These genomes allowed us to model admixture and infer population movements, interactions and turnovers during the peopling of Europe in the Early Upper Paleolithic, and revealed that admixture with preexisting populations in Europe accompanied the population turnover events prior to and following Heinrich Stadial 4 and the Campanian Ignimbrite super-eruption. The inhabitants of Buran-Kaya III shared the highest genomic similarities with Middle-Upper Paleolithic groups found in central and western Europe several thousands of years later and with Caucasus Late Upper Paleolithic groups. Since these remains are associated with the oldest appearance of a cultural assemblage with features reminiscent of the later Gravettian assemblages, these genetic links open up the possibility of technological continuity despite temporal and geographic distances.

Introduction

Archeological and genetic data show anatomically modern humans (AMHs) to be present in Europe by at least 45,000 BP (years Before Present)¹, and possibly earlier². Genomic analyses of AMH remains prior to 40,000 years ago have revealed diverse and poorly characterized populations^{1,3-5}. All of these showed evidence of admixture with Neanderthals either largely predating the 31-52,000-year-old woman from Zlatý Kůň in Czechia⁴, or more recent, local events in the case of genomes from Peștera cu Oase in Romania (Oase1)⁶ and Bacho Kiro in Bulgaria¹. Only the individuals from Bacho Kiro, who were found associated with Initial Upper Paleolithic archeological assemblages (IUP), have demonstrated some genomic relationship with modern human populations, specifically those found in East Asia¹. At some point after 40,000 years ago, these early ancestries appear to vanish from Europe as a new genomic profile appears, a timing concurrent with changing climatic and environmental conditions after the Campanian Ignimbrite super-eruption in Italy, which blanketed southern and eastern Europe in ash 39,490-39,170 BP, at the beginning of the nearly 2,000-year-long Heinrich Stadial 4 cold period (H4) with potentially catastrophic impact on the ecosystems^{7,8}. This post-Campanian

Ignimbrite (post-CI) wave of ancestry first appears after the Heinrich Stadial 4 and is present in Europe in various forms during the Early and Mid-Upper Paleolithic (EUP and MUP, respectively), such as in genomes from the sites of Kostenki14 (~38,000 cal BP)⁹ and Sunghir (~34,000 cal BP)¹⁰ in the eastern European Plain and in two ~35,000-year-old individuals, GoyetQ116-1 from Belgium³ and BK1653¹ from the MUP deposits of Bacho Kiro, Bulgaria, both of whom also contain trace ancestry linked to the previous IUP inhabitants of Bacho Kiro¹. Post-CI ancestry is also found in central Europe beginning at least 31,000 years ago in individuals belonging to the “Věstonice” genetic cluster, found in direct association with Gravettian cultural assemblages³. Unlike the IUP occupants of Europe, this post-CI wave of ancestry shares genetic drift with modern Europeans^{3,9}, but due to the sparseness of genomic data from EUP humans in Europe, the origins of this ancestry, the timing of its appearance(s), and the routes that populations bearing this ancestry took into the European continent, as well as their interactions with preceding IUP-contemporary groups are unknown.

Discovered in 1990, the rock shelter site of Buran-Kaya III in the Crimean Peninsula contains rich deposits of human activity dating from the Middle Paleolithic until the Middle Ages¹¹⁻¹⁴. The assemblages of layers 6-5 to 5-2, dated to ca. 38-34,000 cal BP^{12,14}, have yielded backed microliths, microgravette points, bone tools, ochre, and body ornaments of ivory, shell, and teeth, as well as multiple human skull fragments, some with post-mortem cut marks¹². The similarity of the cultural deposits to later Gravettian assemblages of Europe have led the excavating archeologists to assign this material to the earliest appearance of a Gravettian cultural layer, a designation that has recently been challenged on the basis of the considerable gap of time and geography between these artefacts and those appearing 5,000 years later in Gravettian layers of central Europe¹⁵⁻¹⁸. Other arguments consider that the assemblages of the EUP layers of Buran-Kaya III can be more parsimoniously associated with a cluster of non-Aurignacian contemporaneous EUP layers found in the Caucasus also containing backed lithics^{15,18,19}, such as those at Ortvale Klde, layers 4d-4c²⁰, Mezmaiskaya, layers 1c-1a²¹, and Dzudzuana, layer D²². Still others have proposed a dual approach and termed EUP layers of Buran-Kaya III along with those in the northern Caucasus sites as “proto-Gravettian”^{23,24}, and the similarities of these backed-lithic Caucasus sites with the Early Ahmarian lithics produced in the Levant has been used by some to suggest connections between these industries^{22,25,26}. While a strict correlation between the continuity of cultural and technological practices and the genetic continuity of the creators of archeological material cannot always be expected, genomic

characterization, where possible, of human remains associated with archeological assemblages provides additional keys to understanding the cultural and historical contexts of the material and the behaviors, adaptations and cultural contacts of ancient human groups. To define the population associated with the earliest appearance of this cultural assemblage in Europe, and to determine how the EUP occupants of Crimea fit into our current knowledge of the AMH settlement dynamics of Europe at the close of the Heinrich 4 period 38,000 years ago, we investigated the genetic identity of the human remains of the EUP layer of Buran-Kaya III.

Results

Two cranial fragments (BuKa3A and BuKa3C), one with cut marks (BuKa3A), were aseptically excavated in 2009 from layers 6-1 (BuKa3A) and 6-2 (BuKa3C) of Buran-Kaya III (see Supplementary Information). Radiocarbon (AMS) dates from associated material ranged from 36,260 to 35,280 cal BP (31,900+240-220 ¹⁴C BP) for BuKa3A and 37,415 - 36,245 cal BP (32,450+250-230 ¹⁴C BP) for BuKa3C (Fig. S1 and Table S1). Initial genetic characterization of the samples revealed extremely poor preservation of endogenous DNA, both in fragment length (average 36 bp) and in low quantity relative to environmental DNA (<0.1%). Applying various pretreatments and extraction buffers^{27,28} (see Supplementary Material for details) to the samples, we identified conditions that considerably increased the endogenous DNA content enough to allow genomic analysis (~0.5% for BuKa3A and ~4% for BuKa3C). Enrichment for mitochondrial DNA and shotgun sequencing resulted in 82- (BuKa3A) and 11- (BuKa3C) fold coverage of the mitogenomes. Reads mapping to the mitochondrion were used to determine the posterior probability for contamination with modern human sequences using the program Schmutzi²⁹, which estimated a contamination rate below 1% for BuKa3A, and haplocheck³⁰ detected no mitochondrial contaminants above 1% in either sample (see Methods). Both individuals were determined to be male using the ratios of reads mapping to chromosome X and Y³¹ and chromosome X-mapped reads to the average of autosomal-mapped reads³². The older sample, ~36,800 year-old BuKa3C, belonged to a mitochondrial haplogroup basal to haplogroup U, a lineage not found in modern humans and branching prior to the split that gave rise to U-haplogroups found in E/MUP Europe and present in Kostenki 14, the Sunghir group, Dolni Věstonice in Czechia, and Peștera Muierii, 34,000 year-old remains found in Romania³³ (Fig. S2). The mitochondrial haplogroup of BuKa3A was determined to

belong to an early branch of the N lineage, N1. This assignment falls outside of the lineages previously reported for N in UP Europe, nearly all of which derive from later N branches (U and R haplogroups, Fig. S2). The N1 of BuKa3A is notably distinct from the N haplogroups identified from the 40,000-year-old Oase1 from Romania and those of the 45,000-year-old occupants of Bacho Kiro Cave in Bulgaria, all of which belong to more basal branches that have no modern descendants^{6,34}. Additionally, the BuKa3A mitochondrial DNA sequence contains three of the eight mutations beyond N1 to N1b, a modern rare haplogroup most highly concentrated in the Near East, yet appearing broadly from western Eurasia to Africa. Among ancient samples, the mitochondrial sequence of an 11,000-year-old Epipalaeolithic Natufian from the Levant (Natufian9)³⁵ is also a later derivative of this pre-N1b branch. The Y haplogroup for BuKa3C was determined to belong to a basal F branch that has not been found among modern humans younger than 38,000 years-old, but clustering instead with Y haplogroups previously reported in southeastern Europe for the IUP BK (Bacho Kiro)_F6-620¹ and 40,000 year-old Oase1⁶ (Fig. S3). All three of these Y chromosomes carry ancestral bases at diagnostic sites that exclude their assignment to any modern F lineages, which are principally found in East Asia. A single transversion may also point to a similar basal F haplogroup for BuKa3A but the lower coverage of this sample does not allow a reliable placement beyond haplogroup CT.

Neanderthal admixture

Neanderthal settlements attributed to the Micoquian, Kiik-Koba type were also found at Buran-Kaya III (layer B) and have been dated through faunal bone fragments from 43,200 to 41,000 cal BP¹², demonstrating an earlier Neanderthal presence at the site up to the time of the CI eruption. AMH remains dating to this period found in Romania and Bulgaria have documented local admixture with late Neanderthals in southeastern Europe^{1,6}, leading us to investigate whether admixture with local Neanderthals in Crimea could be detected in early AMHs living at most several thousand years after the last Neanderthals in the region^{12,36}. Neanderthal ancestry calculated using the direct f₄-ratio method³⁷ determined that BuKa3A and BuKa3C possess 3.8 (SE 3.3-5.3) and 2.2 (SE 1.2-3.2) percent Neanderthal ancestry, respectively (Fig. S4). Although the Neanderthal ancestry was slightly higher in BuKa3A than BuKa3C, it is in line with that calculated for other post-CI individuals from Europe using the same method, indicating Neanderthal populations had likely disappeared from the region prior

to the arrival of AMHs ancestral to the population to which the Buran-Kaya III individuals belonged.

Admixture among UP humans in Europe

To maximize the comparable data between BuKa3 and other Paleolithic genomes, single nucleotide polymorphisms (SNPs) were called from 929 high-coverage modern genomes³⁸ and filtered to remove population-specific variants and those occurring in repetitive regions. This left ~13 million genome-wide SNPs, of which 265,510 and 537,076 overlapped with BuKa3A and BuKa3C respectively. The combined ~740,000 SNPs that occurred in at least one BuKa3 individual were then called from 38 previously published shotgun or high-coverage captured western Eurasian genomes from the Paleolithic to the Neolithic (Table S2). Outgroup f_3 -statistics were carried out on sample combinations including a minimum of 10,000 shared SNPs. Overall, both BuKa3 genomes were more similar to genomes younger than 38,000 years than to those older, and were more similar to modern European populations than those outside Europe (Fig. S5). Among the Paleolithic genomes available for analysis, both BuKa3 individuals were found to share the highest amount of drift with the roughly contemporary ~35,000-year-old GoyetQ116-1 from Goyet Cave in Belgium³ and the younger ~31,000-year-old infant twins from Krems-Wachtberg in Austria, KremsWA, which were buried with a Gravettian assemblage³⁹. Although both had a similar genetic distance to the 34,000 year-old individual from Peștera Muierii in Romania and the 35,000 year-old BK1653 from Bacho Kiro in Bulgaria (Fig. 1A, Table S3), the two BuKa3 individuals did not belong to a uniform population, and shared unequal amounts of drift with several individuals older than 30,000 years. BuKa3A showed a higher affinity than BuKa3C to the 40,000-year-old genomes of Oase1 and Tianyuan from East Asia⁴⁰, and also had a slightly greater affinity with several IUP Bacho Kiro individuals (Fig. 1A). Notably, BuKa3A, but not BuKa3C, shared a similar amount of drift with the pre-CI Oase1 as with more contemporary post-CI genomes from Romania and Bulgaria (Table S3). These results, and in particular the higher affinity of the East Asian Tianyuan to BuKa3A than BuKa3C, indicate that, similar to GoyetQ116-1, BuKa3A is also likely to contain ancestry present in Europe prior to the Heinrich 4 Stadial and the CI eruption. BuKa3A also shared more drift than BuKa3C with the ~6,000 year younger Věstonice16 from central Europe, associated with a Gravettian assemblage³, whereas BuKa3C had a slightly higher genetic affinity to the ~38,000-year-old Kostenki14 from the eastern European plain, and a lower

affinity to post-Last Glacial Maximum (LGM) populations from central and western Europe (Fig. 1A.). These $f\beta$ results were explored in greater detail through Multi-Dimensional Scaling (MDS) and plotted on a 3-D scatter plot (Fig. 1B. Fig. S6), which can be used to visualize the global genomic relationships of the ancient populations of Eurasia. In this analysis, the two BuKa3 samples occupy a perimeter of the E/MUP cluster of Europe composed of genomes dating between 38,000 and 29,000 years ago. This cluster corresponds to a wave of ancestry that first occurred in Europe at some point following Heinrich Event 4 including the CI eruption, and appears to have largely replaced the genomically distinct earlier AMH inhabitants of Europe associated or contemporary with IUP technologies^{1,3-5}. Of the E/MUP Europe cluster, the two individuals plotting the farthest along the D3 axis, which differentiates the older pre-CI individuals, are GoyetQ116-1 and BuKa3A, supporting a degree of shared drift between these and the oldest Eurasian AMH genomes.

We attempted to model the possible UP population movements and admixtures involved in the early AMH settlement of Europe using a qpGraph framework that could explain many of the relationships we identified. The best-fitting model with no outlier $f4$ -statistics using the modern African Mbuti population as an outgroup, and having the lowest absolute Z-score of 2.6 is shown in Fig. 2A. This model describes the migration of groups into Europe following Heinrich Event 4 and the CI eruption as having occurred in various waves, originally deriving from a population or populations with a common ancestor that had separated much earlier from the pre-CI ancestry found across Eurasia and represented in Europe by the IUP Bacho Kiro group. The first of these incoming populations admixed with the pre-existing inhabitants, some retaining as much as 18% of the pre-CI ancestry several thousands of years later. GoyetQ116-1, BK1653, and the BuranKaya3 group later descended from these admixed populations. These admixture events appear to be constrained to central and western Europe, since notably, groups in eastern Europe, such as those ancestral to Kostenki14 and Sunghir3, lack a pre-CI ancestry component, detectable either by affinity with East Asian genomes or pre-CI individuals in Europe. Later unions among admixed and unadmixed groups are adequate to explain much of the ancestry identified to date in UP Europe. BuKa3C can be described as descended from a population derived from one ancestral to Kostenki14 and having 17% ancestry from an IUP Bacho Kiro-like group (denoted as “BuKa- α ” in Fig. 2A). BuKa3A could be successfully modeled as having descended from “BuKa- β ”, a population resulting from admixture between “BuKa- α ” and another admixed group ancestral to GoyetQ116-1. This model, although a simplification that does not address all pre-CI-related sources found in

BuKa3A, accounts for the affinities revealed through f_3 - and f_4 -statistics between BuKa3C and Kostenki14, BuKa3A and GoyetQ116-1, both BuKa3C and BuKa3A with KremsWA (Fig. 2B), and the divergent affinities between the BuKa3 individuals and those from IUP Bacho Kiro. The MUP individual from Bacho Kiro, BK1635, is modeled as having derived from an early-arriving population ancestral to the eastern Europe Sunghir3 and Kostenki14 lineages, which admixed with IUP Bacho Kiro-like ancestry, in a process parallel to that which gave rise to GoyetQ116-1. BK1635 is representative of the most likely candidate population to serve as the western source of the Věstonice group, which would later receive additional gene flow from a Sunghir-like population, presumably from the east. The placement of Muierii as a mixture between BK1653-like ancestry and an unsampled population ancestral to Sunghir3 and Kostenki14 gave the only successful model including this sample, and refines the previous placement reported for this individual as a mixture of a pre-Sunghir-like ancestry and a sister group to GoyetQ116-1³³. Post-LGM Epi-Gravettian individuals from northern Italy such as the 17,000-year-old Tagliente⁴¹ (or later Villabruna³) were modeled as a lineage from the same source population as GoyetQ116-1 although with a considerable amount of temporal and genetic distance. This placement agrees roughly with previous models^{3,10,41,42}, but it is evident that there is a need for more clarification considering the genetic history of the ancestors of post-LGM populations in Europe.

Pre-CI gene flow into Buran-Kaya III

To date, two incidents of ancestry present in Europe before 40,000 years ago appearing 5,000 years later in Europe have been reported (GoyetQ116-1 and BK1653)^{1,3}, although little genomic data exist from the earliest period of these admixtures that could assist in better characterizing these events. To investigate this aspect further, we compared BuranKaya3 and other pre-LGM individuals through f_4 -statistics. Overall, BuKa3C was found to be more genetically similar to contemporary and younger groups than older ones, sharing significantly more alleles with any individual younger than 38,000 years when compared with any individual older than 40,000 years ($(f_4(EUP/MUP, IUP; BuKa3C, Mbuti), Z = 3.4 - 11.9)$, Fig. 3A and 3B, Table S4). While BuKa3C was equally distant from most pairs of pre-CI individuals, several significant f_4 -statistical results indicate that BuKa3C contained at least some ancestry found in earlier inhabitants of Europe, showing heightened affinities to Oase1 and BK_CC7_335 in some contexts (Fig.3A). The presence of preCI ancestry in BuKa3A was more pronounced.

Similar to BuKa3C, BuKa3A showed a significantly higher affinity to groups younger than 38,000 years compared to older groups, unless the test involved Oase1 (Fig. 3B), where allele sharing between BuKa3A and Oase1 was found to reduce the affinity of BuKa3A to post-CI genomes. To verify whether this affinity could not be due to elevated Neanderthal content in BuKa3A and Oase1, we recalculated the statistic after masking the increased Neanderthal content found in the Oase1 genome⁵. Excluding excess Neanderthal ancestry increased the shared drift between BuKa3A and Oase1, indicating that this affinity could be traced to shared non-Neanderthal ancestry (Table S5). Through comparisons of pre-CI groups with BuKa3A, significant allele sharing could also be seen with the 31-52,000-year-old woman from Czechia, Zlatý Kůň⁴ with respect to either IUP Bacho Kiro individuals or the 45,000-year-old Ust-Ishim from western Siberia⁴³. In multiple f_4 -statistics, Zlatý Kůň, Oase1, and Tianyuan were found to share roughly the same proportion of alleles with BuKa3A (Table S4). MDS analysis of inverted f_3 -statistics including only UP samples also placed BuKa3A between fellow post-CI individuals and a pre-CI group comprising Oase1 and Zlatý Kůň (Fig. S7). The identification of a shared ancestry between BuKa3A and Oase1 was unexpected since Oase1 ancestry had not previously been detected in any subsequent Eurasian populations^{1,3,6}. A recent analysis of UP genomes, however, could successfully model Oase1 as belonging to a population similar to the IUP Bacho Kiro group with additional Neanderthal admixture⁵, and groups related to Bacho Kiro have contributed ancestry to modern East Asians and GoyetQ116-1^{1,3}. Zlatý Kůň had been previously shown not to have contributed gene flow to later groups⁴, which is contradicted by the allele sharing demonstrated here between Zlatý Kůň and BuKa3A relative to Ust-Ishim. More recent analyses, however, presented admixture models that can support gene flow of between 2-29% from Zlatý Kůň into Bacho Kiro⁵, which would support our observations.

Genetic relationships with post-CI populations

Among the genomes between 38,000 and 30,000 cal BP, which represent the post-CI wave of ancestry arriving into Europe beginning in the EUP, BuKa3C shared similar amounts of drift with the contemporary Kostenki14 as with the younger GoyetQ116 and KremsWA, but shared more ancestry with the 5,000-year younger KremsWA from Central Europe than the more contemporary Sunghir3 to the northeast, revealing a key difference between the otherwise similar Kostenki14 and Sunghir3 genomes (Fig. 2B).

The Věstonice genetic cluster has been associated with Central European Gravettian assemblages³. Two Central European Gravettian burials, Věstonice16 and KremsWA, share an excess of alleles relative to other MUP individuals, yet they behaved differently when tested with respect to both BuKa3 individuals. Although Věstonice16 shows higher affinity towards KremsWA, Sunghir3, BK1653, GoyetQ116-1, and Kostenki14 with respect to one or both BuKa3 genomes, KremsWA could not similarly differentiate any contemporaries from the BuKa3 individuals (Fig. 2B), indicative of gene flow from a BuKa3-like population to KremsWA which is also supported by the high outgroup $f\beta$ -values observed between KremsWA and BuKa3 (Fig. 1A, Table S3). These results show that while the makers of Gravettian assemblages of central Europe exhibited an overall genomic similarity, they can be distinguished by a differential presence of ancestry found thousands of years earlier on the shores of the Black Sea.

Caucasian Hunter Gatherer connections

A Late Upper Paleolithic (LUP) population was identified from 13,300-year-old remains at Satsurbliia Cave in western Georgia, whose distinct ancestry was defined, along with similar nearby 9,800-year-old remains at Kotias Klde, as Caucasus Hunter Gatherer (CHG)⁴⁴, also known as the Satsurbliia genetic cluster³. CHG ancestry has been estimated to have separated from Western Hunter Gatherers (WHG) approximately 45,000 year ago and existed as an isolated population during the LGM, later contributing ancestry to several distinct populations that played roles in developing the modern European genomic identity^{35,44,45}. CHG ancestry has been identified in the same Caucasus region dating back to the MUP, making up roughly half of the ancestry of two 26,000-year-old individuals from Dzudzuana Cave⁴² and 25,000-year-old remains from Kotias Klde⁴⁶, both in Georgia. Given the geographical proximity and cultural links between the contemporary layers of these Caucasus sites and Buran-Kaya III, we additionally investigated whether continuity with the CHG ancestry could extend 10,000 years earlier at BuranKaya3. We show that one or both individuals from BuranKaya3 share significantly more alleles with the CHG cluster than any other post-CI individual older than 30,000 years ($Z = 2.2-6.6$), with the exception of KremsWA ($Z = 0.08-1.9$) (Fig. 3C, Table S4). The identification of CHG-related ancestry at Buran-Kaya III, which appears in early Gravettian-associated ancestry found 5,000 years later in Central Europe, highlights the Caucasus as a potential entry point in repopulating Europe beginning at least 38,000 years ago.

UP samples from the Caucasus and surrounding regions will be required to better resolve these movements and interactions of these groups before and after this transition period.

Discussion

With these two individuals from Buran-Kaya III we genetically characterize one of the earliest facets of the EUP wave of ancestry appearing in Europe between 40,000 and 38,000 cal BP at Crimea, and associate it with a backed-bladed, microlaminar lithic industry, well established in the northern and southern slopes of the Caucasus which has been compared to Gravettian assemblages appearing later in central Europe^{11,13,25}. Disparities in the ancestries of the two individuals, occupying the same place some 1,000 years apart, reflect the non-uniform nature of the relationships among these early migrant groups, and propose a model of complex interactions and exchange not only at the interface of this incoming ancestry, but continuing among their descended lineages. The earlier BuKa3C was minimally admixed with ancestry similar to one found within the IUP Bacho Kiro group and carried an IUP-associated Y haplogroup, whereas the later BuKa3A was more heavily admixed with several pre-CI populations. Both shared more ancestry with individuals living thousands of years later in western and central Europe than with any other contemporary groups found to date. This model reflects a broad pattern of incoming ancestries that had diverged from a similar source, mixing to different degrees with deeper-branching, pre-CI AMH populations who arrived earlier, and interacting with neighboring groups to form the earliest ancestries detectable in present-day Europeans. The elevated pre-CI admixture in the more recent BuKa3A compared to the ~700-years older BuKa3C is reminiscent of population dynamics following the Neolithic migration into Europe, where evidence of admixture increases some time following the initial appearance of new migrants³². The degree of admixture between these incoming AMHs with the pre-CI ancestry of the exiting inhabitants has not previously been appreciated. Of the seven genomes now available from Europe dating to between 38,000 and 34,000 years ago, all five located in southern, central, and western Europe can be modeled as having admixed with the pre-existing AMH groups (Fig. 2A). The regularity of this phenomenon suggests the possibility that the earliest AMH settlers in Europe, perhaps belonging to populations substantially reduced in size by the preceding climatic changes of the Heinrich 4 cold period and the CI eruption, may have been absorbed into the incoming EUP migration wave.

Identifying the geographical areas where these newly arrived populations developed, which had to have been isolated long enough to establish distinct genetic profiles will require a better genetic characterization of EUP populations of the surrounding regions: the Caucasus, eastern Europe, Anatolia and the Levant. Genetic similarities between BuKa3A and BuKa3C with CHG ancestry documented in the Caucasus 10,000 years later, coupled with a similar archeological horizon to that found across several contemporary northern Caucasus sites, argues for the Caucasus as the point of entry of Buran-Kaya III-linked ancestry, which is best modeled as branching from a source population leading to Kostenki-like ancestry after separating from the lineages that would give rise to GoyetQ116-1 and BK_1653, later found in central and western Europe, and Sunghir3 in the east. We also observe that the only groups unadmixed with pre-CI European ancestry, Sunghir3 and Kostenki14, date to before 34,000 years ago and are found only in eastern Europe, which indicate a possible eastern entry point for these lineages from an area that may have been devoid or sparsely inhabited by pre-CI European populations at the time.

The timing of this EUP movement into Europe, coinciding with the close of the lengthy Heinrich 4 period of extreme cold encompassing the CI super-eruption, suggests that climatic changes may have contributed to this migration. Eastern Europe in particular would have been less habitable between 40,000 and 38,000 years ago, and existing IUP populations within Europe may have contracted, as attested by the disappearance of European Neanderthals³⁶, the scarcity of human settlements, and the reduction of forest cover and ungulate populations during this time^{7,47}. Gradually increasing temperatures towards the end of the Heinrich Event 4 and the onset of the warmer Greenland Interstadial 8 (~38,000–35,000 BP)^{7,8} would favor northward population expansions from refuges in the south²⁶.

The population change in Europe during this period is also highlighted by the introduction of EUP lithic innovations and cultural practices. Backed blades and bladelet industries begin to appear in Europe^{48,49}, different from the Levallois blade production techniques characteristic of the IUP⁵⁰, accompanied by an increased diversity in the raw materials incorporated into personal adornments and portable art^{12,48}. An awareness of past population dynamics can deepen the interpretation of the existing archeological record and assist in reconstructing the development and dispersal processes of technologies and artefacts associated with these populations. In this light, it is compelling to contrast the genetic similarities between the BuKa3 group and KremsWA with the similarities between the backed-blade and micro-laminar lithics associated with these remains, as well as the collection of cultural artefacts found at Buran-

Kaya III and contemporary sites in the Caucasus, with Gravettian assemblages appearing later in central Europe. With such a large gap in the archeological record between these sites, in both time and geography, it is difficult to estimate the degree to which these similarities may be due to a continuity of EUP practices at Buran-Kaya III being maintained through tradition in descendant communities over thousands of years. On the other hand, given the interactions between MUP groups in Europe documented through genetic exchange, it is easy to imagine multiple paths through which technical innovations can be communicated across Europe. Similar local and climate-specific functional requirements may also play a role in the correlation of lithic industries²⁴. Addressing the links between central European Gravettian assemblages and the EUP backed microblade industries of Buran-Kaya III and the northern Caucasus, and perhaps also by extension the Early Ahmarian of the Levant as previously suggested²⁶, may complement our growing understanding of EUP population movements into Europe.

Methods

Dating

All radiocarbon dates were recalibrated using the software OxCal v4.4.4 based on the IntCal20 calibration data set⁵¹. The calibrated dates are rounded to 5.

Sample handling, DNA extraction, library construction, and sequencing

Three human parietal fragments were excavated aseptically or semi-aseptically from layer 6-1 and 6-2 during the 2009 excavation season at Buran-Kaya III. All pre-amplification sample preparation was performed in the dedicated ancient DNA facility using decontamination and clean-room protocols as described in Bennett, *et al.*⁵². All buffers and solutions were prepared using water decontaminated by gamma-irradiation (8 kGy). The surface of the bone was wiped three times with a cotton tip soaked in water, then bleach, then water again, and a small part was removed using a flame-sterilized diamond-disc of a Dremel saw (Dremel, Europe, The Netherlands). Bone fragments were then ground to fine powder in a 6775 Freezer/Mill®Spex SamplePrep in liquid nitrogen as described previously³². Powder from each sampling was divided into two equal portions, one of which was subject to a bleach treatment with a 1:20 dilution of hypochlorite followed by a washing step with water, as described²⁷. Phosphate buffer washes, described in Korlević²⁷, were collected and combined for DNA purification. Both

phosphate and non-phosphate buffer treated samples, including reagent-only mocks, were then incubated in 1.5 mL LoBind microcentrifuge tubes (Eppendorf, Hamburg, Germany) with 1 mL 0.5M EDTA, pH 8.0 (Sigma-Aldrich, St. Louis, MO), with 0.25mg/mL proteinase K (Sigma-Aldrich), 0.14M beta-mercaptoethanol and 0.05% UV-irradiated Tween-20 (Sigma-Aldrich), at 37°C for 24 H. This step was repeated for another 24 H and the two fractions were pooled. Following incubation, all tubes were centrifuged at maximum speed for 10 min, and 1 mL supernatant was mixed with 10 times its volume of “2M70” binding buffer (2M guanidine hydrochloride and 70% isopropanol) in a 15 mL tube and passed through QIAquick silica columns (Qiagen, Hilden, Germany) using 25 mL tube extenders (Qiagen) and a vacuum manifold (Qiagen) as described⁵². 2M70 binding buffer has been shown to retain the smaller DNA fragments lost during purification with traditional binding buffers²⁸. Columns were washed twice with 1 mL PE Buffer (Qiagen) then transferred to a micro-centrifuge and dried by spinning 1 minute at 16,100 × g, turning tubes 180° and repeating. DNA was eluted in a total of 60 µl of 10mM EBT (Tris-HCl pH 8.0 containing 0.05% Tween-20) performed in two elution steps of 30 µl each, by spinning 16,100 × g for 1 minute after a 5-minutes incubation. The second 1 mL of extract was purified according to the same protocol but using 5M40 “binding buffer” (5M guanidine hydrochloride and 40% isopropanol). 5M40 binding buffer has been shown to remove more inhibitors than the 2M70 binding buffer but to result in a higher loss of smaller DNA molecules²⁸.

For the first series of screening steps, single-stranded libraries were constructed using either 2 µL (for screening) or 6 µL (for mitochondrial capture) of the eluted DNA, including mocks of all treatments, water only samples, and a positive control oligo following the protocol of Gansauge, *et al.*⁵³ using the splinter oligonucleotide TL110, and eluting in 50 µL EBT. Either 40 µL (for mitochondrial capture) or 4 µL (for screening) of each library was used for bar-coding amplification using dual-barcoded single-stranded library adapters⁵⁴ as primers in the following 100 µL volume reaction: 10 µL 10x PCR Buffer + MgCl₂ (Roche, Basel, Switzerland) 0.4 µM of each primer, 80 µM dNTPs (Roche), and 15 units of FastStart Taq (Roche). Reactions were heated 95°C for 5 minutes, followed by 35 cycles of 95°C 20 s, 53°C for 45 s, 68°C for 45 s, and then 68°C for 5 minutes. Heteroduplexes that could confound size selection were resolved by diluting the PCR product 1:5 in a 100 µL reaction containing 20 µL of the initial reaction, 8 µl PCR Buffer + MgCl₂, 0.4 µM of standard Illumina primers P5 and P7, and 80 µM dNTPs, and amplified a single cycle of 95°C for 1 minute, 60°C for 2 minutes and 68°C for 5 minutes. Products were then purified and size-selected using NucleoMag beads (Macherey-Nagel, Düren, Germany) for two rounds of purification/size selection according to the supplied

protocol at a ratio of bead solution 1.3 times the reaction volume and eluted in 30 μ l EBT. Purified libraries were quantified using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA), Bioanalyzer2100 (Agilent, Santa Clara, California, USA), Qubit 2.0 Fluorometer (Thermo Fisher Scientific), and qPCR reaction. 46 to 148 ng of DNA were enriched for human mitochondrial sequence in two rounds of capture using 1200 ng of biotinylated RNA baits reverse transcribed from human mitochondrial PCR products³² following the protocol described in Massilani *et al.*⁵⁵ but with four changes: (1) for hybridization and wash steps, 60°C was used instead of 62°C. (2) DNA/RNA-bait solution was incubated 96 hours instead of 48 hours. (3) Elution of the bead-bound enriched DNA was performed with a 5 minute incubation in 30 μ L EBT at 95°C followed by a magnetic bead separation and the transfer of the eluate to a new tube rather than a 0.1N NaOH elution followed by silica column purification. (4) All post-capture amplifications were performed for 35 cycles followed by a heteroduplex resolution step as described above. Enriched DNA was then quantified as above, and products from all libraries were pooled in equimolar amounts and sequenced on an Illumina MiSeq using a v3 reagent kit for 2x76 cycles, substituting primer CL72 for the Read1 sequencing primer as described⁵⁴.

Eight additional libraries along with positive and negative controls were made from 3-8 μ L each of the remaining extract BuKa3A_4B, which had the highest relative endogenous DNA content. It should be noted that this extract was derived from the portion of the cranial fragment which included the cranial suture. Libraries were prepared as described above except one library was first treated with USER enzyme (New England Biolabs, Ipswich, Massachusetts, USA) for 30 minutes to remove deaminated cytosine damage. Prior to the barcoding amplification, a 6-cycle amplification of pre-barcoded libraries was performed using 45 μ L of each library with 45 μ L OneTaq 2X Master Mix with Standard Buffer (New England Biolabs), and 0.1 μ M internal primers CL72⁵⁴ and CL130⁵³ with the above PCR conditions. Three pairs of different dual-barcoded adapters were then used to amplify 20 μ L of each amplified library, followed by purification and size selection as described above, which allowed the later pooling of two different enrichment protocols and non-enriched DNA from the same library. An average of 1.2 μ g of DNA to 1 μ g RNA-baits was used for each mitochondrial enrichment, as above, however, an additional alternative “touch-down” hybridization protocol consisting of 60°C for 12 h, 59°C for 12 h, 58°C for 12 h, 57°C for 12 h, and 56°C for 48 h was tested for each library. For selected libraries, an alternative wash protocol described in Fu *et al.*⁴⁰ was also tested. Neither of these alternative protocols had a substantial impact on the results obtained. Enriched and shotgun libraries were then pooled separately and size selected on an E-Gel

SizeSelect 2% agarose gel (Thermo Fisher). Enriched libraries were then sequenced on an Illumina MiSeq, as above, and shotgun libraries were sequenced on an Illumina NextSeq 500/550 High Output Kit v2 (2x75 cycles). The shotgun libraries were further gel-purified on a 6% polyacrylamide gel and sequenced again on the NextSeq as above.

Additional libraries from fresh extractions of BuKa3A and BuKa3C samples were constructed using 6 μ L of extract and the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, Ipswich, MA, USA) after a pretreatment with USER enzyme mix (NEB, Ipswich, MA, USA) as described by the manufacturer. An unindexed Y-shaped adapter was ligated, the ligated libraries were first amplified for 8 cycles with the Illumina Rd1 and Rd2 sequencing primers, then 20% of the library was further amplified for 12 cycles with Illumina unique dual indexed adapter primers. The libraries were purified independently using 1.3 volume of SPRI beads (Macherey-Nagel™ NucleoMag™ NGS Clean-up and Size Select) for two successive rounds of purification followed after pooling of three successive rounds of purification with SPRI beads as above. Sequencing was performed on an Illumina NovaSeq 6000 sequencer in paired-end mode.

Processing of Sequencing Data

All data processing steps were performed using shell instructions and in-house written scripts. Demultiplexing and fastq conversion of the raw output of Illumina sequencers in the BCL format was performed using `bcl2fastq` (Illumina). Sequence quality was assessed using `FastQC` (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Paired-end reads were merged and trimmed for adapters using `leeHom`⁵⁶. Reads larger than 28 base pairs (bp) were selected using `awk` and mapped onto the Human Reference Genome `hs37d5` using `BWA` (Burrows-Wheeler Aligner) version 0.7.17⁵⁷ with a seed length of 18. A minimum length of 28 bp was chosen following analysis to identify the size limit minimizing spurious mapping reads following the approach described in⁵⁸ (Fig. S8). Aligned sequences in BAM format were processed with `Samtools` v.1.9.⁵⁹ to create index files and `flagstat` reports, and to select and sort mapped reads. `Picard` tools (<http://broadinstitute.github.io/picard/>) was used to remove duplicates. Multiple libraries were first merged by `Samtools` v.1.9, followed by filtering for mapping quality (`-q 20`). Reads smaller than 35 bp containing indels were then removed using an `awk` script to minimize spurious alignments⁶⁰. Damage patterns were analyzed using `mapDamage` v.2.2.1⁶¹ (Fig. S9) followed by end-clipping using `BamUtil` (https://genome.sph.umich.edu/wiki/BamUtil:_trimBam). According to the damage pattern,

three to four bases at the ends of each read were clipped for the UDG-treated libraries and 10 bases for non-UDG treated libraries. Library details are given in Table S6.

Sex Identification, Uniparental Haplogroups and Contamination Estimation

The sex of the BuKa individuals was determined by chromosome counting using Qualimap²⁶² and Samtools v.1.9 idxstats using reads mapped with a mapping quality higher than 20. Haplogrep²⁶³ and Haplocheck³⁰ were used to classify the mitochondrial haplogroups and to estimate the contamination level. Schmutzi²⁹ was additionally used to estimate contamination of the higher coverage mitochondrial reads of BuKa3A. Mitochondrial haplogroups were verified manually. A Bayesian tree of 35 ancient, 1 modern and 2 canonical mitochondrial sequences, excluding the hypervariable regions, was constructed using MrBayes (v3.2.6)⁶⁴ using a HKY+i+G nucleotide substitution model, which gave the highest log-likelihood for the tree out of all models tested (GTR and HKY with all combinations of four invariant sites and gamma distributions) using JModelTest2 (v2.1.10)⁶⁵. The chain was run for 1,100,000 iterations, subsampled every 200 after discarding a 9% burn-in period and visualized using FigTree v. 1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree>). A list of sources for the mitochondrial sequences is given in Table S7. Y chromosome haplogroups analysis was performed with a subset of reads mapping to the Y chromosome with a mapping quality 37 or above, and with the terminal two bases trimmed from each end. Y haplogroups were determined by placing alleles onto reference phylogenies to find the lowest supported downstream node using Pathphynder⁶⁶.

Dataset Preparation and Population Genetics Analyses

In order to prepare a SNP panel, the GRCh38/hg38 genomic coordinates of genotype calls coming from 929 publicly available genomes (<ftp://ngs.sanger.ac.uk/production/hgdp>)³⁸ were lifted over the assembly of GRCh37/hg19 by using Picard tool (<http://broadinstitute.github.io/picard/>) based on the UCSC liftOver tool (<http://genome.ucsc.edu/cgi-bin/hgLiftOver>). The variants coming from repetitive regions of the human genome were filtered out with BEDTools⁶⁷ from genotype calls that were lifted over. The heterozygous positions from extracted Western Eurasian and Mbuti populations were filtered by removing the singletons from each population, which resulted in ~13 million biallelic autosomal variants. ~740,000 positions that overlapped with at least one of the BuKa3 genomes were used to curate a 740k SNP panel with a pseudo-haploid SNP calling approach of each

ancient genome using pileupCaller (<https://github.com/stschiff/sequenceTools>). The curation of another SNP panel was carried out for BuKa genomes by HaplotypeCaller tool in GenomeAnalysisTK v.4.1.9⁶⁸ using the variant panel of HGDP 929 genomes which led to ~1.7 million biallelic autosomal positions after exclusion of heterozygous positions coming only from BuKa3.

For the Neanderthal ancestry analysis, the gvcf files of three high-coverage Neanderthal genomes, Altai, Vindija33.19, and Chagyrskaya were obtained from the website of the MPI of Evolutionary Anthropology in Leipzig (<http://cdna.eva.mpg.de/neandertal/>). The three genomes were merged using bcftools 1.9⁶⁹. SNPs that were invariant in any of the Neanderthals and those residing in repeated DNA (as defined by the hg19 repeat mask from <https://www.repeatmasker.org/>) were removed. The overlapping sites between these positions and the 2200k SNP panel (described in Hajdinjak *et al.*¹), amounting to 1.3 M sites, were used to call chimp sequences downloaded from ENA PRJEB36949 and mapped onto the hs37d5 human genome using bwa mem⁵⁷ deduplicated and filtered for a mapping quality higher than 30. SNP calling was performed using bcftools mpileup (options -B -q30 -Q30) and call (option -m). Five Mbuti genome sequences (SRR1157019, SRR1157021, SRR1157023, SRR1157024, and SRR1157055) were processed in the same way. All vcf files were merged and filtered to remove invariant, missing and multiallelic sites, which left ~800k autosomal positions that were used to call ancient genomes, including the two BuKa individuals. All files were then merged together for the estimation of the Neanderthal ancestry.

The outgroup- f_3 and f_4 statistics and admixture graph were computed using qp3P, qpDstat and qpGraph programs from the ADMIXTOOLS package (version 5.0)⁷⁰, respectively. The outgroup- f_3 statistics were calculated in a pairwise manner in the form of $f_3(\text{Mbuti}; \text{ind1}, \text{ind2})$ for 38 individuals. A distance matrix coming from inverted pairwise values $(1-f_3)$ was used to compute Multi-Dimensional Scaling (MDS) analysis (cmdscale package in R). To estimate the Neanderthal ancestry a direct f_4 -ratio was used in the form of $\alpha = (\text{Vindija33.19}, \text{Chimp}; \text{Ancient_test}, \text{Mbuti}) / (\text{Vindija33.19}, \text{Chimp}; \text{Chagyrskaya}, \text{Mbuti})$.

Data and materials availability

Sequence data generated in this study are available through the NCBI Sequence Read Archive SRA ID PRJNAXX and XXX. The mitochondrial sequences are available through GenBank (accession number MKXXX and XXX).

References

1. Hajdinjak, M. *et al.* Initial Upper Palaeolithic humans in Europe had recent Neanderthal ancestry. *Nature* **592**, 253–257 (2021).
2. Slimak, L. *et al.* Modern human incursion into Neanderthal territories 54,000 years ago at Mandrin, France. *Sci. Adv.* **8**, eabj9496 (2022).
3. Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
4. Prüfer, K. *et al.* A genome sequence from a modern human skull over 45,000 years old from Zlatý kůň in Czechia. *Nat. Ecol. Evol.* **5**, 820–825 (2021).
5. Vallini, L. *et al.* Genetics and Material Culture Support Repeated Expansions into Paleolithic Eurasia from a Population Hub Out of Africa. *Genome Biol. Evol.* **14**, evac045 (2022).
6. Fu, Q. *et al.* An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015).
7. Fedele, F. G., Giaccio, B., Isaia, R. & Orsi, G. The Campanian Ignimbrite Eruption, Heinrich Event 4, and palaeolithic change in Europe: A high-resolution investigation. in *Geophysical Monograph Series* (eds. Robock, A. & Oppenheimer, C.) vol. 139 301–325 (American Geophysical Union, 2003).
8. Fitzsimmons, K. E., Hambach, U., Veres, D. & Iovita, R. The Campanian Ignimbrite Eruption: New Data on Volcanic Ash Dispersal and Its Potential Impact on Human Evolution. *PLoS ONE* **8**, e65839 (2013).
9. Seguin-Orlando, A. *et al.* Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113–1118 (2014).
10. Sikora, M. *et al.* Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science* **358**, 659–662 (2017).
11. Prat, S. *et al.* The Oldest Anatomically Modern Humans from Far Southeast Europe: Direct Dating, Culture and Behavior. *PLoS ONE* **6**, e20834 (2011).
12. Péan, S. *et al.* The Middle to Upper Paleolithic Sequence of Buran-Kaya III (Crimea, Ukraine): New Stratigraphic, Paleoenvironmental, and Chronological Results. *Radiocarbon* **55**, 1454–1469 (2013).
13. Yanevich, A. Les occupations gravettiennes de Buran-Kaya III (Crimée): contexte archéologique. *L'Anthropologie* **118**, 554–566 (2014).
14. Prat, S. *et al.* The First Anatomically Modern Humans from South-Eastern Europe. Contributions from the Buran-Kaya III Site (Crimea). *Bull. Mém. Société Anthropol. Paris* (2018) doi:10.3166/bmsap-2018-0032.
15. Hublin, J.-J. The modern human colonization of western Eurasia: when and where? *Quat. Sci. Rev.* **118**, 194–210 (2015).
16. Reynolds, N. & Green, C. Spatiotemporal modelling of radiocarbon dates using linear regression does not indicate a vector of demic dispersal associated with the earliest Gravettian assemblages in Europe. *J. Archaeol. Sci. Rep.* **27**, 101958 (2019).

17. Reynolds, N. Threading the Weft, Testing the Warp: Population Concepts and the European Upper Paleolithic Chronocultural Framework. in *Culture History and Convergent Evolution* (ed. Groucutt, H. S.) 187–212 (Springer International Publishing, 2020). doi:10.1007/978-3-030-46126-3_10.
18. Reynolds, N. The Gravettian is Dead: Against Equivocation and Reification in Chronocultural Studies of the Upper Palaeolithic. in *Les sociétés gravettiennes du Nord-Ouest européen: nouveaux sites, nouvelles données, nouvelles lectures / Gravettian societies in North-western Europe: new sites, new data, new readings* (eds. Touzé, O., Goutas, N., Salomon, H. & Noiret, P.) 309–321 (Presses Universitaires de Liège, 2021).
19. Demidenko, Y. E. Crimean Upper Paleolithic. in *Encyclopedia of Global Archaeology* (ed. Smith, C.) 1782–1791 (Springer New York, 2014). doi:10.1007/978-1-4419-0465-2_1864.
20. Cullen, V. L. *et al.* A revised AMS and tephra chronology for the Late Middle to Early Upper Paleolithic occupations of Ortvale Klde, Republic of Georgia. *J. Hum. Evol.* **151**, 102908 (2021).
21. Golovanova, L. V. *et al.* The Early Upper Paleolithic in the northern Caucasus (new data from Mezmaiskaya Cave, 1997 excavation). *Eurasian Prehistory* **4**, 43–78 (2012).
22. Bar-Yosef, O. *et al.* Dzudzuana: an Upper Palaeolithic cave site in the Caucasus foothills (Georgia). *Antiquity* **85**, 331–349 (2011).
23. Hoffecker, J. F. A New Framework for the Upper Paleolithic of Eastern Europe. (2012).
24. Hoffecker, J. F. & Holliday, V. T. Landscape Archaeology and the Dispersal of Modern Humans in Eastern Europe. in *The Upper Paleolithic of Northern Eurasia and America: Sites, Cultures, Traditions* 140–158 (Petersburg 2014).
25. Svoboda, J. A. The Gravettian on the Middle Danube. *PALEO Rev. Archéologie Préhistorique* **19**, 203–220 (2007).
26. Hoffecker, J. F. The early upper Paleolithic of eastern Europe reconsidered. *Evol. Anthropol. Issues News Rev.* **20**, 24–39 (2011).
27. Korlević, P. *et al.* Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. *BioTechniques* **59**, (2015).
28. Glocke, I. & Meyer, M. Extending the spectrum of DNA sequences retrieved from ancient bones and teeth. *Genome Res.* **27**, 1230–1237 (2017).
29. Renaud, G., Slon, V., Duggan, A. T. & Kelso, J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* **16**, (2015).
30. Weissensteiner, H. *et al.* Contamination detection in sequencing studies using the mitochondrial phylogeny. *Genome Res.* **31**, 309–316 (2021).
31. Skoglund, P., Storå, J., Götherström, A. & Jakobsson, M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J. Archaeol. Sci.* **40**, 4477–4482 (2013).
32. Brunel, S. *et al.* Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proc. Natl. Acad. Sci.* **117**, 12791–12798 (2020).
33. Svensson, E. *et al.* Genome of Peștera Muierii skull shows high diversity and low mutational load in pre-glacial Europe. *Curr. Biol.* **31**, 2973-2983.e9 (2021).

34. Hublin, J.-J. *et al.* Initial Upper Palaeolithic Homo sapiens from Bacho Kiro Cave, Bulgaria. *Nature* **581**, 299–302 (2020).
35. Lazaridis, I. *et al.* Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).
36. Higham, T. *et al.* The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature* **512**, 306–309 (2014).
37. Petr, M., Pääbo, S., Kelso, J. & Vernot, B. Limits of long-term selection against Neanderthal introgression. *Proc. Natl. Acad. Sci.* **116**, 1639–1644 (2019).
38. Bergström, A. *et al.* Insights into human genetic variation and population history from 929 diverse genomes. *Science* **367**, eaay5012 (2020).
39. Teschler-Nicola, M. *et al.* Ancient DNA reveals monozygotic newborn twins from the Upper Palaeolithic. *Commun. Biol.* **3**, 650 (2020).
40. Fu, Q. *et al.* DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl. Acad. Sci.* **110**, 2223–2227 (2013).
41. Bortolini, E. *et al.* Early Alpine occupation backdates westward human migration in Late Glacial Europe. *Curr. Biol.* **31**, 2484–2493.e7 (2021).
42. Lazaridis, I. *et al.* *Paleolithic DNA from the Caucasus reveals core of West Eurasian ancestry.* <http://biorxiv.org/lookup/doi/10.1101/423079> (2018) doi:10.1101/423079.
43. Fu, Q. *et al.* Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445–449 (2014).
44. Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* **6**, 8912 (2015).
45. Jeong, C. *et al.* The genetic history of admixture across inner Eurasia. *Nat. Ecol. Evol.* **3**, 966–976 (2019).
46. Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).
47. Hoffecker, J. F. *Desolate landscapes: Ice-Age settlement in Eastern Europe.* (Rutgers University Press, 2002).
48. Conard, N. J. & Bolus, M. Radiocarbon dating the appearance of modern humans and timing of cultural innovations in Europe: new results and new challenges. *J. Hum. Evol.* **44**, 331–371 (2003).
49. Hoffecker, J. F. The spread of modern humans in Europe. *Proc. Natl. Acad. Sci.* **106**, 16040–16045 (2009).
50. Kuhn, S. L. & Zwyns, N. Rethinking the initial Upper Paleolithic. *Quat. Int.* **347**, 29–38 (2014).
51. Reimer, P. J. *et al.* The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757 (2020).
52. Bennett, E. A. *et al.* Library construction for ancient genomics: single strand or double strand? *BioTechniques* **56**, 289–298 (2014).
53. Gansauge, M.-T. *et al.* Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic Acids Res.* **45**, e79 (2017).

54. Gansauge, M.-T. & Meyer, M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat. Protoc.* **8**, 737–748 (2013).
55. Massilani, D. *et al.* Past climate changes, population dynamics and the origin of Bison in Europe. *BMC Biol.* **14**, 93 (2016).
56. Renaud, G., Stenzel, U. & Kelso, J. leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res.* **42**, e141–e141 (2014).
57. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–60 (2009).
58. van der Valk, T. *et al.* Million-year-old DNA sheds light on the genomic history of mammoths. *Nature* **591**, 265–269 (2021).
59. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–9 (2009).
60. de Filippo, C., Meyer, M. & Prüfer, K. Quantifying and reducing spurious alignments for the analysis of ultra-short ancient DNA sequences. *BMC Biol.* **16**, (2018).
61. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
62. Okonechnikov, K., Conesa, A. & García-Alcalde, F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **btv566** (2015) doi:10.1093/bioinformatics/btv566.
63. Weissensteiner, H. *et al.* HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–W63 (2016).
64. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755 (2001).
65. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**, 772 (2012).
66. Martiniano, R., De Sanctis, B., Hallast, P. & Durbin, R. Placing Ancient DNA Sequences into Reference Phylogenies. *Mol. Biol. Evol.* **39**, msac017 (2022).
67. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
68. McKenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
69. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
70. Patterson, N. *et al.* Ancient Admixture in Human History. *Genetics* **192**, 1065–1093 (2012).

Acknowledgments

The paleogenomic facility of the Institut Jacques Monod obtained support from the University Paris Diderot within the program “Actions de recherches structurantes”. The sequencing facility of the Institut Jacques Monod, Paris, is supported by grants from the University Paris Diderot, the Fondation pour la Recherche Médicale (DGE20111123014), and the Région Ile-de-France (11015901). Moreover, we acknowledge support from the French national research center CNRS and EUR G.E.N.E. (ANR-17-EURE-0013 ; IdEx #ANR-18-IDEX-0001 l'Université de Paris ; Programme d'Investissements d'Avenir) for supporting the PhD fellowship extension of OP. We thank the National Academy of Sciences of Ukraine for permission to excavate at Buran-Kaya III and all the team members of the excavations. We also thank the French National Research Agency (ANR-05-JCJC-0240-01), the Fyssen Foundation, the Muséum national d'Histoire naturelle, the Centre National de la Recherche Scientifique (CNRS) for their financial supports (excavations and anthropological analyses). We thank Olivier Gorgé for assistance with some of the sequencing.

Author contributions

EMG, SaP and StP initiated the project. EMG and TG supervised the study. EAB, EMG and TG produced data. OP and EAB performed formal analyses. OP, EAB, TG and EMG interpreted the data. SaP, StP, LC and AY provided samples and information to the archaeological context. EAB wrote the paper with input from all authors.

Figures

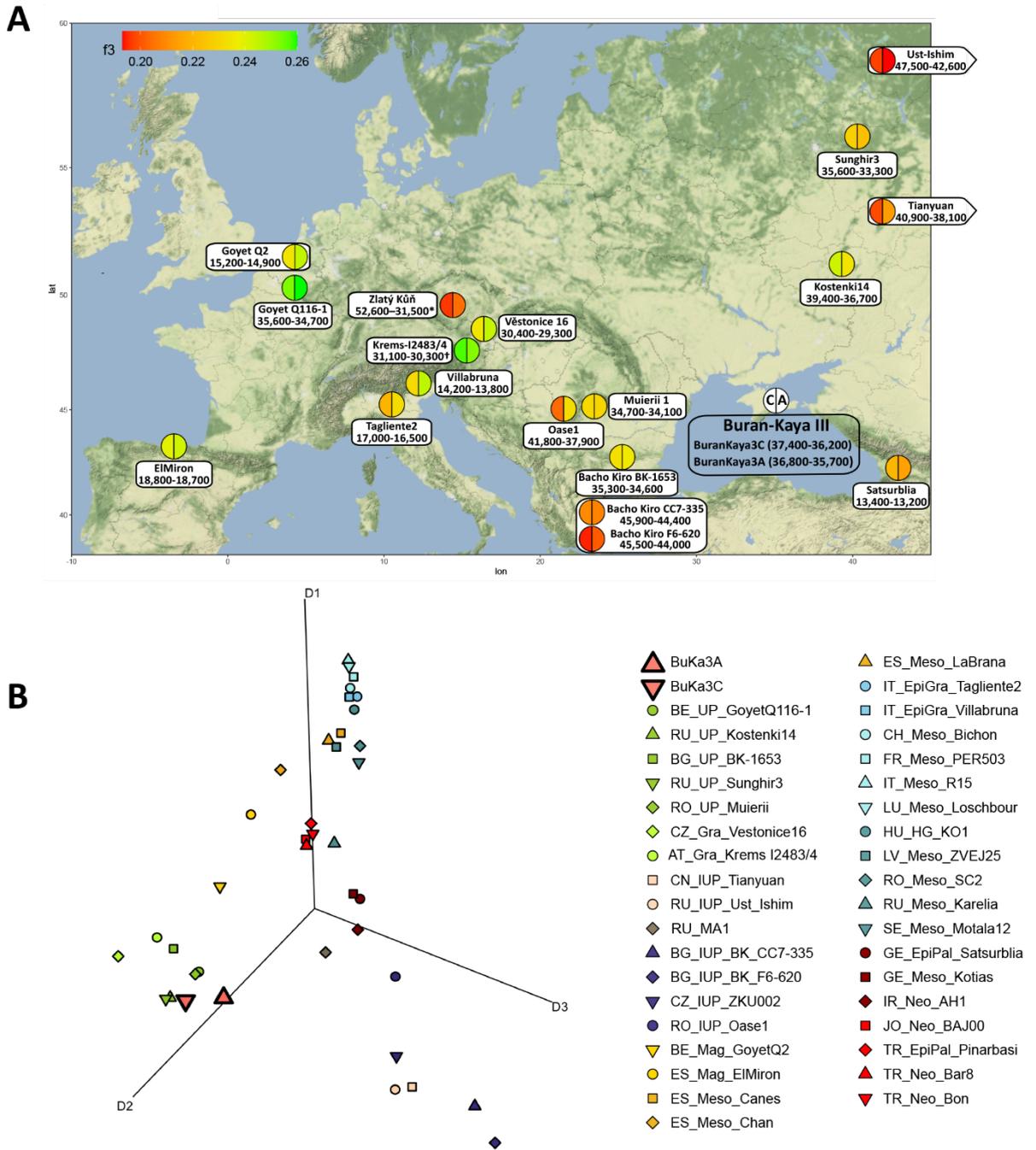


Figure 1. Genetic affinities between Buran-Kaya III and other ancient individuals. **A**, Heatmap showing the shared genetic drift calculated by f_3 -statistics in the form of $f_3(\text{Mbuti}; \text{BuKa3}, \text{ancient test})$. Higher f_3 values indicate greater allele sharing between Buran-Kaya III (BuKa3) and other individuals. The left half of the circles show f_3 values tested with BuKa3C and the right side with BuKa3A. Individuals that do not fit on the map due to geographical distance or overlapping locations of archaeological sites are placed as close as

possible to their correct locations. **B**, Multidimensional scaling (MDS) of ancient individuals calculated using the pairwise genetic distance matrix coming from inverted f_3 values ($1-f_3$) in the form of $f_3(\text{Mbuti}; \text{ancient1}, \text{ancient2})$. The three main dimensions are represented. Additional angles are visible on Fig. S6. A video of the rotating three dimensional MDS is available as supplementary information (SV1).



Figure 2. Population history of Buran-Kaya III individuals. **A**, Estimation of admixture graph using qpGraph that models the different waves of AMHs into Europe and their population history. The model has no outlier (highest $|Z\text{-score}| = 2.6$). Hypothetical populations nodes are written in italics. **B**, Genetic affinities within the post-CI cluster using f_4 -statistics in the form of $f_4(\text{Mbuti, Goyet Q116-1/Kostenki14/Krems 1283-4/Sunghir3/Vestonice16; BuKa3A/BuKa3C, test_post-CI})$. Black outlines indicate $|Z\text{-score}| \geq 2$.

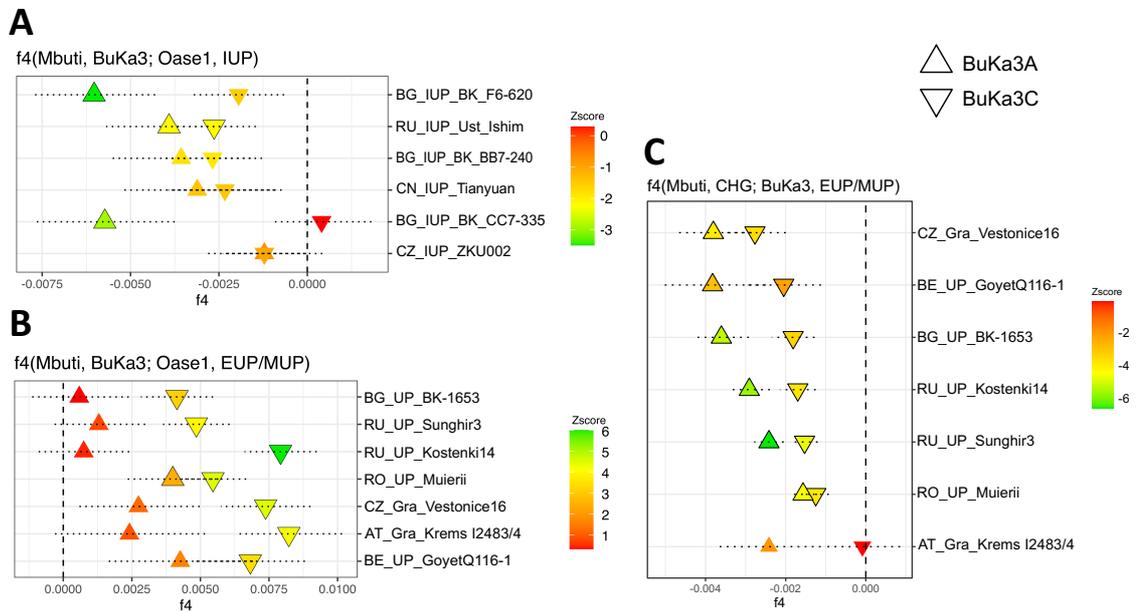


Figure 3. Detection of Buran-Kaya III gene-flow events. **A**, Gene flow from an Oase1-like population into BuKa3, tested using f_4 statistics in the form of $f_4(\text{Mbuti, BuKa3A/BuKa3C; Oase1, IUP})$. IUP represents individuals dating before $\sim 40,000$ cal BP. **B**, Genetic affinities and possible gene flow from Oase1-like population to BuKa3 tested using f_4 statistics in the form of $f_4(\text{Mbuti, BuKa3A/BuKa3C; Oase1, EUP/MUP})$. EUP/MUP represents individuals dating between $\sim 39 - 30,000$ cal BP. Neanderthal alleles from recent admixture were masked in Oase1. **C**, Genetic affinities between BuKa3 and the CHG cluster tested using f_4 statistics in the form of $f_4(\text{Mbuti, CHG; BuKa3A/BuKa3C, EUP/MUP})$. CHG represented by Kotias and Satsurblia individuals. Each test is highlighted with heatmap of Z-score values. Black outlines indicate $|Z\text{-score}| \geq 2$.

Supplementary Information pdf file:

Figure S1. Recalibration of comparative AMS ¹⁴C dates of Buran-Kaya III

Figure S2. Bayesian phylogenetic tree of mitochondrial sequences

Figure S3. Positions for BuKa3C, Oase1 and BK_F6-620 (Bacho Kiro) on Y chromosome tree

Figure S4. Neanderthal ancestry calculated for ancient individuals

Figure S5: Genetic affinities between Buran-Kaya III and present-day populations.

Figure S6. Additional angles of the 3D MDS scatterplot featured in Figure 1B

Figure S7. Multidimensional scaling (MDS)

Fig. S8. Observed ratio of mismatches to the reference sequence by read length (in bp) of Buran-Kaya III libraries

Figure S9. Terminal damage patterns for different BuKa3 library treatments

Figure S10. Remains from Buran-Kaya III

Figure S11. Sampling of human remains from Buran-Kaya III

Figure S12. Cranial fragment *BuKa3C*

Supplementary Video SV1: Rotating 3D MDS scatterplot featured in Figure 1B

Table S2. References for genomes used in this analysis

Table S7. References for mitochondrial sequences

Supplementary Text 1: Buran-Kaya III site, chronological, cultural and anthropological information

Supplementary Text 2: Strategy of collection and sampling protocol

Supplementary excel tables:

Table S1: Comparative ¹⁴C dates

Table S3: Summary of outgroup f3 statistics

Table S4: Summary of f4 statistics

Table S5: Summary of f4 statistics before and after masking the Neanderthal alleles

Table S6: Summary of library counts

Supplementary Information

Table of Contents

Figure S1. Recalibration of comparative AMS ¹⁴C dates of Buran-Kaya III

Figure S2. Bayesian phylogenetic tree of mitochondrial sequences

Figure S3. Y chromosome tree

Figure S4. Neanderthal ancestry calculated for ancient individuals

Figure S5: Genetic affinities between Buran Kaya III and present-day populations

Figure S6. Additional angles of the 3D MDS scatterplot featured in Figure 1

Figure S7. Multidimensional scaling (MDS)

Figure S8. Observed ratio of mismatches of Buran Kaya III libraries

Figure S9. Terminal damage patterns for different BuKa3 library treatments

Table S2. References for genomes used in this analysis

Table S7. References for mitochondrial sequences

Supplementary Text 1: Buran Kaya III site, chronological, cultural and anthropological information

Supplementary Text 2: Strategy of collection and sampling protocol

References

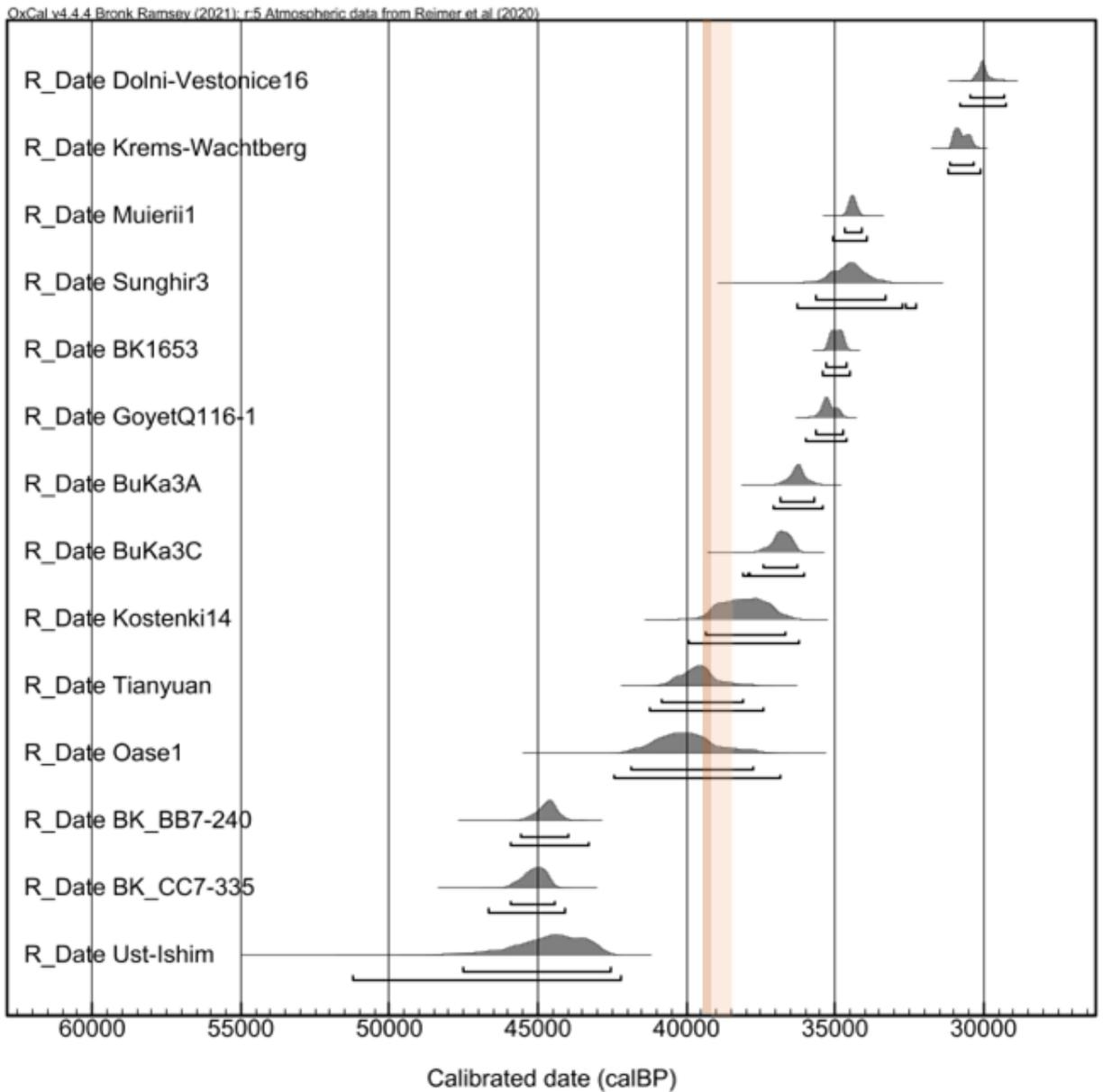


Figure S1. Recalibration of comparative AMS ^{14}C dates of Buran-Kaya III, layers 6-1 and 6-2 with Upper Paleolithic samples. All dates were recalibrated using the software OxCal v4.4.4 based on the IntCal20 calibration data set. Climatic events are represented by the orange band, the Campanian Ignimbrite (CI) volcanic super eruption event by the darker orange band, the Heinrich Event 4 by the lighter orange band.

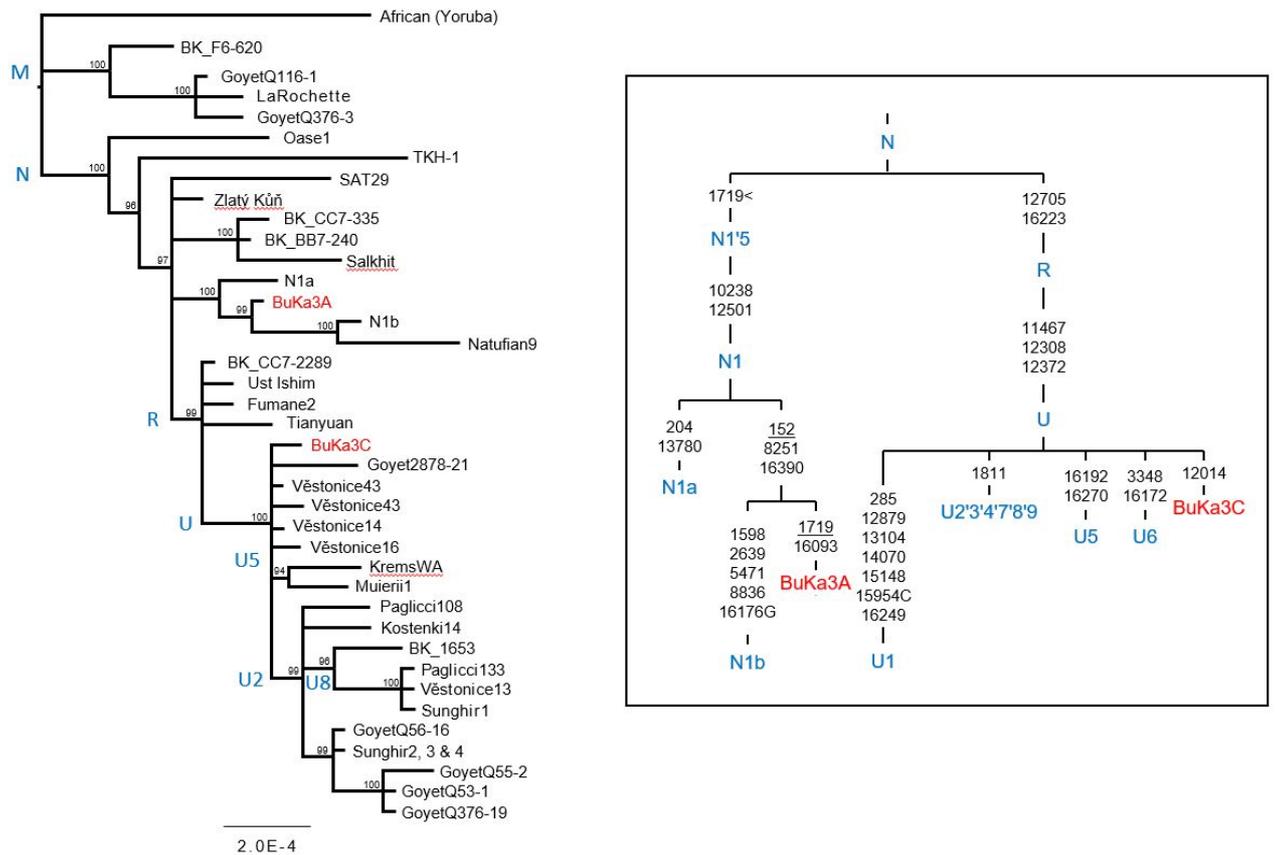


Figure S2. Bayesian phylogenetic tree of mitochondrial sequences, excluding Hyper-Variable Regions, from Early and Mid-Upper Palaeolithic individuals including those from Buran-Kaya III (red). Posterior probability indicated at the nodes. Scale bar denotes substitutions per site. Sources for mitochondrial sequences are listed in Table S3. Additional ancient and canonical N sequences are included to give branching details. Inset gives detailed mutation information from mtpHYL v.5.003 (<https://sites.google.com/site/mtpHYL/home>) for both BuKa3 individuals. References for mitochondrial sequences are given in Table S7.

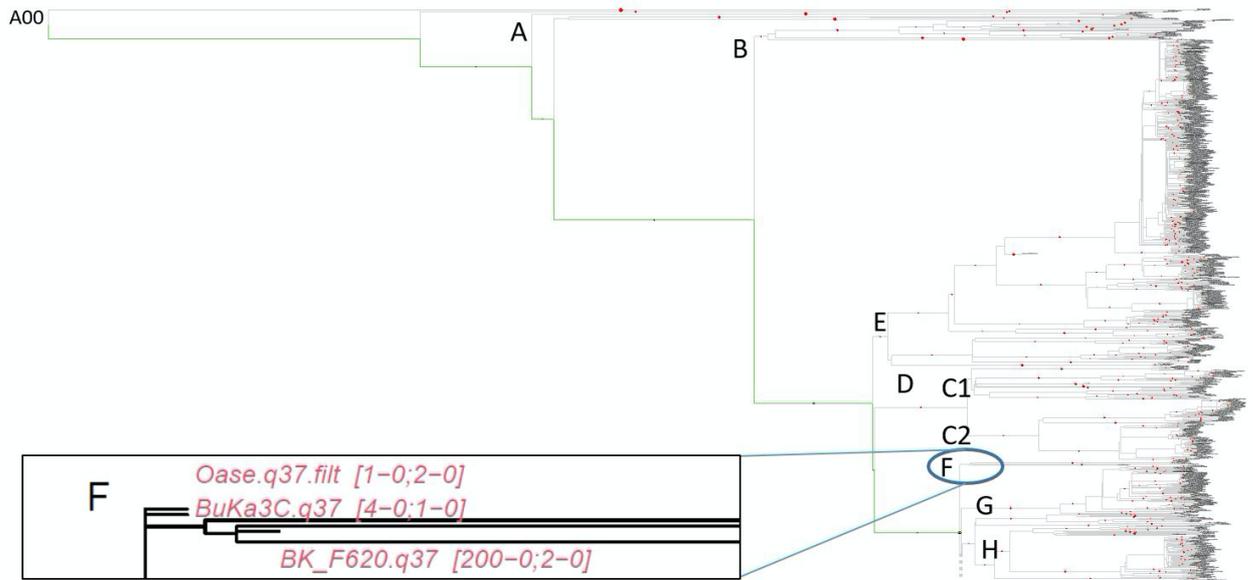


Figure S3. Positions for BuKa3C, Oase1 and BK_F6-620 (Bacho Kiro) on Y chromosome tree following Best Path analysis of pathPhynder¹. Green points on the tree show the numbers of supporting markers and red points, the numbers of non-supporting markers for the given branches for BuKa3C. Numbers following the sample names give supporting and non-supporting marker counts, respectively, for designations above the current branch followed by those for the current branch.

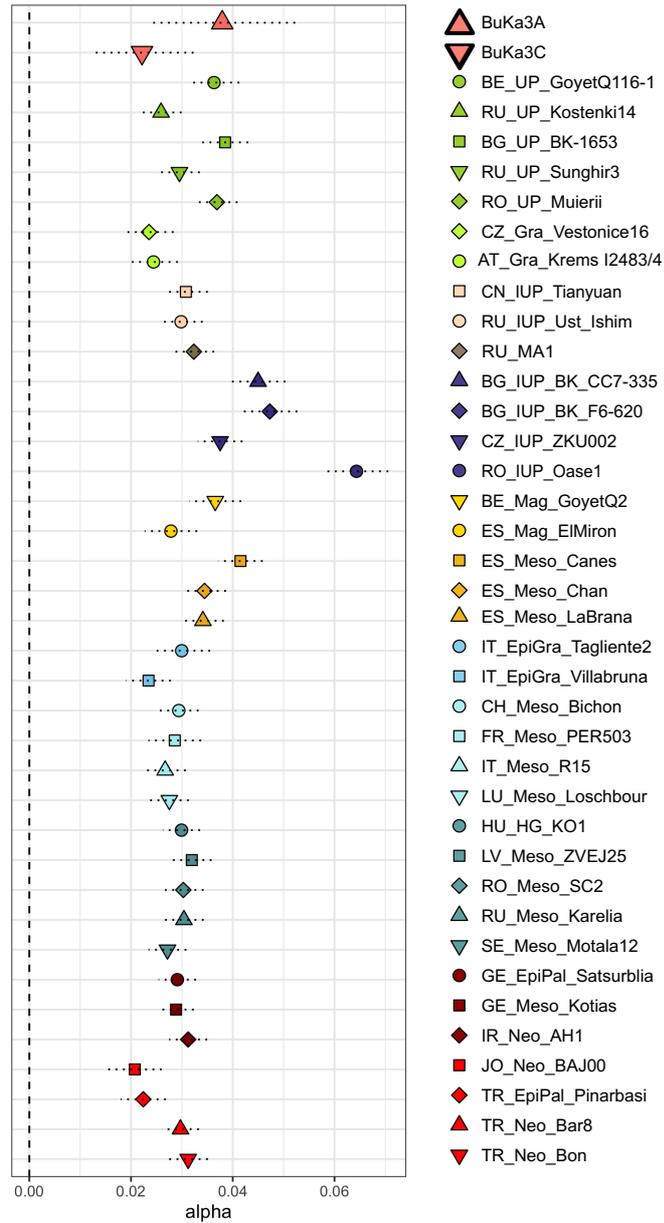


Figure S4. Neanderthal ancestry calculated for ancient individuals using direct f_4 -ratio test². α values calculated in the form of $(\text{Vindija33.19, Chimp; Ancient_test, Mbuti})/(\text{Vindija33.19, Chimp; Chagyrskaya, Mbuti})$.

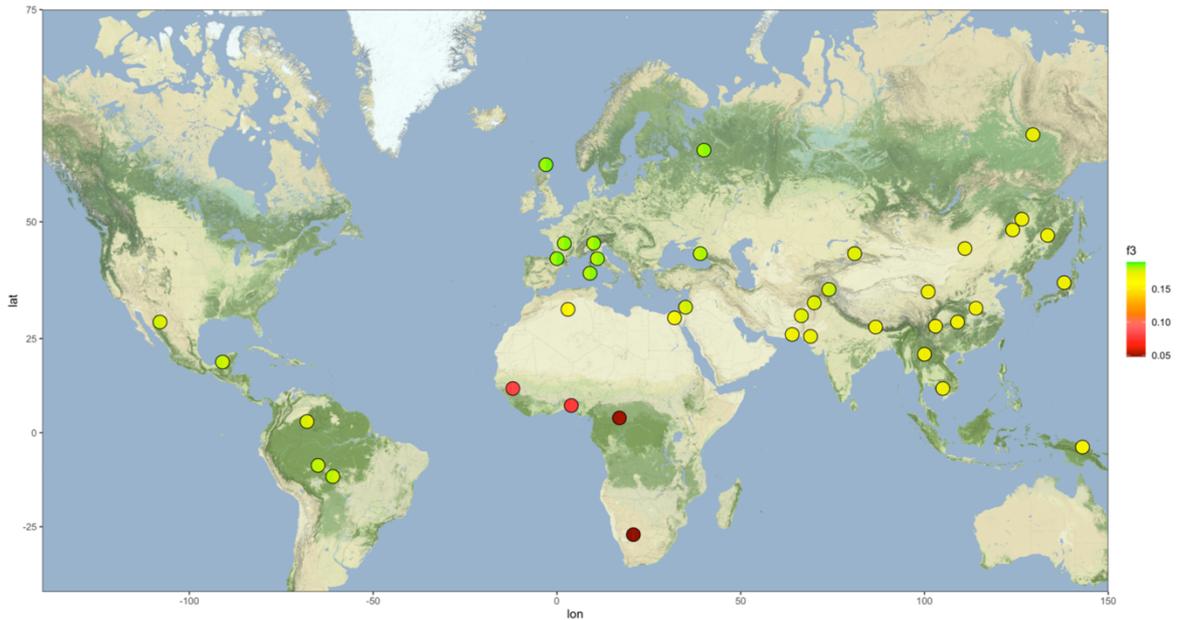


Figure S5: Genetic affinities between Buran Kaya III and present-day populations. Heatmap showing the shared genetic drift calculated by f_3 -statistics in the form of $f_3(\text{Mbuti}; \text{BuKa3}, \text{modern})$. Higher f_3 values indicate greater allele sharing between Buran Kaya III (BuKa3) and populations from HGDP dataset. The left half of the circles shows f_3 values tested with BuKa3C while right side show those tested with BuKa3A.

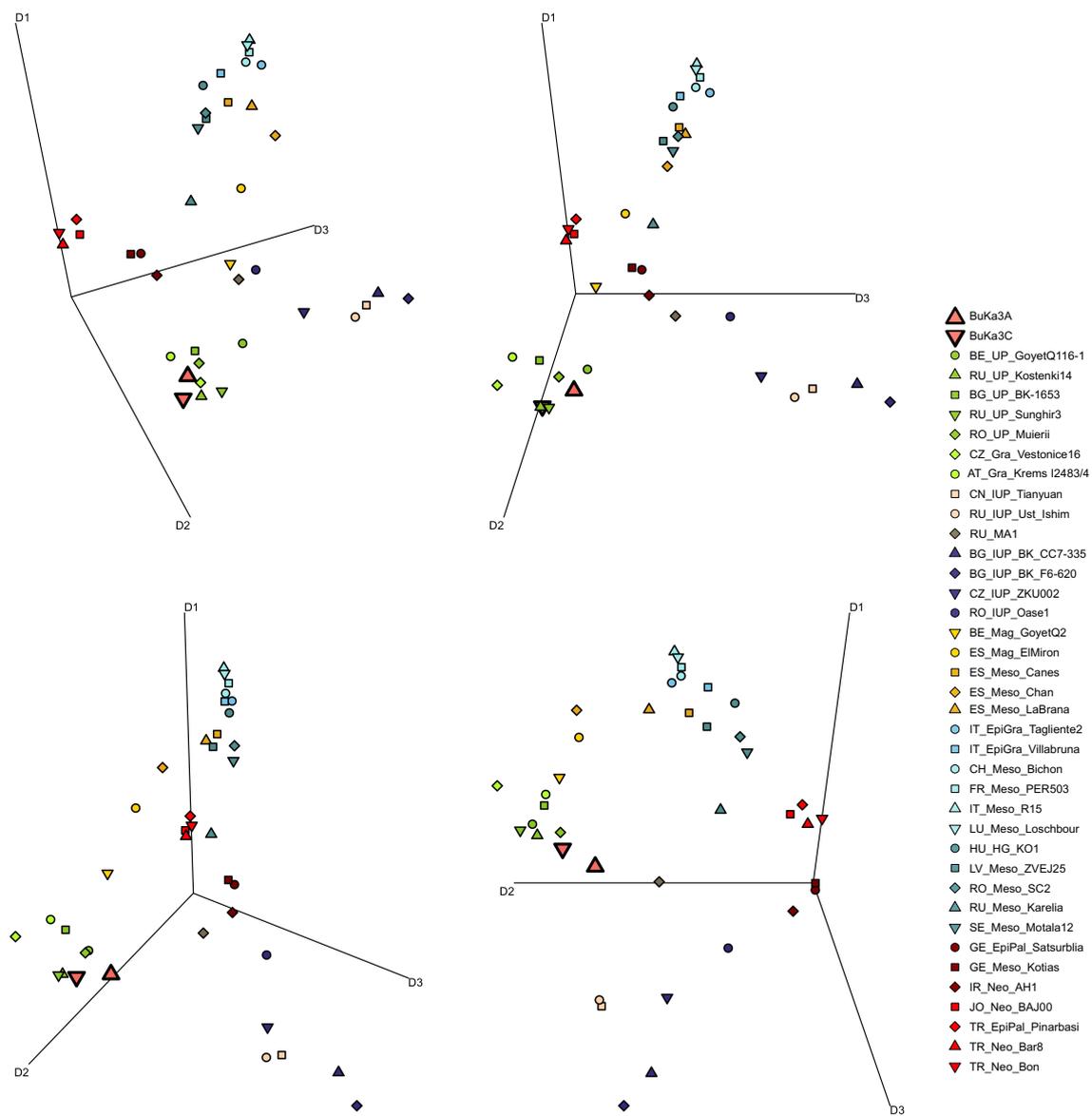


Figure S6. Additional angles of the 3D MDS scatterplot featured in Figure 1.

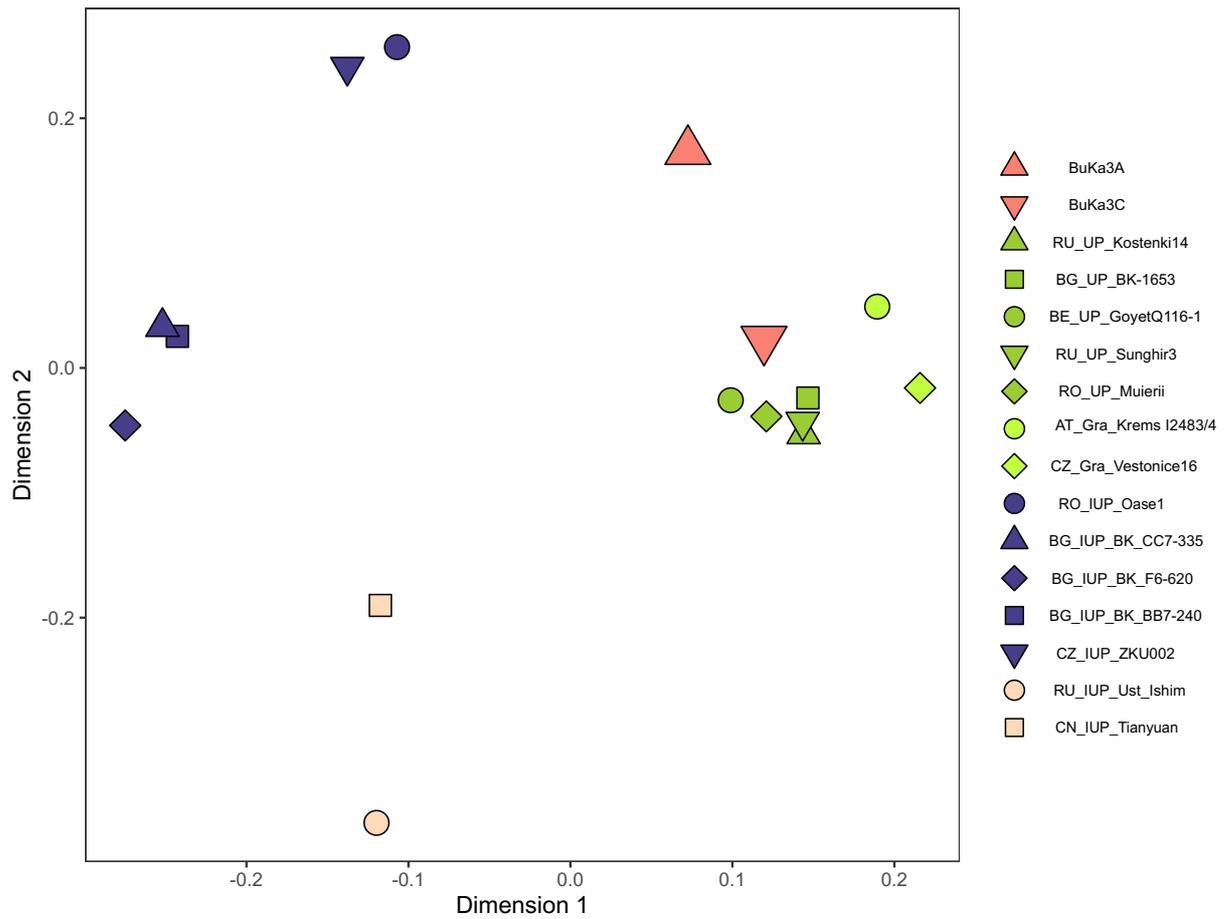


Figure S7. Multidimensional scaling (MDS) calculated using the pairwise genetic distance matrix coming from inverted $f\beta$ values ($1-f\beta$) in the form of $f\beta$ (Mbuti; ancient1, ancient2), for high-covered individuals older than 30,000 years only, showing the attraction between BuKa3A and a group consisting of Oase1 and Zlatý Kůň. Recent Neanderthal admixture regions have been masked for Oase1.

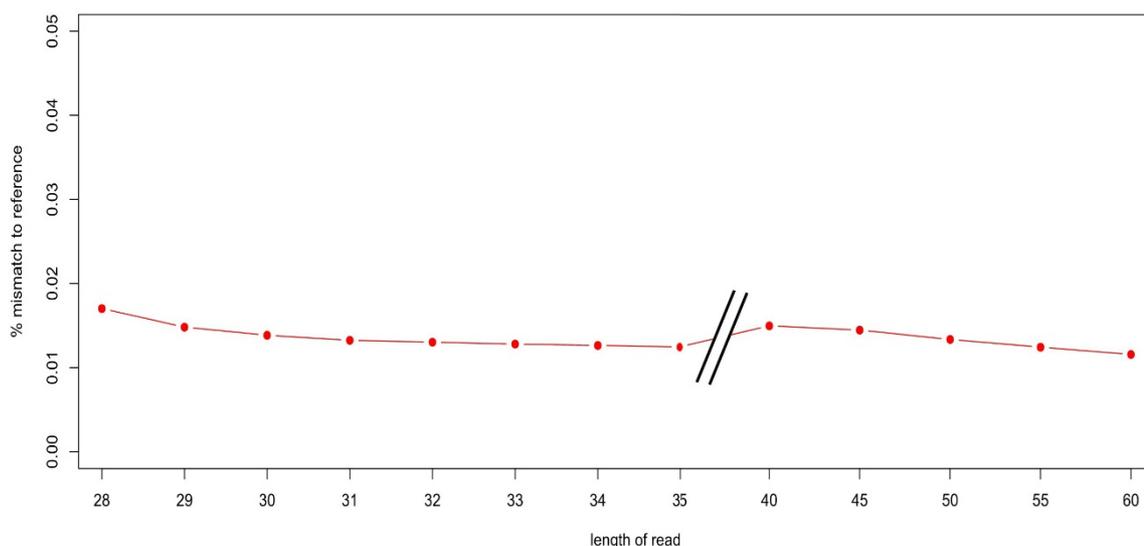


Fig. S8. Observed ratio of mismatches to the reference sequence by read length (in bp) of Buran Kaya III libraries. The increase mismatch percentage of shorter reads over the average of longer reads shows the rate of spurious mapping of shorter fragments. For Buran Kaya III libraries, this rate was 1.39% for reads 28-34 bp compared to an average of 1.32% for reads 35-60 bp.

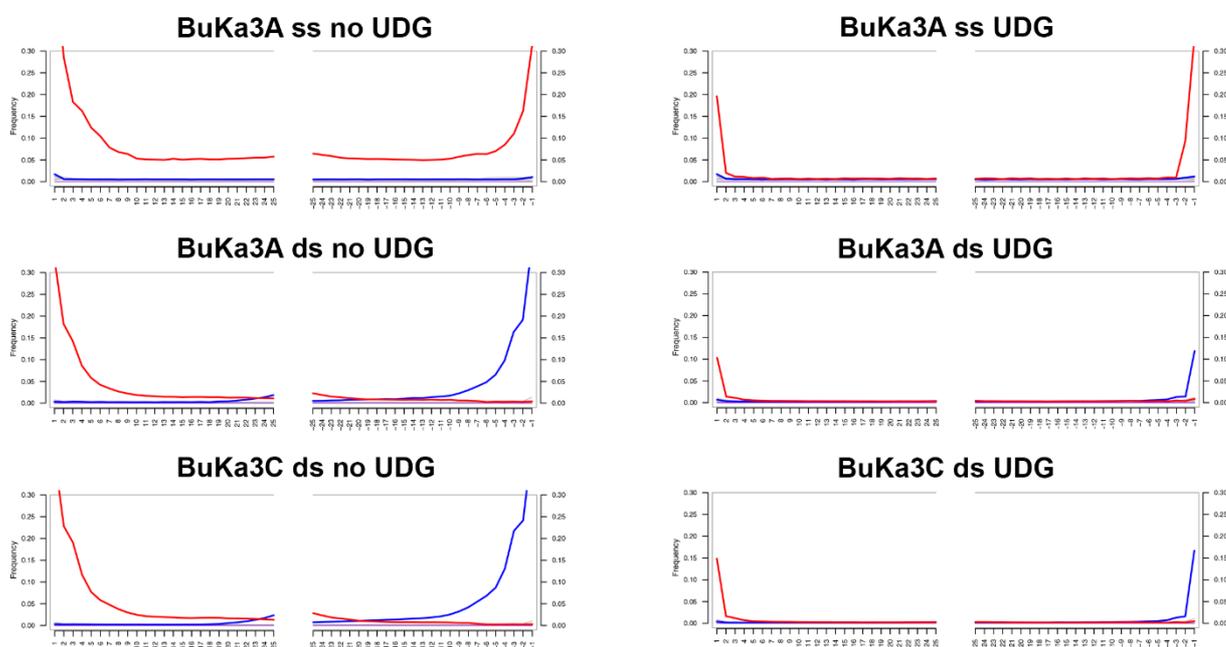


Figure S9. Terminal damage patterns for different BuKa3 library treatments, ss= single stranded, ds= double stranded libraries. UDG= Uracil-DNA Glycosylase.

Supplementary Text 1. Buran Kaya III site, chronological, cultural and anthropological information

By Sandrine Prat¹

¹UMR 7194 (HNHP), MNHN/CNRS/UPVD, Alliance Sorbonne Université, Musée de l'Homme, Paris, France

Chronological, cultural context and settlement patterns

Buran-Kaya III is a rock shelter located on the eastern part of Crimea, in the Belogorsk region, in the middle basin of the Burulcha River (4 km south from the city of Aromatne). This site was discovered in 1990 by A. Yanevich and excavated initially until 2001 by a team directed by A. Yanevich (Neolithic to Palaeolithic layers) and A. Marks (Middle and Early Upper Palaeolithic layers) with the participation of V. Chabai, Y. Demidenko, K. Monigal, M. Otte and Y. Yamada, followed by new fieldwork conducted by A. Yanevich and S. Péan between 2009 and 2011. This rockshelter unearthed an exceptional stratigraphic sequence from the Middle Palaeolithic to the Neolithic which lead to multiple studies^{9,10,50-71}.

Buran-Kaya III is a key site for understanding the Middle to Upper Palaeolithic transition and the arrival and dispersal of Anatomical Modern Humans (AMHs) in Europe as well their potential biological and/or cultural interactions with Neanderthals. During the excavation season of 2001 and 2009-2011, this site has yielded a Middle Palaeolithic layer (Micoquian, Kiik-Koba type, layer B), located unexpectedly above an Early Upper Paleolithic layer (Streletskian or eastern Szeletian, layer C) and below three Aurignacian layers (layers 6-5 to 6-3), followed by three Gravettian layers (layers 6-2, 6-1 and 5-2) and one Swiderian layer (layer 4), all densely layered across 1 to 2 meters of depth. The stratigraphical and radiochronological analyses¹⁰ had emphasized an early chronological framework for the Middle to Upper Paleolithic period in southeastern Europe. These studies show the presence of Early Upper Paleolithic (early Streletskian or eastern Szeletian, layer C) populations in Crimea, before 40.0 ka cal BP confirming a previous dating of 44.2-39.2 ka cal BP (36.7 ± 1500 ka BP, OxA-6868) of the Streletskaya or eastern Szeletian layer⁵² and providing early evidence of this transitional industry. Based on the new radiocarbon dates, the late Middle Palaeolithic settlements

(Micoquian layer, layer B) range between 35,390 +290/-270 (GrA-47319) and 37,700 ± 900 BP¹⁰ (OxA-25879) (43.2-39.9 cal BP), showing that no Neanderthal settlements occurred at Buran-Kaya III after the Campanian Ignimbrite eruption around 39 ka cal BP. The two Upper Paleolithic cultural traditions, related to the Aurignacian and Gravettian complexes (layers 6-5 to 5-2), were successively present in Crimea during a relatively short time span between 38.7 and 34.4 ka cal BP. During the following period of about 20,000 years, no depositional processes occurred or were preserved and the site was presumed abandoned. The Final Paleolithic populations (Swiderian, layer 4) settled in Crimea between 11.8 and 11.3 ka cal BP, i.e. during MIS 1.

The layers 6-2, 6-1 and 5-2 yielded a large collection of human remains, and, during the 2001 field season, more than 28,000 lithic remains, representing a large spectrum of tool types (e.g. burins, end-scrapers, and backed microliths). The lithic industries exhibited a high percentage of microliths (more than 80%) and some microgravettes⁶⁸, which are consistent with the Gravettian technocomplex. They do not present Dufour or pseudo-Dufour blades, often associated with Aurignacian technocomplex, which are present in layer 6-3, 6-4 and 6-5, stratigraphically located below⁶⁸. As shown by the sedimentological and stratigraphic analyses conducted by S. Puaud during the 2009 and 2010 field seasons, there is continuity in the stratigraphic sequence from the lower layers (Aurignacian technocomplex: 6-5, 6-4, 6-3) to the upper layers (Gravettian technocomplex: 6-2, 6-1 and 5-2). The attribution of the upper layers to the Gravettian tradition is free from material sedimentary overlap.

The ornaments excavated from the layers 5-2, 6-1 and 6-2 consist of perforated mollusc shells, fox and red deer teeth, perforated mammoth ivory pendants and a thin blade of mammoth tusk ivory with engraved "meander". Concerning the bone industry, more than 60 bone tools have been discovered in the Gravettian layers corresponding to projectile points, awls, arrowheads, and assegai points⁶⁸. The bone industry was probably imported from another site of production as indicated by the absence of manufacturing waste. This evidence, combined with the zooarchaeological results⁶⁶, suggests a pattern of recurrent short-term occupations, such as seasonal hunting or temporary butchery camps, associated with exploitation of small or medium-sized mammals, such as saiga antelope (the main game exploited in these layers). The settlement of AMHs associated with the Gravettian layers of Buran-Kaya III occurred under interstadial climatic conditions, which became cooler and drier leading to an open steppe environment. The open and steppe-like periglacial environment, with a local transition zone between steppes and mountains located on the migration path of several species, provided a

diversified faunal spectrum (e.g., saiga antelope, hare, red and polar foxes, reindeer, woolly rhinoceros, marmot, wild cat and brown bear)^{10,66,70}. The micromammal and palynological analyses^{58,72} reveal a steppe environment with a climatic evolution towards an increasing aridity as recorded in the mammalian sequence. In the layer 6-2, the climate was relatively cold, but not harsh, and relatively dry. The climatic conditions recorded in layer 6-1, which yielded the most diversified faunal spectrum of the sequence, became significantly drier than before, with less-contrasting seasons.

Human remains

At Buran-Kaya III, more than 160 human remains were discovered in three well-documented Gravettian layers (6-2, 6-1 and 5-2). Two of them unearthed from layer 6-1 and 6-2 have been directly dated to 31,900+240/-220 BP (36.7-35.7 ka cal BP), layer 6-1 (GrA-37938)⁹, and 32,450 +250/-230 BP (37.4-36.2 ka cal BP), layer 6-2 (GrA-50457)¹⁰ (Table S1). They are among the oldest direct evidence of AMHs in Europe in a well-documented archaeological context (Gravettian *sensu lato*)^{9,10}. The specimens from Buran-Kaya III represent, together with Bacho Kiro, Bulgaria (BK_CC7-335: 42,450 +/-510 BP, 45.9-44.4 ka cal BP (ETH-86772); BK_BB7-240: 41,850 +/- 480 BP, 45.5-43.9 ka cal BP(ETH-86770)⁴, Kostenki-Borshchevo, Russia (Kostenki 14: 33,250 +/-500 BP; 39.4-36.7 ka cal BP)⁸, and Peștera cu Oase, Romania (Oase 1: 34,290 +970/-870 BP, 41.8-36.9 ka cal BP, GrA-22810, 34,950 BP (combined age: 41.8-37.9 ka cal BP))⁶ (Table S1), the earliest occurrence of AMHs in Eastern Europe. AMHs occurred later in Western Europe, e.g., at Goyet Cave in Belgium (Goyet Q116-1: 30,880 +170-160 BP; 35.6-34.7 ka cal BP, GrA-46175)¹¹, in Russia, e.g., at Sunghir (Sunghir SI: 28,890 +/- 430 BP; 34.2-32.0 ka cal BP, OxA-X-2464-12; Sunghir SII: 30,100+/-550; 35.8-33.4 ka cal BP, OxA-A-2395-6; Sunghir SIII: 30,000 +/-550 BP; 35.6-33.3 ka cal BP, OxA-X-2395-7; Sunghir SIV: 29820 +/- 280 BP; 34.8-33.7 ka cal BP, OxA-X-2462-52)^{8,18} and in the Czech Republic e.g., at Dolní Věstonice (Dolní Věstonice 16, associated with Gravettian tradition: 25,740 +/-210 BP (on charcoal); 30.4-29.3 ka cal BP, GrN-15277)⁸ (Table S1).

Layer 6-1 has yielded the richest assemblage of human specimens with 150 highly fragmented remains (mostly cranial parts and teeth (95%) and several hand phalanges), corresponding to at least five individuals⁶⁷. Due to the fragmentation of the remains, the taxonomic assignment as AMH could only be performed on an occipital bone fragment and the permanent teeth. Based on the combinations of morphological features on these anatomical

elements (Figure S11) (namely, absence of an occipital “bun” or a bilaterally transverse torus; lack of a well-developed metaconid and a transverse crest on the lower premolars; lack of shovelling, labial convexity and the presence of well-developed lingual tubercles on the upper first incisors; the lack of a well-developed mid-trigonid crest and a large anterior fovea on the lower molars), these individuals from layer 6-1 are allocated to AMHs^{9,70}.

Isotopic analyses of bone collagen revealed that the diet of these individuals consisted mainly of terrestrial resources, mammoth meat playing a preponderant role in the protein intake, and of plant consumption⁶⁹.

Among the human remains from the layer 6-1 (2001, 2009 and 2010 field seasons), only a few bones exhibit human modifications, such as cut marks. The human skulls were intentionally selected in association with post-mortem treatments of the corpses (ritual cannibalism or a specific mortuary practice, such as post-mortem disarticulation processing of corpses for secondary disposal)^{9,66,70}. Based on these taphonomic observations, the specimens from layer 6-1 represent the oldest Upper Palaeolithic AMHs from Eastern Europe showing post-mortem treatment of the corpses.

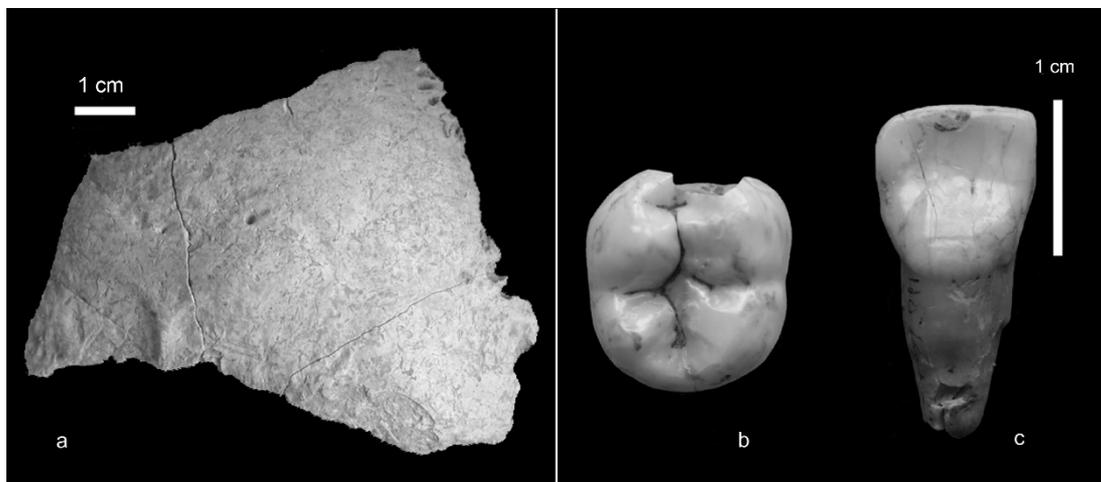


Figure S10 Remains from Buran-Kaya III (layer 6-1). From left to right, a) occipital bone; b) right lower second molar; c) right first upper incisor. Scale bar = 1 cm; photos © L.Crépin/S. Prat.

Supplementary Text 2: Strategy of collection and sampling protocol

By Laurent Crépin

UMR 7194 (HNHP), MNHN/CNRS/UPVD, Muséum national d'Histoire naturelle, Alliance Sorbonne Université, Institut de Paléontologie Humaine, Paris, France

The excavations performed between 2009 and 2011 incorporated a multi-sampling strategy to better contextualize the site of Buran-Kaya III. Sampling for sedimentological, biogeochemical, and radiometric studies were performed on all Gravettian layers, and in 2009, the sampling for the palaeogenomic analyses was carried out focused on the two layers that contained most of the human remains (6-1 and 6-2). To this end, a one half-square meter section was excavated incorporating precautions against contamination: a maximum of two excavators dug the square wearing masks and gloves, and when potential human remains were discovered, one excavator put on a new full body protection suit, including masks and gloves, to perform the removal of the material using new tools (Figure SI3a).

This protocol was maintained over eight days, during which time layers 6-1 and 6-2 were entirely cleared. In total, six bones were identified and inventoried using the described procedure, of which four were certified as human remains after verification: three from layer 6-1 (two cranial fragments and one tooth) and one for layer 6-2 (a cranial fragment). These remains, separated from other archaeological remains, were transferred directly after the mission to the high containment laboratory at the “Institut Jacques Monod” (IJM), where they were analyzed and stored at -20°C. Additionally, all those working on the site during the excavation campaign provided a cheek-swab sample to control for possible contamination of the samples.

The anthropological and taphonomic study of the human remains was then carried out directly in the high containment palaeogenomic laboratory at the IJM using typical precaution measures for ancient DNA research (Figure SI3b). Finally, we decided to carry out the paleogenomics analysis on the bone fragments that we could identify anatomically in an unambiguous manner and that appeared to be the best preserved with a minimum of visible alteration: artefact number N.170, square 9Z, layer 6-1.a, called here *BuKa3A* (Figure S12c) and

artefact number N.473, square 9Z, layer 6-2.a, called here *BuKa3C*. *BuKa3A* is a posterior cranial bone fragment presumably belonging to a mature individual, measuring a maximum of 21 x 19 mm.

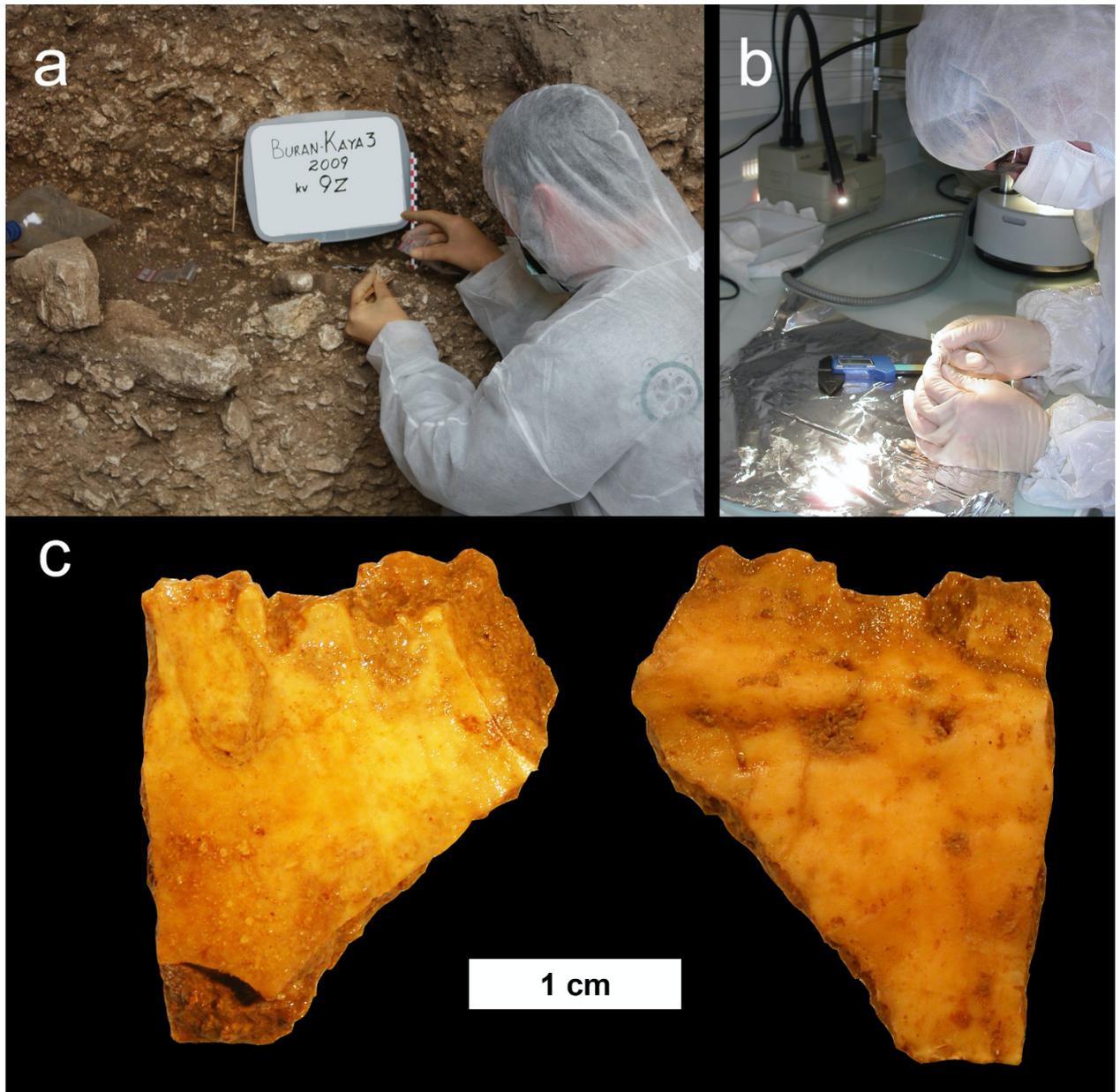


Figure S11. a) Sampling of human remains during the 2009 excavation season at Buran Kaya III; photo © A. Yanevich. b) Anthropological study of human remains in the high containment laboratory at the IJM; photo © E.M. Geigl. c) Images of the human cranial fragment from 6-1.a, *BuKa3A*, analyzed in the present study, taken under a stereomicroscope (Leica MZ FLIII, PLANAPO 0.63X); photo © E.M. Geigl.

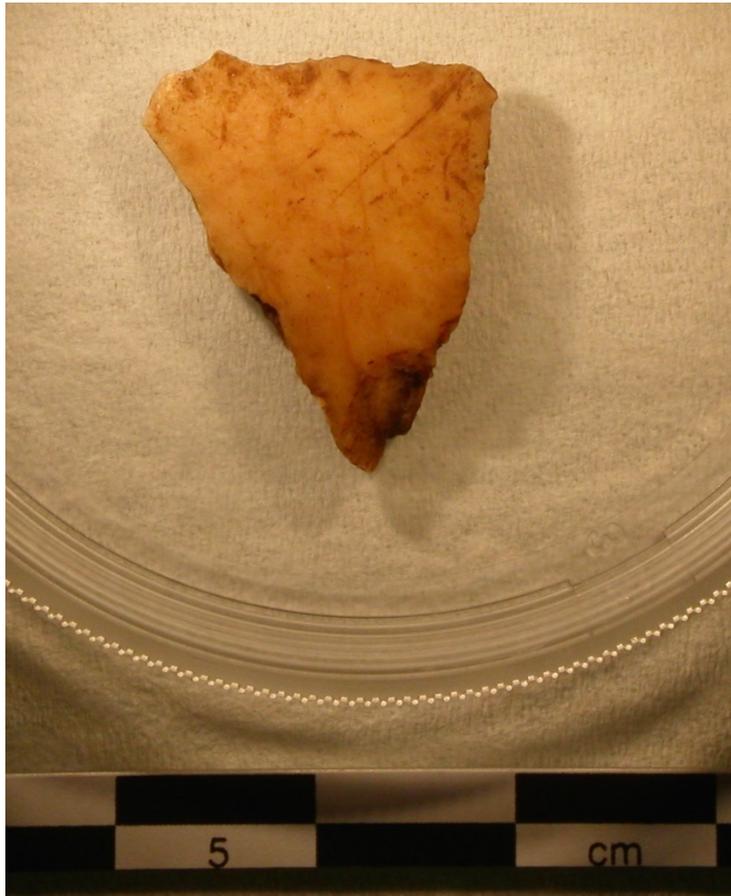


Figure S12. *BuKa3C* cranial fragment from square 9Z, layer 6-2.a, taken under a stereomicroscope (Leica MZ FLIII, PLANAPO 0.63X); photo © E.M. Geigl

Table S2. References for genomes used in this analysis

Name	Reference
AT_Gra_KremsWA12483-4	Teschler-Nikola, <i>et al.</i> ²²
BE_Mag_GoyetQ2	Fu, <i>et al.</i> ²⁶
BE_UP_GoyetQ116-1	Fu, <i>et al.</i> ²⁶
BG_IUP_BK_CC7-335	Hajdinjak, <i>et al.</i> ²⁷
BG_IUP_BK_F6-620	Hajdinjak, <i>et al.</i> ²⁷
BG_UP_BK-1653	Hajdinjak, <i>et al.</i> ²⁷
BuKa3A	This study
BuKa3C	This study
CH_Meso_Bichon	Jones, <i>et al.</i> ²⁵
CN_UP_Tianyuan	Fu, <i>et al.</i> ²⁸
CZ_Gra_Vestonice16	Fu, <i>et al.</i> ²⁶
CZ_IUP_ZKU002	Prüfer, <i>et al.</i> ³
ES_Mag_ElMiron	Fu, <i>et al.</i> ²⁶
ES_Meso_Canes	Gonzalez-Fortes, <i>et al.</i> ²⁹
ES_Meso_Chan	Gonzalez-Fortes, <i>et al.</i> ²⁹
ES_Meso_LaBrana	Olalde, <i>et al.</i> ³⁰
FR_Meso_PER503	Brunel, <i>et al.</i> ³¹
GE_EpiPal_Saturlia	Jones, <i>et al.</i> ²⁵
GE_Meso_Kotias	Jones, <i>et al.</i> ²⁵
HU_HG_KO1	Gamba, <i>et al.</i> ³²
IR_Neo_AH1	Broushaki, <i>et al.</i> ³³
IT_EpiGra_Tagliente2	Bortolini, <i>et al.</i> ³⁴
IT_EpiGra_Villabruna	Fu, <i>et al.</i> ²⁶
IT_Meso_R15	Antonio, <i>et al.</i> ³⁵
JO_Neo_BAJOD	Feldman, <i>et al.</i> ³⁶
LU_Meso_Loschbour	Jones, <i>et al.</i> ²⁵
LV_Meso_ZVEJ25	Jones, <i>et al.</i> ²⁵
RO_IUP_Oase1	Fu, <i>et al.</i> ³⁷
RO_Meso_SC2	Gonzalez-Fortes, <i>et al.</i> ²⁹
RO_UP_Muierii	Svensson, <i>et al.</i> ³⁸
RU_IUP_Ust-Ishim	Fu, <i>et al.</i> ⁵
RU_MA1	Raghavan, <i>et al.</i> ³⁹
RU_Meso_Karelia	Raghavan, <i>et al.</i> ³⁹
RU_UP_Kostenki14	Seguin-Orlando, <i>et al.</i> ⁷
RU_UP_Sunghir3	Sikora, <i>et al.</i> ⁴⁰
SE_Meso_Motala12	Lazaridis, <i>et al.</i> ⁴¹
TR_EpiPal_Pinarbasi	Feldman, <i>et al.</i> ³⁶
TR_Neo_Bar8	Hofmanová, <i>et al.</i> ⁴²
TR_Neo_Bon	Kılınc, <i>et al.</i> ⁴³

Table S3. Summary of outgroup f3 statistics in the form of f3(Mbuti; Ancient1, Ancient2)

Ancient1	Ancient2	Outgroup	f3	std_err	Z	SNPs
BuKa3A	AT_Gra_Krems I2483/4	Mbuti	0,253	0,008	30,2	12501
BuKa3A	BE_Aur_GoyetQ116-1	Mbuti	0,260	0,008	31,8	12911
BuKa3A	BE_Mag_GoyetQ2	Mbuti	0,245	0,008	31,6	11365
BuKa3A	BG_IUP_BB7-240	Mbuti	0,214	0,006	36,2	23048
BuKa3A	BG_IUP_BK_CC7-335	Mbuti	0,208	0,006	34,2	21385
BuKa3A	BG_IUP_BK_F6-620	Mbuti	0,201	0,005	38,4	35221
BuKa3A	BG_UP_BK-1653	Mbuti	0,236	0,006	41,8	31235
BuKa3A	CH_Meso_Bichon	Mbuti	0,221	0,004	56,0	82289
BuKa3A	CN_IUP_Tianyuan	Mbuti	0,213	0,007	31,7	20273
BuKa3A	CZ_Gra_Vestonice16	Mbuti	0,245	0,007	33,6	15961
BuKa3A	CZ_IUP_ZKU002	Mbuti	0,204	0,004	52,4	77014
BuKa3A	DE_Meso_BDB001	Mbuti	0,234	0,007	33,2	16988
BuKa3A	ES_Mag_ElMiron	Mbuti	0,243	0,008	30,0	12350
BuKa3A	ES_Meso_Canes	Mbuti	0,227	0,004	51,6	44891
BuKa3A	ES_Meso_Chan	Mbuti	0,218	0,004	51,9	56642
BuKa3A	ES_Meso_LaBrana	Mbuti	0,220	0,004	55,8	71521
BuKa3A	FR_Meso_PER503	Mbuti	0,214	0,006	35,9	22294
BuKa3A	GE_EpiPal_Satsurbliá	Mbuti	0,221	0,004	54,0	56849
BuKa3A	HU_HG_KO1	Mbuti	0,218	0,004	56,5	81551
BuKa3A	IE_Meso_SRA62	Mbuti	0,217	0,004	55,2	82926
BuKa3A	IR_Neo_AH1	Mbuti	0,220	0,004	52,9	56150
BuKa3A	IT_EpiGra_Tagliente2	Mbuti	0,232	0,007	34,3	18445
BuKa3A	IT_EpiGra_Villabruna	Mbuti	0,245	0,006	38,7	20749
BuKa3A	IT_Meso_R15	Mbuti	0,224	0,004	55,8	79667
BuKa3A	JO_Neo_BAJ00	Mbuti	0,225	0,011	21,0	7174
BuKa3A	LU_Meso_Loschbour	Mbuti	0,218	0,004	54,0	82881
BuKa3A	LV_Meso_ZVEJ25	Mbuti	0,226	0,004	56,0	78125
BuKa3A	LV_Meso_ZVEJ32	Mbuti	0,226	0,005	48,1	50356
BuKa3A	RO_IUP_Oase1	Mbuti	0,226	0,009	25,9	9363
BuKa3A	RO_Meso_OC1	Mbuti	0,218	0,004	50,3	63276
BuKa3A	RO_Meso_SC2	Mbuti	0,223	0,004	52,5	74452
BuKa3A	RO_UP_Muierii	Mbuti	0,230	0,004	56,3	82647
BuKa3A	RU_IUP_Ust_Ishim	Mbuti	0,193	0,004	48,4	83103
BuKa3A	RU_MA1	Mbuti	0,217	0,004	53,1	62792
BuKa3A	RU_Meso_Karelia	Mbuti	0,218	0,004	52,3	68187
BuKa3A	RU_UP_Kostenki14	Mbuti	0,234	0,005	48,1	43670
BuKa3A	RU_UP_Sungghir3	Mbuti	0,224	0,004	55,1	82545
BuKa3A	SE_Meso_Motala12	Mbuti	0,216	0,004	53,9	74865
BuKa3A	TR_EpiPal_Pinarbasi	Mbuti	0,218	0,007	31,8	17462
BuKa3A	TR_Neo_Bar8	Mbuti	0,220	0,004	58,1	81905
BuKa3A	TR_Neo_Bon	Mbuti	0,219	0,004	56,9	78374
BuKa3C	AT_Gra_Krems I2483/4	Mbuti	0,258	0,006	42,3	26825
BuKa3C	BE_Aur_GoyetQ116-1	Mbuti	0,253	0,006	42,2	27700
BuKa3C	BE_Mag_GoyetQ2	Mbuti	0,234	0,006	36,9	24555
BuKa3C	BG_IUP_BB7-240	Mbuti	0,203	0,004	45,2	49534
BuKa3C	BG_IUP_BK_CC7-335	Mbuti	0,210	0,005	45,3	45880
BuKa3C	BG_IUP_BK_F6-620	Mbuti	0,194	0,004	48,3	76001
BuKa3C	BG_UP_BK-1653	Mbuti	0,233	0,004	52,0	66764
BuKa3C	CH_Meso_Bichon	Mbuti	0,220	0,004	62,3	178363
BuKa3C	CN_IUP_Tianyuan	Mbuti	0,199	0,005	43,3	43345
BuKa3C	CZ_Gra_Vestonice16	Mbuti	0,235	0,005	45,9	34597
BuKa3C	CZ_IUP_ZKU002	Mbuti	0,196	0,004	56,0	166815
BuKa3C	DE_Meso_BDB001	Mbuti	0,228	0,005	42,7	36575
BuKa3C	ES_Mag_ElMiron	Mbuti	0,240	0,006	39,5	26512
BuKa3C	ES_Meso_Canes	Mbuti	0,226	0,004	61,2	95557
BuKa3C	ES_Meso_Chan	Mbuti	0,221	0,004	61,4	121867
BuKa3C	ES_Meso_LaBrana	Mbuti	0,221	0,004	62,6	154393
BuKa3C	FR_Meso_PER503	Mbuti	0,218	0,005	44,6	47806
BuKa3C	GE_EpiPal_Satsurbliá	Mbuti	0,215	0,004	60,8	122965
BuKa3C	HU_HG_KO1	Mbuti	0,220	0,003	63,6	176760
BuKa3C	IE_Meso_SRA62	Mbuti	0,218	0,003	62,5	179784
BuKa3C	IR_Neo_AH1	Mbuti	0,211	0,004	58,3	121256
BuKa3C	IT_EpiGra_Tagliente2	Mbuti	0,220	0,005	43,8	40054
BuKa3C	IT_EpiGra_Villabruna	Mbuti	0,231	0,005	47,1	44357
BuKa3C	IT_Meso_R15	Mbuti	0,225	0,003	66,5	172898
BuKa3C	JO_Neo_BAJ00	Mbuti	0,219	0,008	28,1	15476
BuKa3C	LU_Meso_Loschbour	Mbuti	0,221	0,004	62,7	179891
BuKa3C	LV_Meso_ZVEJ25	Mbuti	0,223	0,003	66,3	169204
BuKa3C	LV_Meso_ZVEJ32	Mbuti	0,221	0,004	58,2	109868
BuKa3C	RO_IUP_Oase1	Mbuti	0,203	0,006	33,9	20018
BuKa3C	RO_Meso_OC1	Mbuti	0,218	0,003	62,3	137418
BuKa3C	RO_Meso_SC2	Mbuti	0,221	0,003	64,0	161626
BuKa3C	RO_UP_Muierii	Mbuti	0,227	0,004	62,0	178949
BuKa3C	RU_IUP_Ust_Ishim	Mbuti	0,197	0,003	57,5	179888
BuKa3C	RU_MA1	Mbuti	0,214	0,004	60,4	135990
BuKa3C	RU_Meso_Karelia	Mbuti	0,217	0,004	59,6	148432
BuKa3C	RU_UP_Kostenki14	Mbuti	0,243	0,004	59,0	93744
BuKa3C	RU_UP_Sungghir3	Mbuti	0,228	0,004	63,2	178544
BuKa3C	SE_Meso_Motala12	Mbuti	0,217	0,004	61,0	162284
BuKa3C	TR_EpiPal_Pinarbasi	Mbuti	0,220	0,005	41,5	37960
BuKa3C	TR_Neo_Bar8	Mbuti	0,214	0,004	61,0	177423
BuKa3C	TR_Neo_Bon	Mbuti	0,210	0,003	63,1	169617

Table S5. Summary of f4 statistics before and after masking the Neanderthal alleles in the form of f4(UP, BuKa/Oase; BuKa/Oase, Mbuti)

W	X	Y	Z	Oase1				Oase1 Masked Neandertal			
				f4	stderr	Zscore	snps	f4	stderr	Zscore	snps
Mbuti	BuKa3A	RO_IUP_Oase1	BE_UP_GoyetQ116-1	0,004814	0,002415	1,993	9400	0,004263	0,002649	1,61	8229
Mbuti	BuKa3A	RO_IUP_Oase1	RU_UP_Kostenki14	0,002669	0,00156	1,71	18897	0,000746	0,001659	0,45	16524
Mbuti	BuKa3A	RO_IUP_Oase1	BG_UP_BK-1653	0,001878	0,001668	1,126	17922	0,000584	0,001802	0,324	15668
Mbuti	BuKa3A	RO_IUP_Oase1	RU_UP_Sunghir3	0,002482	0,001597	1,554	19981	0,001298	0,001689	0,769	17462
Mbuti	BuKa3A	RO_IUP_Oase1	RO_UP_Muierii	0,004038	0,001544	2,615	19958	0,003993	0,00166	2,405	17448
Mbuti	BuKa3A	RO_IUP_Oase1	CZ_Gra_Vestonice16	0,003439	0,002093	1,643	11880	0,002742	0,002232	1,228	10455
Mbuti	BuKa3A	RO_IUP_Oase1	AT_Gra_Krems I2483/4	0,00288	0,002533	1,137	9218	0,002414	0,002754	0,876	8079
Mbuti	BuKa3A	RO_IUP_Oase1	BG_IUP_BK_CC7-335	-0,004535	0,001863	-2,434	14039	-0,005726	0,001976	-2,898	12298
Mbuti	BuKa3A	RO_IUP_Oase1	BG_IUP_BK_F6-620	-0,005093	0,001624	-3,135	18648	-0,006029	0,001725	-3,495	16283
Mbuti	BuKa3A	RO_IUP_Oase1	BG_IUP_BB7-240	-0,002641	0,001806	-1,462	15131	-0,003567	0,001963	-1,817	13233
Mbuti	BuKa3A	RO_IUP_Oase1	CZ_IUP_ZKU002	-0,001196	0,001503	-0,131	19200	-0,001215	0,00164	-0,741	16779
Mbuti	BuKa3A	RO_IUP_Oase1	CN_IUP_Tianyuan	-0,00238	0,002034	-1,17	13294	-0,003114	0,002176	-1,431	11635
Mbuti	BuKa3A	RO_IUP_Oase1	RU_IUP_Ust_Ishim	-0,003204	0,001631	-1,964	19971	-0,003911	0,001779	-2,199	17454
Mbuti	BuKa3C	RO_IUP_Oase1	BE_UP_GoyetQ116-1	0,007831	0,001855	4,221	19515	0,00682	0,001988	3,43	17053
Mbuti	BuKa3C	RO_IUP_Oase1	RU_UP_Kostenki14	0,008302	0,001219	6,809	39470	0,007923	0,001316	6,02	34400
Mbuti	BuKa3C	RO_IUP_Oase1	BG_UP_BK-1653	0,00429	0,001252	3,428	37477	0,004145	0,001316	3,15	32655
Mbuti	BuKa3C	RO_IUP_Oase1	RU_UP_Sunghir3	0,005024	0,001193	4,21	41729	0,00485	0,001215	3,992	36323
Mbuti	BuKa3C	RO_IUP_Oase1	RO_UP_Muierii	0,005853	0,001123	5,212	41682	0,005456	0,001211	4,505	36285
Mbuti	BuKa3C	RO_IUP_Oase1	CZ_Gra_Vestonice16	0,007364	0,001528	4,82	24955	0,007378	0,001636	4,509	21846
Mbuti	BuKa3C	RO_IUP_Oase1	AT_Gra_Krems I2483/4	0,008865	0,001838	4,823	19101	0,008224	0,001939	4,24	16714
Mbuti	BuKa3C	RO_IUP_Oase1	BG_IUP_BK_CC7-335	0,001365	0,001304	1,047	29113	0,000404	0,001409	0,287	25431
Mbuti	BuKa3C	RO_IUP_Oase1	BG_IUP_BK_F6-620	-0,001156	0,001218	-0,949	38953	-0,001943	0,001288	-1,508	33911
Mbuti	BuKa3C	RO_IUP_Oase1	BG_IUP_BB7-240	-0,001291	0,0013	-0,993	31645	-0,002678	0,001396	-1,918	27560
Mbuti	BuKa3C	RO_IUP_Oase1	CZ_IUP_ZKU002	-0,000242	0,001093	-0,221	40007	-0,001214	0,00118	-1,029	34813
Mbuti	BuKa3C	RO_IUP_Oase1	CN_IUP_Tianyuan	-0,001612	0,001519	-1,061	27799	-0,002233	0,001604	-1,452	24353
Mbuti	BuKa3C	RO_IUP_Oase1	RU_IUP_Ust_Ishim	-0,001508	0,001121	-1,345	41713	-0,002637	0,001183	-2,228	36303
Mbuti	RO_IUP_Oase1	BuKa3C	BE_UP_GoyetQ116-1	0,000384	0,001778	0,216	19515	0,000382	0,001921	0,199	17053
Mbuti	RO_IUP_Oase1	BuKa3C	BE_Mag_GoyetQ2	0,001213	0,00178	0,681	15638	0,001243	0,001931	0,644	13645
Mbuti	RO_IUP_Oase1	BuKa3C	ES_Mag_ElMiron	-0,001214	0,001772	-0,685	19751	-0,000724	0,001906	-0,38	17308
Mbuti	RO_IUP_Oase1	BuKa3C	RU_IUP_Ust_Ishim	-0,003379	0,001193	-2,832	41713	-0,002953	0,001273	-2,32	36303
Mbuti	RO_IUP_Oase1	BuKa3C	AT_Gra_Krems I2483/4	0,003186	0,001652	1,929	19101	0,003096	0,001778	1,741	16714
Mbuti	RO_IUP_Oase1	BuKa3C	RU_UP_Kostenki14	-0,001519	0,001083	-1,402	39470	-0,001434	0,001168	-1,228	34400
Mbuti	RO_IUP_Oase1	BuKa3C	BG_UP_BK-1653	-0,000752	0,001136	-0,662	37477	-0,000559	0,001226	-0,456	32655
Mbuti	RO_IUP_Oase1	BuKa3C	CZ_Gra_Vestonice16	-0,001584	0,001429	-1,108	24955	-0,001462	0,001502	-0,973	21846
Mbuti	RO_IUP_Oase1	BuKa3C	CZ_IUP_ZKU002	-0,001429	0,001109	-1,288	40007	-0,001337	0,001217	-1,098	34813
Mbuti	RO_IUP_Oase1	BuKa3C	RU_MA1	-0,000407	0,001233	-0,33	33136	-0,000087	0,00134	-0,065	28867
Mbuti	RO_IUP_Oase1	BuKa3C	CN_IUP_Tianyuan	-0,002762	0,00145	-1,905	27799	-0,002412	0,001565	-1,541	24353
Mbuti	RO_IUP_Oase1	BuKa3C	RO_UP_Muierii	-0,000857	0,001095	-0,782	41682	-0,000848	0,001197	-0,708	36285
Mbuti	RO_IUP_Oase1	BuKa3C	RU_UP_Sunghir3	-0,001143	0,00108	-1,058	41729	-0,000265	0,001159	-0,228	36323
Mbuti	RO_IUP_Oase1	BuKa3C	IT_EpiGra_Tagliente2	-0,004044	0,002276	-1,935	9700	-0,004822	0,002438	-1,978	8495
Mbuti	RO_IUP_Oase1	BuKa3C	IT_EpiGra_Villabruna	0,001342	0,001304	1,029	30396	0,001853	0,001379	1,344	26575
Mbuti	RO_IUP_Oase1	BuKa3C	ES_Meso_Chan	-0,00129	0,001249	-1,033	33460	-0,000446	0,001336	-0,334	29263
Mbuti	RO_IUP_Oase1	BuKa3C	RU_Meso_Karelia	0,001821	0,001202	1,515	36298	0,001377	0,001321	1,042	31552
Mbuti	RO_IUP_Oase1	BuKa3C	CH_Meso_Bichon	0,000634	0,001165	0,544	41679	0,000573	0,001259	0,455	36276
Mbuti	RO_IUP_Oase1	BuKa3C	LU_Meso_Loschbour	0,000206	0,001065	0,194	41731	0,000937	0,001131	0,828	36321
Mbuti	RO_IUP_Oase1	BuKa3C	ES_Meso_Chan	-0,00129	0,001249	-1,033	33460	-0,000446	0,001336	-0,334	29263
Mbuti	RO_IUP_Oase1	BuKa3C	IR_Neo_AH1	0,000161	0,001229	0,131	31280	0,000143	0,001314	0,109	27211
Mbuti	RO_IUP_Oase1	BuKa3C	TR_Neo_Bon	-0,000261	0,00112	-0,233	40059	-0,000083	0,001195	-0,07	34888
Mbuti	RO_IUP_Oase1	BuKa3C	RO_Meso_OC1	0,000381	0,001313	0,29	31343	0,000723	0,001438	0,503	27349
Mbuti	RO_IUP_Oase1	BuKa3C	SE_Meso_Motala12	0,000205	0,001148	0,179	38061	0,000135	0,001252	0,108	33132
Mbuti	RO_IUP_Oase1	BuKa3C	GE_EpiGra_Satsurblia	0,000048	0,00127	0,038	29292	-0,00035	0,001367	-0,256	25489
Mbuti	RO_IUP_Oase1	BuKa3C	GE_Meso_Kotias	0,001438	0,001082	1,329	41480	0,001478	0,001173	1,26	36099
Mbuti	RO_IUP_Oase1	BuKa3C	BG_IUP_BB7-240	-0,000248	0,00137	-0,181	31645	-0,000555	0,00144	-0,386	27560
Mbuti	RO_IUP_Oase1	BuKa3C	BG_IUP_BK_CC7-335	0,002104	0,001438	1,463	29113	0,002692	0,001527	1,763	25431
Mbuti	RO_IUP_Oase1	BuKa3C	BG_IUP_BK_F6-620	-0,001037	0,001178	-0,88	38953	-0,000274	0,001265	-0,217	33911
Mbuti	RO_IUP_Oase1	BuKa3A	BE_UP_GoyetQ116-1	-0,000759	0,002352	-0,323	9400	-0,000851	0,002526	-0,337	8229
Mbuti	RO_IUP_Oase1	BuKa3A	BE_Mag_GoyetQ2	-0,000248	0,002648	-0,094	7526	-0,000551	0,002859	-0,193	6597
Mbuti	RO_IUP_Oase1	BuKa3A	ES_Mag_ElMiron	0,000731	0,002477	0,295	9557	0,001259	0,002717	0,464	8403
Mbuti	RO_IUP_Oase1	BuKa3A	RU_IUP_Ust_Ishim	-0,001511	0,001602	-0,943	19971	-0,001902	0,001782	-1,068	17454
Mbuti	RO_IUP_Oase1	BuKa3A	AT_Gra_Krems I2483/4	0,00043	0,002514	0,171	9218	-0,00001	0,002797	-0,004	8079
Mbuti	RO_IUP_Oase1	BuKa3A	RU_UP_Kostenki14	-0,002993	0,001584	-1,89	18897	-0,002984	0,001739	-1,716	16524
Mbuti	RO_IUP_Oase1	BuKa3A	BG_UP_BK-1653	-0,002749	0,001641	-1,676	17922	-0,003804	0,001835	-2,073	15668
Mbuti	RO_IUP_Oase1	BuKa3A	CZ_Gra_Vestonice16	0,000609	0,002105	0,289	11880	0,001693	0,002333	0,726	10455
Mbuti	RO_IUP_Oase1	BuKa3A	CZ_IUP_ZKU002	-0,001673	0,001584	-1,056	19200	-0,00184	0,001702	-1,081	16779
Mbuti	RO_IUP_Oase1	BuKa3A	RU_MA1	0,001801	0,001733	1,039	15944	0,002149	0,001861	1,155	13951
Mbuti	RO_IUP_Oase1	BuKa3A	CN_IUP_Tianyuan	-0,002376	0,002097	-1,133	13294	-0,00252	0,002279	-1,106	11635
Mbuti	RO_IUP_Oase1	BuKa3A	RO_UP_Muierii	-0,000461	0,001482	-0,311	19958	-0,000176	0,001629	-0,108	17448
Mbuti	RO_IUP_Oase1	BuKa3A	RU_UP_Sunghir3	0,00058	0,001533	0,378	19981	0,000626	0,001694	0,37	17462
Mbuti	RO_IUP_Oase1	BuKa3A	IT_EpiGra_Tagliente2	0,006318	0,003002	2,105	4706	0,004606	0,003375	1,365	4132
Mbuti	RO_IUP_Oase1	BuKa3A	IT_EpiGra_Villabruna	0,001695	0,001803	0,94	14561	0,001791	0,001922	0,932	12765
Mbuti	RO_IUP_Oase1	BuKa3A	ES_Meso_Chan	0,000496	0,001759	0,282	15987	0,000342	0,001912	0,179	14045
Mbuti	RO_IUP_Oase1	BuKa3A	RU_Meso_Karelia	0,00244	0,001634	1,494	17381	0,00152	0,001763	0,862	15189
Mbuti	RO_IUP_Oase1	BuKa3A	CH_Meso_Bichon	0,00042	0,00148	0,284	19934	0,000154	0,001625	0,095	17426
Mbuti	RO_IUP_Oase1	BuKa3A	LU_Meso_Loschbour	-0,000604	0,001561	-0,387	19982	-0,000751	0,001703	-0,441	17463
Mbuti	RO_IUP_Oase1	BuKa3A	ES_Meso_Chan	0,000496	0,001759	0,282	15987	0,000342	0,001912	0,179	14045
Mbuti	RO_IUP_Oase1	BuKa3A	IR_Neo_AH1	-0,00015	0,001629	-0,092	15015	-0,00037	0,001769	-0,209	13166
Mbuti	RO_IUP_Oase1	BuKa3A	TR_Neo_Bon	-0,000723	0,001564	-0,463	19121	-0,001358	0,001659	-0,804	16719
Mbuti	RO_IUP_Oase1	BuKa3A	RO_Meso_OC1	0,000626	0,001658	0,377	14945	-0,000029	0,001789	-0,016	13071
Mbuti	RO_IUP_Oase1	BuKa3A	SE_Meso_Motala12	-0,000197	0,001642	-0,12	18232	-0,00101	0,001768	-0,571	15915
Mbuti	RO_IUP_Oase1	BuKa3A	GE_EpiGra_Satsurblia	0,00073	0,001736	0,42	14116	0,000005	0,001901	0,003	12337
Mbuti	RO_IUP_Oase1	BuKa3A	GE_Meso_Kotias	-0,000845	0,001406	-0,601	19863	-0,00102	0,001506	-0,677	17367
Mbuti	RO_IUP_Oase1	BuKa3A	BG_IUP_BB7-240	-0,001157	0,00176	-0,657	15131	-0,000693	0,001923	-0,36	13233
Mbuti	RO_IUP_Oase1	BuKa3A	BG_IUP_BK_CC7-335	-0,0005	0,001937	-0,258	14039	-0,000264	0,002106	-0,125	12298
Mbuti	RO_IUP_Oase1	BuKa3A	BG_IUP_BK_F6-620	-0,002832	0,001639	-1,728	18648	-0,00324	0,001801	-1,798	16283

Table S6. Summary of library counts

Sample	Library type	Total merged Reads	Mapped Reads	%Mapped Reads	No Duplicates	% Duplicates	Q20	%Q20	no indel	Bp covered	no indel	Bp covered
BuKa3A	SS	434067516	1363797	0,31%	1308612	4,05%	968913	74,04%	949222	34646603	1,61E+006	5,88E+007
	DS	386553501	1077374	0,28%	958509	11,03%	668375	69,73%	660757	24117631		
BuKa3C	DS	307786326	4228673	1,37%	4196156	11,24%	3134038	74,69%	3124721	115614677	3,12E+006	1,16E+008

Table S7. References for mitochondrial sequences

	Haplogroup	Date*	Location	Reference
Yoruba	L3	modern	West Africa	AF347014.1
BK_BB7-240	N	44690	Bulgaria	Hublin, <i>et al.</i> ⁴
BK_CC7-335	N	45120	Bulgaria	Hublin, <i>et al.</i> ⁴
BK_CC7-2289	R	43690	Bulgaria	Hublin, <i>et al.</i> ⁴
BK_F6-620	M	43110	Bulgaria	Hublin, <i>et al.</i> ⁴
BK_1653	U8	34950	Bulgaria	Hublin, <i>et al.</i> ⁴
BuKa3A	N1	35770	Crimea	this study
BuKa3C	U	36830	Crimea	this study
Cioclovina1	U	32660	Romania	Fu, <i>et al.</i> ²⁶
Fumane2	R	39805	Italy	Benazzi, <i>et al.</i> ⁴⁴
Goyet2878-21	U5	26662	Belgium	Fu, <i>et al.</i> ²⁶
GoyetQ116-1	M	35210	Belgium	Fu, <i>et al.</i> ²⁶
GoyetQ376-19	U2	27515	Belgium	Fu, <i>et al.</i> ²⁶
GoyetQ376-3	M	33540	Belgium	Fu, <i>et al.</i> ²⁶
GoyetQ53-1	U2	27975	Belgium	Fu, <i>et al.</i> ²⁶
GoyetQ55-2	U2	27520	Belgium	Fu, <i>et al.</i> ²⁶
GoyetQ56-16	U2	26320	Belgium	Fu, <i>et al.</i> ²⁶
Kostenki14	U2	38050	Russia	Seguin-Orlando, <i>et al.</i> ⁷
KremsWA	U5	31020	Austria	Teschler-Nikola, <i>et al.</i> ²²
LaRochette	M	27592	France	Posth, <i>et al.</i> ¹¹
Muierii1	U6	35204	Romania	Hervella, <i>et al.</i> ⁴⁵
N1a	N1a	NA	NA	N1a canonical
N1b	N1b	NA	NA	N1b canonical
Natufian9	N1	10800	Israel	Lazaridis, <i>et al.</i> ⁴⁶
Oase1	N	39580	Romania	Fu, <i>et al.</i> ³⁷
Paglicci108	U2'3'4'7'8'9	28396	Italy	Fu, <i>et al.</i> ²⁶
Paglicci133	U8c	32260	Italy	Fu, <i>et al.</i> ²⁶
Salkhit	N	34425	Mongolia	Massilani, <i>et al.</i> ⁴⁷
SAT29	N	28543 †	Georgia	Gelabart, <i>et al.</i> ⁴⁸
Sunghir1	U8	33270	Russia	Sikora, <i>et al.</i> ⁴⁰
Sunghir2	U2	34610	Russia	Sikora, <i>et al.</i> ⁴⁰
Sunghir3	U2	34520	Russia	Sikora, <i>et al.</i> ⁴⁰
Sunghir4	U2	34320	Russia	Sikora, <i>et al.</i> ⁴⁰
Tianyuan	R	39570	China	Fu, <i>et al.</i> ²⁸
TKH-1	N	6978	Lybia	Vai, <i>et al.</i> ⁴⁹
Ust-Ishim	R	45020	Siberia	Fu, <i>et al.</i> ⁵
Věstonice13	U8	31140	Czechia	Fu, <i>et al.</i> ²⁶
Věstonice14	U5	30870	Czechia	Fu, <i>et al.</i> ²⁶
Věstonice16	U5	31250	Czechia	Fu, <i>et al.</i> ²⁶
Věstonice43	U5	31150	Czechia	Fu, <i>et al.</i> ²⁶
Zlatý Kůň	N	42036 †	Romania	Prüfer, <i>et al.</i> ³

*median value (cal BP) of dates from reference.

†median value from mtDNA tip dating.

References

1. Martiniano, R., De Sanctis, B., Hallast, P. & Durbin, R. Placing Ancient DNA Sequences into Reference Phylogenies. *Mol. Biol. Evol.* **39**, msac017 (2022).
2. Petr, M., Pääbo, S., Kelso, J. & Vernot, B. The limits of long-term selection against Neandertal introgression. (2018) doi:10.1101/362566.
3. Prüfer, K. *et al.* A genome sequence from a modern human skull over 45,000 years old from Zlatý kůň in Czechia. *Nat. Ecol. Evol.* **5**, 820–825 (2021).
4. Hublin, J.-J. *et al.* Initial Upper Palaeolithic Homo sapiens from Bacho Kiro Cave, Bulgaria. *Nature* **581**, 299–302 (2020).
5. Fu, Q. *et al.* Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445–449 (2014).
6. Trinkaus, E. *et al.* An early modern human from the Peștera cu Oase, Romania. *Proc. Natl. Acad. Sci.* **100**, 11231–11236 (2003).
7. Seguin-Orlando, A. *et al.* Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113–1118 (2014).
8. Marom, A., McCullagh, J. S. O., Higham, T. F. G., Sinitsyn, A. A. & Hedges, R. E. M. Single amino acid radiocarbon dating of Upper Paleolithic modern humans. *Proc. Natl. Acad. Sci.* **109**, 6878–6881 (2012).
9. Prat, S. *et al.* The Oldest Anatomically Modern Humans from Far Southeast Europe: Direct Dating, Culture and Behavior. *PLoS ONE* **6**, e20834 (2011).
10. Péan, S., Puaud, S., Crépin, L., Prat, S. & Quiles, A. The Middle to Upper Paleolithic Sequence of Buran-Kaya III (Crimea, Ukraine): New Stratigraphic, Paleoenvironmental, and Chronological Results. *Radiocarbon* **55**, 1454–1469 (2013).
11. Posth, C. *et al.* Pleistocene Mitochondrial Genomes Suggest a Single Major Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. *Curr. Biol.* **26**, 827–833 (2016).
12. Bader, O.N. *Sungir': Verkhnepaleoliticheskaya stoyanka [Sungir': An Upper Paleolithic site]*. (Nauka, 1978).
13. Vassilyev, S., Sinitsyn, A., & Marcel, O. *Le Sungirien. Actes du colloque de Saint-Petersbourg*. vol. ERAUL 147 (Etudes et Recherches Archéologiques de l'Université de Liège, 2017).
14. Anikovich, M. Early Upper Paleolithic industries of Eastern Europe. *J. World Prehistory* **6**, 205–245 (1992).
15. Bradley, B., Anikovich, M. & Engenii, G. Early Upper Paeolithic in the Russian Plain: Streletskayan flaked stone artefacts and technology. *Antiquity* **69**, 989–998 (1995).
16. Pettitt, P. B. & Bader, N. O. Direct AMS radiocarbon dates for the Sungir mid Upper Palaeolithic burials. *Antiquity* **74**, 269–270 (2000).
17. Dobrovolskaya, M., Richards, M.-P. & Trinkaus, E. Direct radiocarbon dates for the Mid Upper Paleolithic (eastern Gravettian) burials from Sunghir, Russia. *Bull. Mém. Société Anthropol. Paris* **24**, 96–102 (2012).

18. Nalawade-Chavan, S., McCullagh, J. & Hedges, R. New Hydroxyproline Radiocarbon Dates from Sungir, Russia, Confirm Early Mid Upper Palaeolithic Burials in Eurasia. *PLoS ONE* **9**, e76896 (2014).
19. Soficaru, A., Doboş, A. & Trinkaus, E. Early modern humans from the Peştera Muierii, Baia de Fier, Romania. *Proc. Natl. Acad. Sci.* **103**, 17196–17201 (2006).
20. Trinkaus, E. & Svoboda, J. *Early modern human evolution in Central Europe: the people of Dolní Věstonice and Pavlov*. (Oxford University Press., 2006).
21. Einwögerer, T. *et al.* Upper Palaeolithic infant burials. *Nature* **444**, 285–285 (2006).
22. Teschler-Nicola, M. *et al.* Ancient DNA reveals monozygotic newborn twins from the Upper Palaeolithic. *Commun. Biol.* **3**, 650 (2020).
23. Shang, H., Tong, H., Zhang, S., Chen, F. & Trinkaus, E. An early modern human from Tianyuan Cave, Zhoukoudian, China. *Proc. Natl. Acad. Sci.* **104**, 6573–6578 (2007).
24. Pinhasi, R. *et al.* Satsurblia: New Insights of Human Response and Survival across the Last Glacial Maximum in the Southern Caucasus. *PLoS ONE* **9**, e111271 (2014).
25. Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* **6**, 8912 (2015).
26. Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
27. Hajdinjak, M. *et al.* Initial Upper Palaeolithic humans in Europe had recent Neanderthal ancestry. *Nature* **592**, 253–257 (2021).
28. Fu, Q. *et al.* DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl. Acad. Sci.* **110**, 2223–2227 (2013).
29. González-Forbes, G. *et al.* Paleogenomic Evidence for Multi-generational Mixing between Neolithic Farmers and Mesolithic Hunter-Gatherers in the Lower Danube Basin. *Curr. Biol.* **27**, 1801-1810.e10 (2017).
30. Olalde, I. *et al.* Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225–228 (2014).
31. Brunel, S. *et al.* Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proc. Natl. Acad. Sci.* **117**, 12791–12798 (2020).
32. Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 5257 (2014).
33. Broushaki, F. *et al.* Early Neolithic genomes from the eastern Fertile Crescent. *Science* **353**, 499–503 (2016).
34. Bortolini, E. *et al.* Early Alpine occupation backdates westward human migration in Late Glacial Europe. *Curr. Biol.* **31**, 2484-2493.e7 (2021).
35. Antonio, M. L. *et al.* Ancient Rome: A genetic crossroads of Europe and the Mediterranean. *Science* **366**, 708–714 (2019).
36. Feldman, M. *et al.* Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nat. Commun.* **10**, 1218 (2019).
37. Fu, Q. *et al.* An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015).
38. Svensson, E. *et al.* Genome of Peştera Muierii skull shows high diversity and low mutational load in pre-glacial Europe. *Curr. Biol.* **31**, 2973-2983.e9 (2021).

39. Raghavan, M. *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87–91 (2014).
40. Sikora, M. *et al.* Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science* **358**, 659–662 (2017).
41. Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
42. Hofmanová, Z. *et al.* Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc. Natl. Acad. Sci.* **113**, 6886–6891 (2016).
43. Kılınc, G. M. *et al.* The Demographic Development of the First Farmers in Anatolia. *Curr. Biol.* **26**, 2659–2666 (2016).
44. Benazzi, S. *et al.* The makers of the Protoaurignacian and implications for Neandertal extinction. *Science* **348**, 793–796 (2015).
45. Hervella, M. *et al.* The mitogenome of a 35,000-year-old *Homo sapiens* from Europe supports a Palaeolithic back-migration to Africa. *Sci. Rep.* **6**, (2016).
46. Lazaridis, I. *et al.* Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).
47. Massilani, D. *et al.* Denisovan ancestry and population history of early East Asians. *Science* **370**, 579–583 (2020).
48. Gelabert, P. *et al.* Genome-scale sequencing and analysis of human, wolf, and bison DNA from 25,000-year-old sediment. *Curr. Biol.* **31**, 3564-3574.e9 (2021).
49. Vai, S. *et al.* Ancestral mitochondrial N lineage from the Neolithic ‘green’ Sahara. *Sci. Rep.* **9**, 3530 (2019).
50. Yanevich, A., Stepanchuk, V. & Cohen, Y. Buran-Kaya III and Skalistiy rockshelter: two new dated Late Pleistocene sites in the Crimea. *Préhistoire Eur.* **9**, 315–324 (1996).
51. Janevic, A. Buran-Kaya 3–Neue Angaben zur Kulturgliederung des Jungpaläolithikums der Krim. *Préhistoire Eur.* **13**, 133–148 (1998).
52. Marks, A. E. A new Middle to Upper Paleolithic “transitional” assemblage from Buran-Kaya III, level C, eastern Crimea: a preliminary report. in *Préhistoire d’Anatolie, Genèse de deux mondes, I* (ed. Otte, M.) vol. 1 353–366 (ERAUL, 1998).
53. Pettitt, P. Middle and early upper Palaeolithic Crimea: the radiocarbon chronology. in *Préhistoire d’Anatolie. Genèse des deux mondes/Anatolian Prehistory. At the Crossroads of Two Worlds* (ed. Otte, M.) vol. 1 329–338 (ERAUL, 1998).
54. Marks, A. & Monigal, K. The Middle to Upper Palaeolithic interface at Buran-Kaya III, Eastern Crimea. in *Neanderthals and modern humans-Discussing the transition: Central and Eastern Europe from 50,000-30,000 BP. from* (eds. Orschiedt, J., Weniger, G. C.) 212–226 (Mettmann: Neanderthal Museum, 2000).
55. Chabai, V. P. The Late Middle and Early Upper Paleolithic in Crimea (Ukraine). in *Actes du Colloque de la Commission VIII de l’UISPP* vol. 17 25–36 (Trabalhos de Arqueologia, 2001).
56. Chabai, V. The chronological and industrial variability of the Middle to Upper Paleolithic transition in Eastern Europe. in *Proceedings of Symposium 6.1 of the XIVth Congress of the UISPP* vol. 6 71–288 (2003).

57. Chabai, V., Monigal, K. & Marks, A. The Middle Paleolithic and early Upper Paleolithic of Eastern Crimea, Volume 3 (ERAUL 104). *Liège Univ. Liège* (2004).
58. Gerasimenko, N. Vegetational history of Buran-Kaya III. *Paleolit. Crimea Middle Paleolit. Early Up. Paleolit. East. Crimea* **3**, 19–34 (2004).
59. Markova, A. K. Small mammal fauna from Buran-Kaya III. *Paleolit. Crimea Middle Paleolit. Early Up. Paleolit. East. Crimea* **3**, 35–48 (2004).
60. Monigal, K. Introduction to the site of Buran-Kaya III. *Paleolit. Crimea Middle Paleolit. Early Up. Paleolit. East. Crimea* **3**, 3–18 (2004).
61. Patou-Mathis, M. Archeozoological analysis of large mammal fauna from Buran-Kaya III layer B. *Middle Paleolit. Early Up. Paleolit. East. Crimea Liege ERAUL* **104**, 95–111 (2004).
62. Demidenko, Y. E. The Early and Mid-Upper Palaeolithic of the North Black sea region: an overview. *Quartär* **55**, 99–114 (2008).
63. Patou-Mathis, M. The subsistence behaviours of the Last Crimean Neanderthals. in *Sourcebook of Paleolithic Transitions* 441–454 (Springer, 2009).
64. Yanevich, A. *et al.* Upper Palaeolithic settlements in Buran-Kaya 3 (Crimea, Ukraine): new interdisciplinary researches of the layers 5-2, 6-1 and 6-2. in *Archeological Almanac* vol. 20 187–202 (Donetsk: Donbass, 2009).
65. Chabai, V. The volcanic winter, Buran-Kaya cannibals and the fate of the last eastern European Neanderthals. *Arkheologiya* **3**, 5–26 (2012).
66. Crépin, L., Péan, S. & Lázničková-Galetová, M. Comportements de subsistance au Paléolithique supérieur en Crimée: analyse archéozoologique des couches 6-2, 6-1 et 5-2 de Buran-Kaya III. *L'Anthropologie* **118**, 584–598 (2014).
67. Prat, S. Stress physiologique et état de santé des plus anciens Hommes anatomiquement modernes du sud-est de l'Europe (données dentaires, couche 6-1, Buran-Kaya III, Crimée). *L'Anthropologie* **118**, 567–583 (2014).
68. Yanevich, A. Les occupations gravettiennes de Buran-Kaya III (Crimée): contexte archéologique. *L'Anthropologie* **118**, 554–566 (2014).
69. Drucker, D. G. *et al.* Isotopic analyses suggest mammoth and plant in the diet of the oldest anatomically modern humans from far southeast Europe. *Sci. Rep.* **7**, (2017).
70. Prat, S. *et al.* The First Anatomically Modern Humans from South-Eastern Europe. Contributions from the Buran-Kaya III Site (Crimea). *Bull. Mém. Société Anthropol. Paris* (2018) doi:10.3166/bmsap-2018-0032.
71. Yanevich, A. Buran-kays'ka kul'tura hravetu Krymu [The Buran-Kaya culture of Gravettian of Crimea]. *Arkheolohiya* **2**, 11–20 (2000).
72. Gerasimenko, N. Environmental changes in the Crimean mountains during the Last Interglacial–middle pleniglacial as recorded by pollen and lithopedology. *Quat. Int.* **164–165**, 207–220 (2007).

Chapter II

Genomics of the European Neolithization Process Paralleled by Fine-Scale Population Differentiation

Oğuzhan Parasayan¹, Corinne Thevenet², Jérôme Dubouloz², Michael Ilett², Denis Bouquin³, Nicolas Garmond², Ivan Praud², Lamys Hachem², Christophe Laurelut², Elisabeth Panloups², William Devriendt⁴, Corina Liesau von Lettow-Vorbeck⁵, Thierry Grange^{1*}, Eva-Maria Geigl^{1*}

¹Université Paris Cité, CNRS, Institut Jacques Monod, Paris, France

²UMR 8215 Trajectoires (CNRS-Université Paris I)

³Service Archéologique, Direction de l'Urbanisme, de la Planification, de l'Aménagement et de l'Archéologie (DUPAA), UMR 6298 ARTEHIS-Université de Bourgogne, France

⁴Direction de l'Archéologie, Béthune-Bruyais, Belgium

⁵Dpto. De Prehistoria y Arqueología, Facultad de Filosofía y Letras, Universidad Autónoma de Madrid, Spain

*co-senior

The Neolithic expansion from the Balkans to the Paris Basin along the continental route was a relatively slow process that took over 1,000 years. This left ample time for regional differentiation of both genomes and cultures. While some of the regional characteristics have been analyzed genetically in Central Europe, there is a lack of data for northern France. Therefore, we wanted to fill the empty space on the map and investigate the genomics of the neolithization process in this area at high resolution. We focused on the Paris Basin and sampled 168 individuals from the Early, Middle and Late Neolithic. Fine-scale analysis of newly generated genome data allowed us to shed light on fundamental populational processes, such as migration and admixture. The interpretation of the obtained genomic results in light of archaeological data, which is still ongoing, will modify our view of the neolithization process in western Europe.

Results & Discussion

We report the genomic analysis of 109 newly sequenced ancient individuals from 20 sites in the geographically defined Paris Basin including the present-day Picardy, Champagne and Nord-Pas-de-Calais regions of northern France. We also enriched the sample of Neolithic Iberian shotgun genomes with a new genome of a Late Neolithic individual. All the individuals studied here are associated with Neolithic cultures spanning a period from 5100 to 2500 BCE except one Mesolithic individual dating to ~6200 BCE (Fig. 1a, b). The majority of the sites are located along the Aisne Valley in Picardy and correspond to the remains of Neolithic villages dating between 5100 and 4900 BCE and associated with the final phase of the Linearbandkeramik (LBK) culture in Western Europe, formerly known as the “Rubané Récent du Bassin Parisien”- RRBP, and at present as terminal LBK¹. We also analyzed data obtained from three individuals from the site of Saint-Martin-sur-le-Pré in the Marne Valley Southeast of Reims. We also sequenced and analyzed individuals from post-LBK periods in the Aisne Valley including the Blicquy/Villeneuve-Saint-Germain (BVSG) culture dating between 4900-4700 BCE, as well as from later periods associated with the Cerny culture, the Michelsberg culture (Michelsberger Kultur - MK) and the post-Michelsberg period (post-MK) approximately dating to 4700-4300 BCE, 4250- 3700 BCE and 3600-3,00 BCE, respectively²⁻⁴.

We further included seven individuals, dating between 4126 and 3866 calBCE, from the Middle Neolithic II site of Escalles “Mont d’Hubert” from the most Western extension of the Paris Basin in the Flanders region (Cap Blanc Nez) with its particular “Escalles” culture. One analyzed individual was buried in the Middle Neolithic III “Balloy” culture site at Haut-de-Launois, Pont-sur-Seine in the Aube department Southeast of Paris and dated to 4715 ± 40 BP (GrA-47169) (3641-3380BCE). Another one originates from the site of Mairy “Les Hautes Chanvières”, a Middle Neolithic III site close to Sedan in the Ardennes located at the tipping point between the Rhine zone and the Western area of the “Michelsberg” culture and dating to 4770 ± 45 BP (3631-3372BCE). Finally, we analyzed 18 individuals from two hypogea at Chouilly « la Grifaine » in the Marne Valley dating to between 3400-3300 and 2900-2800 BCE, i.e., during the Late Neolithic (Table S1). The Iberian individual from Camino de las Yeseras in the Madrid region dates to the 3rd millennium BCE. We generated genomic data from these 110 ancient individuals by shotgun sequencing resulting in a coverage range from 0.1 to 4.2 X with a median of 0.6 X genomic coverage. 78 individuals with a coverage higher than 0.3 X were imputed and phased together with publicly available ancient genomes generated by shotgun sequencing (Table S1, S3) based on the high accuracy performance for common alleles (MAF >5%) reported by Rubinacci et al. Lower coverage genomes and those that had been genotyped using sequence capture with the 1240k SNP array were treated separately and considered as pseudo-haploid genomes (See Materials and Methods).

Uniparental Markers and Kinship

We determined the uniparental markers in all individuals. The mitochondrial haplogroups comprise the diversity of haplogroups that were previously reported in Neolithic individuals from France, i.e., H, J, K, N, T, U (Table S1)⁵⁻⁸. Similarly, Y haplogroup diversity of these newly reported Neolithic individuals comprise the common Neolithic and Mesolithic haplogroups C1a2, G2a2, H2, I2a1a, I2a1b, J2a1a, R1b. We encountered one exceptional case of I1 in a Late Neolithic individual from the Chouilly hypogea F11, which is a haplogroup commonly found in present-day Northern European populations⁹ and to our knowledge reported only once in an LBKT Neolithic site in Hungary¹⁰. The Y chromosome sequences from both the newly generated and the previously published genomes were placed into the reference Y chromosome phylogeny (Fig S1)¹¹.

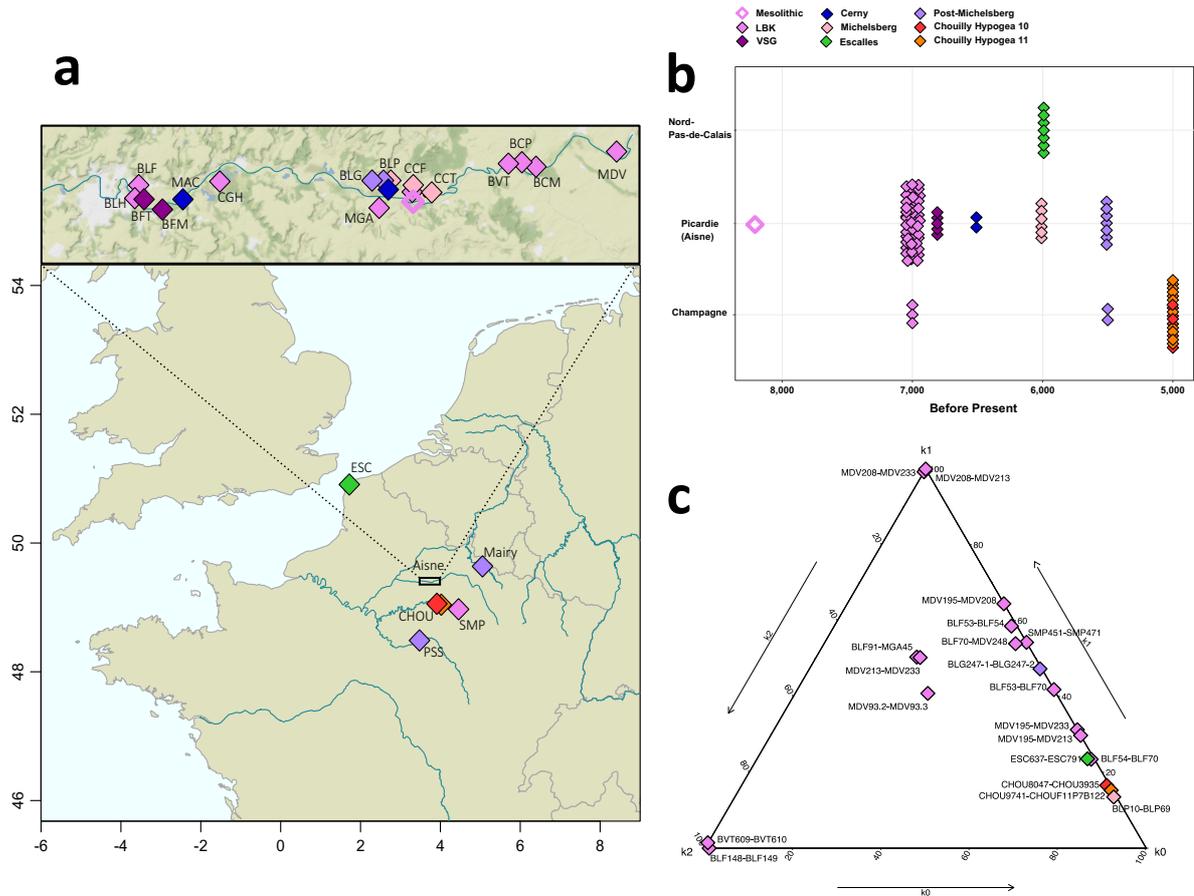


Figure 1. Spatiotemporal distribution and genetic relatedness of the analyzed Neolithic individuals

a, Map showing geographical locations of archaeological sites analyzed in this study. The Aisne valley is shown in the inset at higher resolution. **b**, Chronological timeline of the individuals grouped by archaeological cultures and geographical regions in northern France. **c**, Genetic relatedness among analyzed individuals. Cotterman coefficients (k_2 , k_1 , and k_0) of pairs of related individuals estimated by NgsRelate indicating probabilities of sharing 2, 1 or 0 autosomal alleles that are identical-by-descent.

In order to provide insights about the burial practices and social interactions within and between Neolithic sites and to exclude closely related individuals from downstream population analyses, we estimated the pairwise kinship coefficients for autosomes and X chromosomes of all individuals (Table S2). We calculated Cotterman coefficients (k_0 , k_1 , k_2), i.e., the genome fraction between two individuals sharing 0, 1 and 2 alleles in the identity-by-descent (IBD) approach, using NgsRelate¹² on imputed genomes. This analysis allowed us to infer the kin relationships between related individuals in comparison with previously reported pedigree simulations (Fig. 1c)¹³. We report two pairs of twins, BVT609 – BVT610 and BLF148- BLF149, in the sites of Berry-au-Bac "Le Vieux Tordoir" (BVT) and Bucy-le-Long "La Fosselle" (BLF),

respectively (Fig 1c.). To increase the genomic information from these individuals, the raw data of the twins were merged as if they were different genomic libraries of the same individuals. With the support of the uniparental haplogroups and the anthropological age of death estimation as well as the Cotterman coefficient distribution we infer a family including a mother, MDV208, and her two children, adult female MDV233 and juvenile male MDV213, that were buried in LBK burials in close vicinity in the site Menneville “Derrière le Village” (MDV)^{14,15} as the parent offspring pairs share a single allele in the IBD analysis at each position and the siblings are positioned in the middle of the ternary diagram (Fig. 1c). However, we do not detect any degree of kinship from other individuals that were buried in close vicinity to this family. There are two other pairs of sisters that can be deduced from their position in the ternary diagram and their mitochondrial haplogroups. The pair MDV93.2, an adult female, and MDV93.3, a female infant, was buried together in Menneville “Derrière le Village”. Strikingly, another pair of adult sisters were found to be buried in different sites roughly 30 km apart, one in Maizy “Les Grands Aisements” (MGA45) and one in Bucy-le-Long “La Héronnière” (BLF91) illustrating connections between LBK villages and suggesting some female mobility. We found also various inter- and intra-site second to third degrees of kinship that can be deduced from the distribution of k_0 and k_1 values. One notable case is a pair from the LBK site Saint-Martin-sur-le-Pré, a male SMP471 and a female SMP451. These individuals shared the mitochondrial haplogroup and were inferred based on the autosomal and X-chromosomal kinship values to have third- and second-degree kin relations (Table S2). These results indicate that they were half siblings sharing the same mother.

Mesolithic hunter-gatherers and Neolithic farmers in western Europe

Previous studies that compared the various aspects of individuals from Early Neolithic populations and their contributions to the populations of later periods in Western Europe relied only on genomes from LBK sites in Central Europe or made assumptions directly based on post-LBK populations in Western Europe^{7,8,16}. The conclusions drawn, however, might be misleading since they did not take into account the genome evolution of the westwards moving LBK groups leading to genetic differentiation. Instead, genetic differences between the northern and southern Neolithic genomes were explained exclusively by varying hunter-gatherer

proportions^{7,17,18}. It is questionable, however, that the moving LBK groups did only evolve through admixture. Genomic data from populations in Northwestern Europe, and in particular the Paris Basin, and a higher resolution analysis were needed to better inform the biological and social evolution during the LBK and the following Neolithic periods.

The newly produced genomes from Early Neolithic individuals from the territory of present-day France, in particular from the Paris Basin were deemed to fill this gap and to increase the resolution of the analysis. We excluded the genomes of related individuals as well as low coverage genomes and analyzed the remaining 269 imputed and phased Mesolithic and Neolithic genomes via ChromoPainterv2/fineSTRUCTURE analysis¹⁹ comprising 65 new genomes being produced in the present study. This analysis was followed by principal component analysis (PCA) performed on the linked coancestry matrix (Fig. 2a). Hunter-gatherers and farmers are separated by PC1 (right and left, respectively) whereas PC2 separates hunter-gatherer ancestries in a way roughly correlated with geography. We further focused our investigation on the differences in the genetic structure between the various hunter-gatherer genomes from both Mesolithic and Neolithic contexts applying the same pipeline (Fig. S2). The ensuing fineSTRUCTURE dendrogram reveals an east-to-west cline (Fig. S2a). In western Europe, the hunter-gatherer genomes are divided into two clusters corresponding to the presence of the previously well-defined dual ancestries associated with either the Magdalenian or the Epi-Gravettian cultures²⁰. The Epi-Gravettian ancestry, which was found in the early post LGM individuals from northern Italy, Tagliente and Villabruna dating back to respectively ~ 17 and ~ 14 kya^{20,21}, is represented in the present dataset by the genomes of slightly later individuals from the Bichon and Kendrick's Caves (Bichon and Kendrick, both $\sim 13,7$ kya)^{22,23}. The Magdalenian ancestry, which was previously found in Spain (El Miron, ~ 19 kya) and Belgium (Goyet Q2, ~ 15 kya)^{20,24}, is represented by the earliest individual from Gough's Cave (Gough, $\sim 14,8$ kya)²³. The fineSTRUCTURE dendrogram separates these two earliest individuals with Epi-Gravettian ancestry from the later western European Mesolithic individuals. The newly reported Mesolithic individual CCF711 from Cuiry-lès Chaudardes in the Aisne Valley (Fig. 1a,b) is both genetically and geographically closest to the cluster of Mesolithic hunter-gatherers from present-day France, Great Britain, and Luxembourg who all are close to the highly related Mesolithic hunter-gatherers from Italy, altogether associated with the earlier Bichon and Kendrick individuals (Fig. 2b). This cluster attests to the presence of genetic ancestry of the earlier Epi-Gravettian cultures in western Europe and thus genetic

continuity until the Mesolithic. However, this “Epi-Gravettian” group is distinct from the Mesolithic individuals from Iberia that have retained the genetic ancestry associated with Magdalenian hunter-gatherers, here represented by the Gough’s Cave individual^{23,24}.

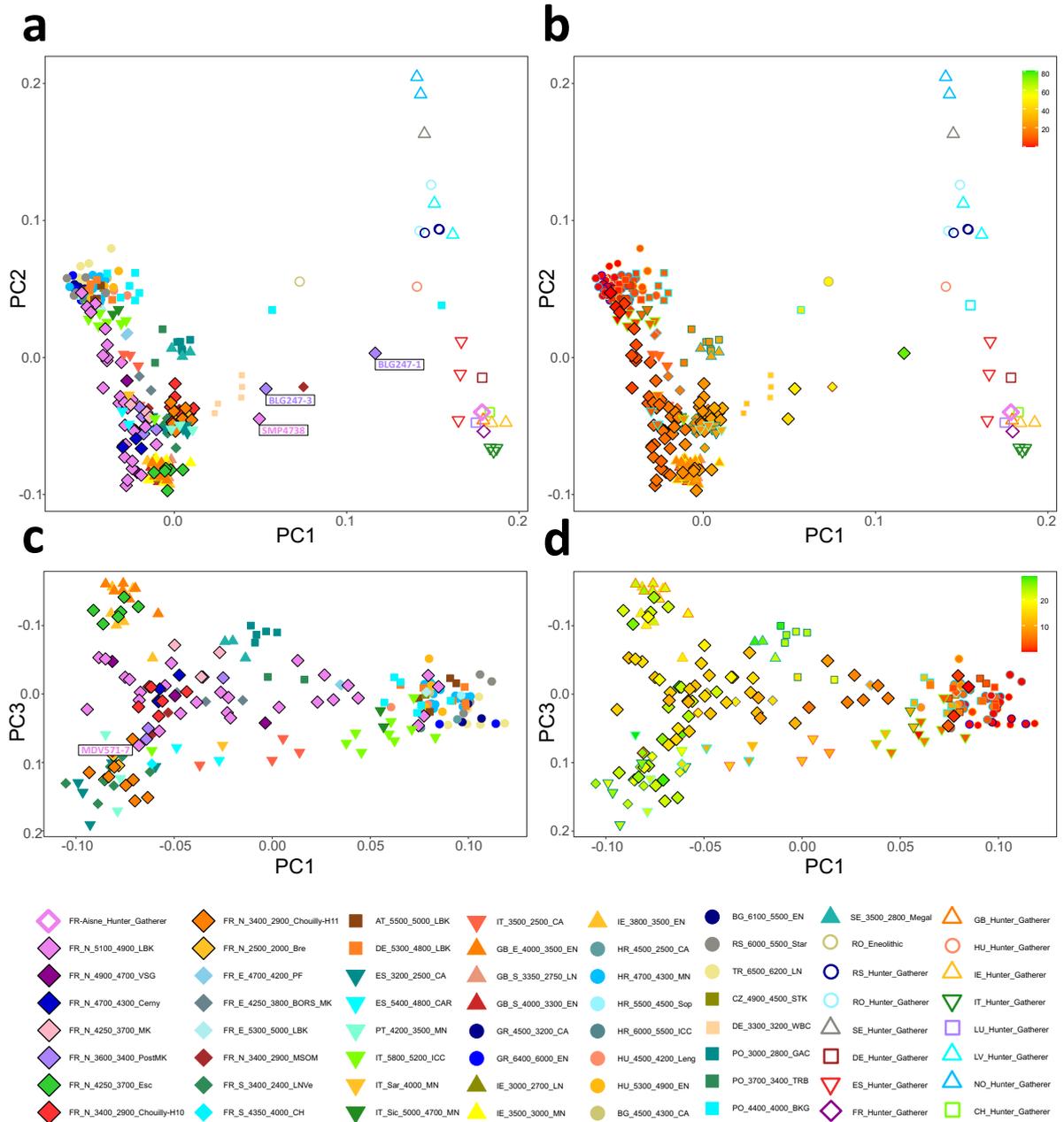


Figure 2. Fine-scale structure of Neolithic populations in Europe.

a, Principal Component Analysis (PCA) based on ChromoPainter haplotypic coancestry matrix calculated on the whole dataset including both farmer and hunter-gatherer genomes. The fraction of variance explained by PC1 and PC2 is, respectively, 64.9% and 9.1%. Empty symbols indicate hunter-gatherer genomes. Diamonds with black outlines indicate Neolithic individuals from northern France reported here. **b**, PCA plot as in Fig. 2a but showing individuals colored by their hunter-gatherer ancestry proportions estimated by

SOURCEFIND. The hunter-gatherer genomes are represented as in Fig. 2A. **c**, PCA based on ChromoPainter haplotypic coancestry matrix calculated on a subset of individuals excluding hunter-gatherer and outlier farmer genomes with high hunter-gatherer ancestry ($> \sim 35\%$). The later Neolithic individuals from British Isles are excluded from the dataset in order to reduce the effect of clustering formed by the extreme haplotype sharing (see Figure S4). The fraction of variance explained by PC1 and PC3 is, respectively, 38,9% and 2,8% (PC2: 5.5%, see Fig. S5). **d**, PCA plot as in Fig. 2c but showing individuals colored by their hunter-gatherer ancestry proportions estimated by SOURCEFIND. Names of the individuals referred to in the text are indicated. The black outlines indicate newly reported genomes. Two-letter country codes according to the International Organization for Standardization (ISO) are used in the group labels. (FR: France, AT: Austria, DE: Germany, ES: Spain, PT: Portugal, IT: Italy, GB: Great Britain (-E :England, -S: Scotland), IE: Ireland, GR: Greece, HR: Croatia, HU: Hungary, BG: Bulgaria, RS: Serbia, TR: Turkey, CZ: Czechia, PO: Poland, SE: Sweden, RO: Romania, NO: Norway, LV: Latvia, LU: Luxembourg, CH: Switzerland)

It has been shown that the Last Glacial Maximum (LGM) had a detrimental effect on the diversity of European hunter-gatherer populations, that the resulting bottleneck was followed by population expansions and that diversity and population size in farmer populations was higher in the Holocene²⁵⁻²⁸. To assess the recent demographic history of the ancestral population of the Mesolithic individual from the Paris Basin, CCF711, together with other Mesolithic groups, we calculated the distribution of runs of homozygosity (ROH) segments based on their total number of segments and their cumulative lengths. To render this estimation more robust and to decrease the number of false-positive homozygous segments, we restricted our analyses to segments longer than 1 centimorgan (cM). The ratio of cumulative length and number of ROH segments shows larger proportions of ROHs in Mesolithic individuals from western Europe compared to contemporary southeastern individuals (Fig. S3a). This suggests a higher “background relatedness” in the west due to possibly smaller recent population sizes. CCF711, however, has shorter cumulative ROHs compared to the other western Mesolithic individuals with Epi-Gravettian ancestry described above, to which he shows both genetic and geographical proximity. When the shorter ROH segments between 1 to 2 cM are excluded, the ROH proportion for this individual is drastically decreased indicating a lack of segments larger than 2 cM (Fig. S3b). This shift detected after exclusion of the ROH segments between 1 to 2 cM suggests that CCF711 originated from a population that was relatively larger compared to those of other western Mesolithic individuals. We predicted the pigmentation pattern of Mesolithic and Neolithic individuals in the imputed dataset through genotyping using the forensic HIrisplex-S assay²⁹. The Mesolithic hunter-gatherer from the Aisne Valley, CCF711, probably had phenotypic traits similar to Loschbour, the closest neighbor of CCF711 both

genetically and geographically, with blue eyes, dark black to brown hair and intermediate skin color (Table S7). This is in contrast to other Mesolithic hunter gatherers who probably had dark to black skin such the Mesolithic individuals from the British Isles^{16,30}.

The first farmers of western Europe – Genomic differences between the two Neolithic migration routes into Europe

The comparison of hunter-gatherers and European Neolithic farmers reveals the various levels of hunter-gatherer ancestries to be the main source of differentiation between European Neolithic farmers along PC1 as visible when the proportions of HG ancestry in the genomes is represented (Fig. 2b). PC2 also separates the Neolithic farmer groups from the Marmara (TR), Aegean (GR), the Balkans (BG, RS, HR), and Central Europe (DE, HU, PO) from those of the British Isles. Individuals from other regions are distributed between these two poles in a manner that is not related to their geographic origin (Fig 2a). Genomes from the Iberian Peninsula (ES, PT) and northern Europe (PO, SE) cluster together while those from the northern France LBK individuals (FR_N_5100_4900_LBK) are widely spread and appear to encompass the whole European Neolithic variability, as visible along PC2. Furthermore, some of these northern France LBK individuals are similar to Southern Mediterranean Neolithic individuals associated with the Impresa-Cardial (ICC) culture. This uncovers strong links between LBK groups in the Paris Basin and ICC groups from the Western Mediterranean.

Since this differentiation of Neolithic farmer does not follow their hunter-gatherer proportions, we aimed at better characterizing it by performing the PCA in the absence of hunter-gatherer-rich genomes ($>\sim 35\%$) (Fig. 2c, d, S4b). PC1 separates Neolithic farmers according to the combination of the extent of hunter-gatherer ancestry and an approximate east-to-west cline (Fig. 2c, d). PC3 separates the Neolithic farmers with hunter-gatherer ancestry according to a south-to-north cline (Iberian Peninsula and British Isles, respectively). The corresponding genetic differentiation seems to have started with the beginning of the Neolithic migrations since the earliest ICC groups in Italy have already diverged from the Balkan Neolithic population (See Fig. S4-5 for alternative PCA representations). The correlation

between the position of individuals on this PCA plot and their geographical locations reveals the progression of the genetic differentiation of Neolithic farmers as they move along the two European migration routes in Europe. In the PCA, the genomes on the right side correspond to the original Aegean Neolithic population before the separation into two main groups. Individuals of group I associated with the ICC moved from east-to-west (right-to-left in the PCA plot) over the Mediterranean via Italy (intermediate positions on the PCA plot) ending on the Iberian Peninsula (lower left corner). Genomes of the individuals belonging to the Neolithic group II moving northward through the Balkans and to the following LBK groups moving westward spread on the PCA plot to the upper left corner with central European individuals from Poland occupying an intermediate position. From these results we can conclude that the migrating groups on both routes diverged from the same source population in the southern Balkans as they moved into Europe and admixed with hunter-gatherers. The LBK population in the Paris Basin has a heterogenous distribution on the PCA testifying the early Neolithization of Europe as a transition period during which the farmer populations differentiated. Indeed, the presence of both eastern and western ancestries in the earliest farmer group of the Paris Basin highlights the ongoing formation of a local ancestry signature in this founder population. The spread along PCA of the later Neolithic individuals from the Paris Basin is more restricted as it is the case for the other western European individuals. Along PC3, they occupy an intermediate position between the Iberian Peninsula/southern France and the British Isles. In conclusion, our results show the ChromoPainterv2/fineSTRUCTURE-based PCA to be powerful enough to uncover the genomic differences between the southern ICC and the northern LBK groups.

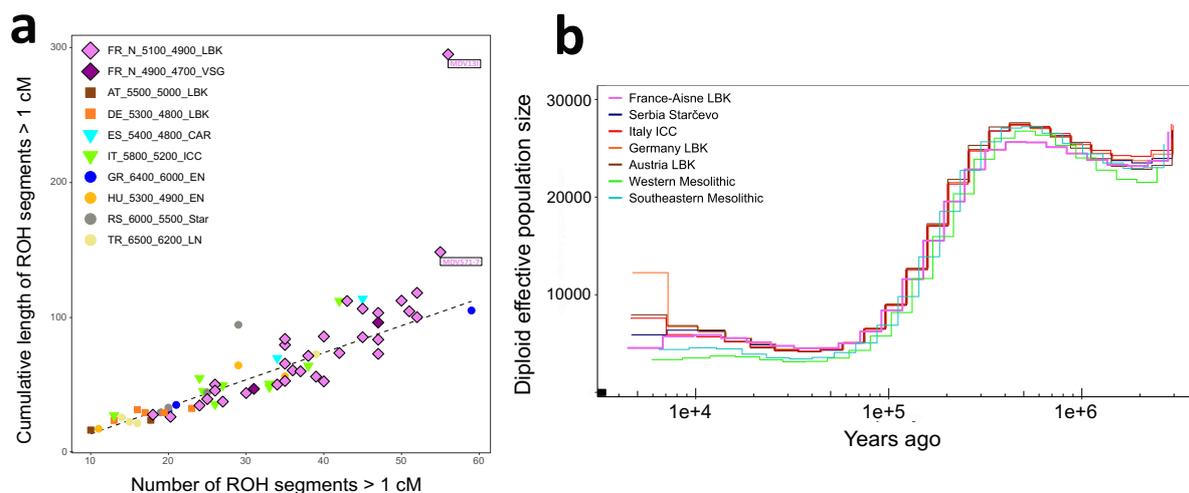


Figure 3. Demographic history of Early Neolithic European populations.

a, Runs of Homozygosity (ROH) computed for Early European farmers. The ratio of cumulative length and number of ROH segments longer than 1 centi-Morgan (cM) was used to estimate the inbreeding and background population relatedness for assessing the recent demographic history of individuals. The names of individuals referred to in the text are indicated. **b**, Past effective population size changes inferred by MSMC2. Six haplotypes per Early Neolithic or Mesolithic population were used for the estimations.

To assess the recent demographic history of Mesolithic and Early Neolithic farmer groups, we calculated the distribution of ROH segments based on their total number of segments and their cumulative lengths, restricting the analyses to segments longer than 1 centimorgan (cM) (see above). We find that the distribution of ROH segments in the LBK population from the Paris Basin varies considerably with values that are on average higher than those of the LBK individuals from Germany (Fig. 3a). This suggests that the Paris Basin was colonized by smaller groups compared to those in central Europe causing a loss of genetic diversity that occurred independently of the extent of hunter-gatherer ancestry (Fig. S6). Individual MDV13-I is exceptional in that it shows a high number of long ROH segments, which are signals of recent consanguinity (Fig. 3a). To infer past effective population size changes in Early Neolithic groups, we used Multiple Sequentially Markovian Coalescent (MSMC2), a method based on a pairwise hidden Markov model (HMM) approach, for the estimation of coalescent rates³¹. We used six haplotypes per population consisting of the individuals with the highest coverage to minimize the impact of imputation errors. The MSMC2 results show population size fluctuations over the last 50 kya (Fig. 3b), similar to those previously reported for Eurasia³². We observe a differentiation of the population dynamics between the Mesolithic and Neolithic populations

with the former experiencing only reduced post-LGM expansion and then conserving a small population size during the first half of the Holocene. The lower effective population size estimated for the western Mesolithic populations is in agreement with the differences in the ROH segment distribution between these forager and farmer groups (Fig. 3a, Fig. S3a). The population size of the southeastern Mesolithic/Epipaleolithic populations, however, increases after ~18 kya suggesting that the population dynamics in Western and Southeastern Mesolithic hunter-gatherer populations were different (Fig. 3b). This result is in agreement with the ROH size distribution in the different populations (Fig. S3a). The ancestral Neolithic farmer populations started to expand right after the LGM, earlier than what was previously reported²⁸, a trend that was never reversed up to now in most of Europe. In the most recent phase, 7-4 kya, the population size of the various Early Neolithic populations all over Europe steadily increases except in the LBK population of the Aisne Valley, whose size decreases. This decrease is corroborated by the higher proportion of ROH observed in this population (Fig. 3).

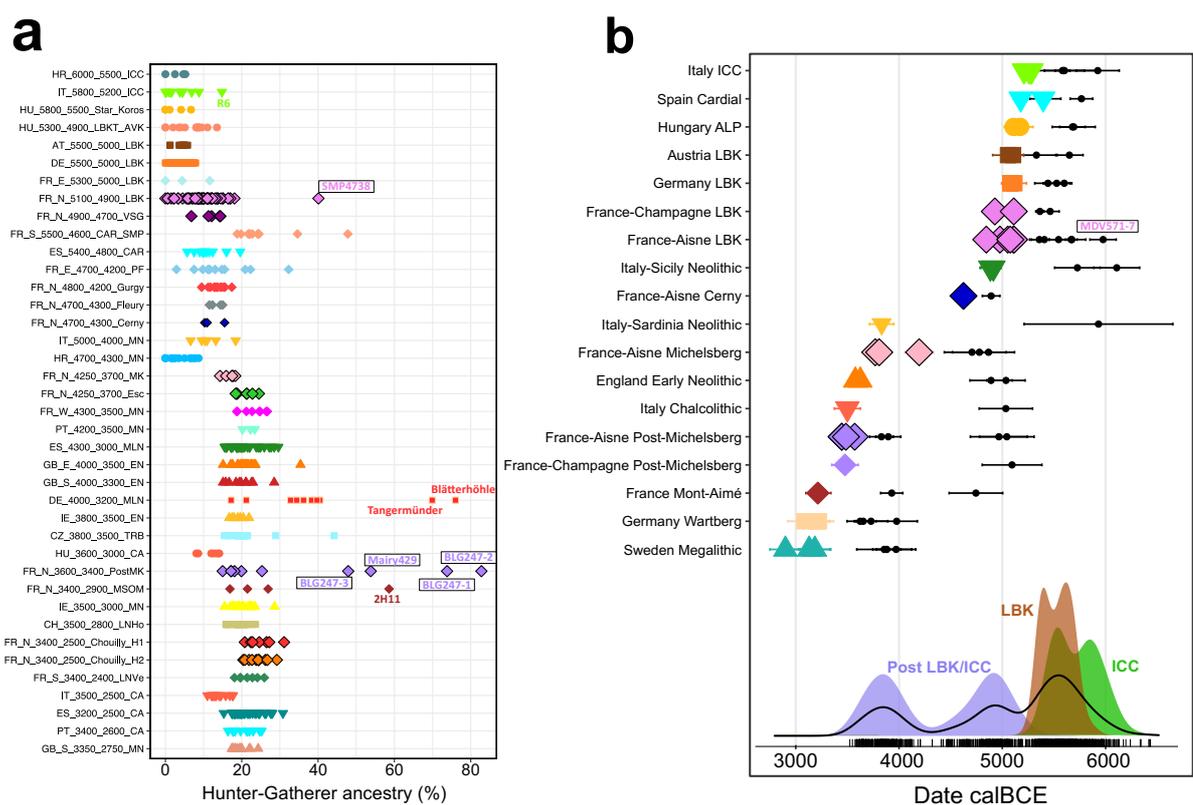


Figure 4. Evolution of hunter-gatherer ancestry in Neolithic Europeans.

a, Proportion of hunter-gatherer ancestry calculated per individual by two-way admixture modelling of qpAdm. The individuals were grouped by location, archaeological context and dates. The names of individuals referred to in the text are indicated and newly reported genomes are boxed. **b**, top: Admixture dates and 95% confidence intervals (in black) inferred

with fastGLOBETROTTER and mean radiocarbon dating values of the individuals. bottom: Density curve of inferred dates. The black line indicates the cumulative distribution computed from all genomes whereas the colored density plots correspond to individuals grouped as ICC, LBK and post LBK/ICC as indicated on the figure. The LBK outlier MDV571-7 that has an ancestry related to the Mediterranean expansion route and an admixture date shared with three other ICC was included in the ICC group for density plot calculation.

Although the archaeological evidence for the interaction between early migrating LBK farmers and autochthonous foragers is limited³³, genetic data provided clear evidence for their admixture¹⁷. To reconstruct the admixture history between early European farmers and autochthonous forager groups, we inferred the admixture dates using fastGLOBETROTTER, a robust method relying on phased haplotypes³⁴. The archaeological contexts in the LBK villages from which the analyzed individuals originate are well characterized and precisely dated based on both stylistic criteria on the pottery and radiocarbon dates. The villages in the Aisne Valley were occupied during a relatively short period of ~150 years between 5100 to 4900 BCE, i.e., for ~5 generations. Based on this narrow, reliable time frame, we inferred the dates for the admixed individuals regardless of the presence of radiocarbon dating information to have occurred in their ancestors between 9 to 33 generations before their birth corresponding to an average value of 21 generations (Table S5). Based on these data and assuming 28 years per generation, we can propose that the major admixture wave occurred on average between 5600-5400 BCE, i.e., at the time of the emergence of the LBK in Transdanubia and its early expansion into Central Europe. For the radiocarbon-dated individuals we plotted the inferred admixture dates with 95% CI obtained from a jack-knifing resampling method together with the calibrated mean age of the samples (Fig. 4b). The most recent number of generations is detected in the LBK individual SMP4738 dating to 5100 cal BCE, for whom we demonstrate elevated hunter-gatherer ancestry (Fig. 2b, 4a), indicating admixture dates between 5400-5300 cal BCE. This suggests that the admixture occurred in the ancestors of SMP4738 before they arrived in the Paris Basin. When all early European Neolithic individuals are compared, the LBK individuals from France and Germany indicate an approximate admixture date range corresponding to the early LBK expansion.

ICC individuals from Spain and Italy display diverse and recent admixture events. However, admixture dates estimated for R6, a Neolithic ICC individual from Central Italy dated to 5242 - 5177 cal BCE³⁵ with the highest hunter-gatherer ancestry of all ICC genomes

available, ~20%, (Table S4), is in the range of 5923 – 6132 calBCE, i.e., earlier than what was estimated for LBK individuals. This is in line with ICC expansions having started around 500 years earlier in the central Mediterranean³³, i.e., contemporaneous with the admixture date estimated for this individual. A similar range of admixture dates was estimated for Neolithic individuals from Sicily and Sardinia (Fig. 4b). Furthermore, recently reported admixture dates of Neolithic groups from continental Italy and Sardinia³⁶ and southwestern France¹⁸ by a method relying on the ancestry covariance pattern between the test genome (pseudo-haploid in this case) and those of two source populations coincide with these earlier admixture dates inferred herein for ICC individuals. A striking case is the woman MDV571-7, buried in the ditch enclosure of the LBK site of Menneville “Derrière le Village” in the Aisne Valley, that probably died from a blow to her head. The inferred admixture date is around 33 generations before her birth (95% CI: 28–36) around 5975 – 6102 calBCE (Fig. 4b). This range fully overlaps with the dates inferred from the Italian ICC individual R6. Moreover, MDV571-7 clusters together with the ICC individuals from Italy and Spain indicating similar genetic ancestry as the members of the Mediterranean expansion (Fig. 2c). Interestingly, the ROH distribution profile of MDV571-7 diverges slightly from the profile of the other LBK individuals of the Aisne Valley, further supporting the interpretation of a different origin of her ancestry (Fig. 3a). These different genetic lines of evidence suggest that MDV571-7 has a different origin than the other LBK individuals in the Aisne Valley. Using a method adapted to 1.2 million SNP array and pseudo-haploid data²⁶, higher ROH proportions for individuals from southern France and smaller effective population sizes compared to continental expansion route have been reported which supports the hypothesis of farming expansion via the Mediterranean route taking place in smaller-sized groups¹⁸. Therefore, it seems legitimate to hypothesize that MDV571-7 or her ancestors came from an ICC-associated context in southern Europe. This is the first reported genetic evidence of interaction between LBK and ICC groups. This individual probably had been killed and then was buried with various wild and domestic animals suggesting that she occupied a particular position in the village, possibly related to her origin.

To obtain a global view of the distribution of the admixture dates of the various genomes analyzed here, we plotted density curves of these dates together with their uncertainty distribution. When all Neolithic individuals are considered, three major admixture waves are visible centered around 5500, 4900 and 3800 BCE. Since these individuals correspond to different migrations routes and/or cultural affiliations and span a large time period, we divided them in three groups corresponding to the two European Neolithic migration routes, the central

European/LBK route and the Mediterranean/ICC route, as well as a third group corresponding to later Neolithic periods. This allowed us to achieve better resolution of the timing of admixture. The ICC-associated individuals show two admixture waves, one corresponding to the early wave discussed above, including MDV571-7, and a second wave centered around 5500 BCE. The LBK individuals also experienced two major admixture waves but occurring at distinct times, one centered around 5600 BCE, the second around 5400 BCE. The post-LBK individuals also show two peaks that will be discussed later.

We further investigated the genomic distribution of hunter-gatherer segments in admixed Neolithic genomes. We inferred local ancestry based on ELAI that uses a two-layer HMM for learning local haplotype structures³⁷. The inferred local ancestry pattern per admixed individual was plotted as an ancestry karyogram to visualize the genomic segments of hunter-gatherer or Early European farmer (EEF) ancestry (Fig 5, Fig S9-10). The “outlier” LBK individual SMP4738 has an elevated hunter-gatherer ancestry (~45%) associated with a more recent admixture date (95% CI: 8–10 generations), carries longer genomic segments of hunter-gatherer origin compared to another LBK individual from the Paris Basin, BLF148-9, with limited hunter-gatherer ancestry (~6%) and a similar number of admixture generations (95% CI: 8–15) (Fig 5a). The differences in the distribution of long segments indicate a gene flow event in the ancestors of SMP4738 that is potentially more recent than the inferred number of generations for the major admixture event that occurred at a similar date than for other individual with smaller segment sizes. This suggests that recent gene flow events leading to a few long haplotype blocks do not contribute as much to admixture date estimation with fastGLOBETROTTER than many shorter haplotype blocks do.

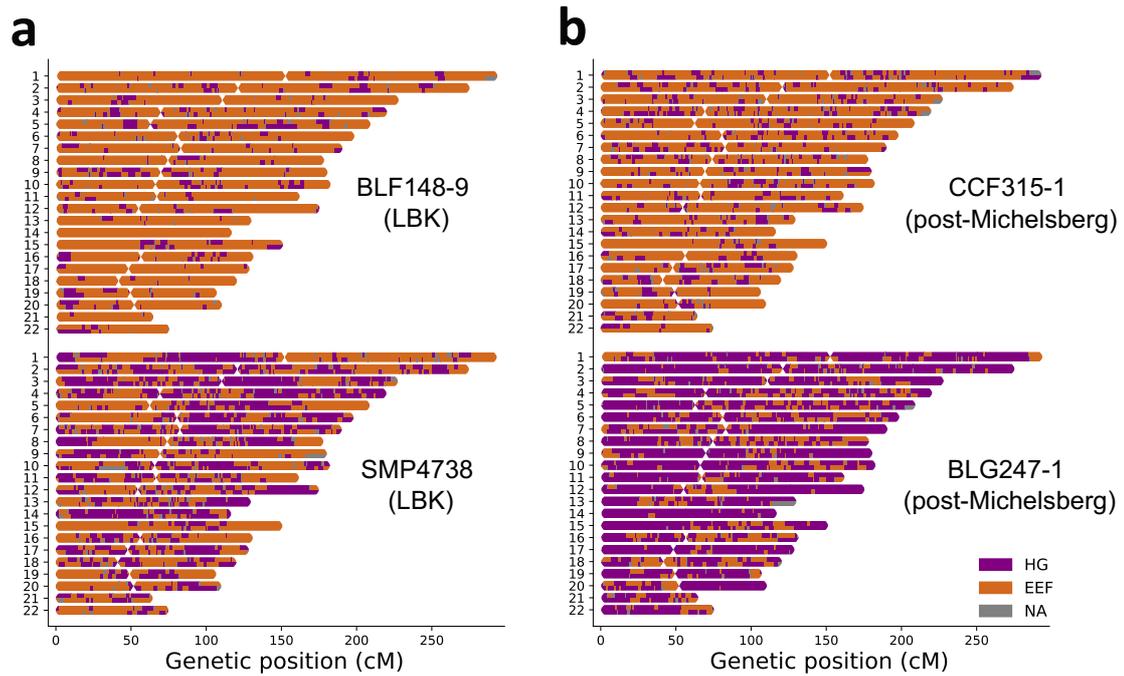


Figure 5. Genomic distribution of hunter-gatherer segments in Neolithic farmer genomes

Karyogram showing local ancestry pattern inferred with ELAI for admixed Neolithic individuals. The genomic regions with mean local ancestry dosages higher than 90% for either hunter-gatherers (HG) or Early European farmers (EEF) are shown in purple or orange, respectively. The grey color, NA, indicates the regions with less reliable local ancestry inference (<90%). Ancestry karyograms with different admixture histories in Early Neolithic (a) and Late Neolithic (b) farmers from the Paris Basin. See Figure S9 and S10 for karyograms of other Neolithic individuals.

The inference of local ancestry segments in admixed individuals allowed us to analyze the background hunter-gatherer ancestry and investigate the potential source for gene flow into Early Neolithic farmers. To do so, we extracted the hunter-gatherer segments present in admixed Neolithic genomes and constructed for each individual a new “semi-diploid” genome. The underlying procedure is that positions on segments inferred to originate from hunter-gatherers on only one haplotype are pseudo-diploidized whereas those inferred to originate from hunter-gatherers on both haplotypes remain in the diploid state. The reconstituted partial genomes were analyzed separately together with hunter-gatherer genomes filtered to keep only shared positions. The ChromoPainter2/fineSTRUCTURE pipeline was used to produce coancestry matrices. The resulting PCA plots allowed us to characterize the hunter-gatherer ancestry in Neolithic farmers. To explore the accuracy of this approach, we analyzed the hunter-gatherer ancestries of early farmers from either the continental/LBK or

Mediterranean/ICC expansions. Strikingly, we see a notable clustering between the hunter-gatherer ancestries of some individuals and local hunter-gatherers (Fig S8). The approach used is powerful enough to detect different signals of local admixture in different regions of Europe. An early Neolithic individual from Hungary (NE1), for instance, has HG ancestry that clusters closely to a local hunter-gatherer from Hungary (Fig. S8d). Likewise, a Cardial individual from southern Spain, Mur (5296-5076 cal BCE), is found to be the product of a local admixture between Cardial farmers and Iberian hunter-gatherers (Fig S8e). Not all Neolithic farmers could be associated, however, to a single, previously described, source of hunter-gatherer ancestry. For example, the hunter-gatherer ancestry of another Cardial farmer from northeastern Spain, CB13 (5469-5327 calBCE) , who lived around 200 years earlier than the Cardial individual from Mur, does not appear to be similarly close to Iberian Mesolithic individuals (Fig. S8f), extending previous observations at lower resolution³⁸. CB13 occupies a position on the PCA suggesting a dual influence of western and Iberian hunter-gatherer ancestry (Fig. S8f). We detect a hunter-gatherer ancestry with a close proximity to western European Mesolithic individuals in two Italian ICC individuals, R6 and R17 (Fig. S8a-b). The ICC individuals cluster together on the whole genome PCA despite their hunter-gatherer ancestries originating from different sources since (Fig. 2). This suggests that the main source for the differentiation between western European Neolithic farmers are not different hunter-gatherer ancestries. When this approach was applied to the LBK individuals from the Paris Basin, we detected various hunter-gatherer ancestry affinities (Fig. S7). The aforementioned SMP4738 individual, a carrier of elevated hunter-gatherer ancestry and large segments, presumably originating from recent admixture events, clusters with western European forager groups and is close to the earlier CCF711 Mesolithic individual from the same region (Fig. S7a), thus indicating a recent admixture event with local individuals with high hunter-gatherer ancestry. Interestingly, MDV571-7, who displayed ICC-like genetic ancestry, carried hunter-gatherer ancestry with closer proximity to the Iberian Mesolithic individuals than observed for most of the other Paris Basin Neolithic individuals. Another LBK individual, MDV188-4-89, has also hunter-gatherer ancestry that shows proximity to Iberian Mesolithic individuals (Fig. 7e). This individual, in contrast to MDV571-7, does not show genomic proximity to southern ICC individuals (Fig. 2c) providing additional evidence that the whole genome differentiation of western Neolithic individuals detected by the ChromoPainter PCA is not solely driven by differences in the source of HG ancestry. The complex geographical distribution of Magdalenian ancestry, which has so far been found in Iberian Mesolithic individuals (Fig. S2), is yet to be fully characterized to obtain a better understanding of its presence in farmer genomes of later periods. When the hunter-

gatherer ancestry in a given Neolithic individual occupies an intermediate position between currently described hunter-gatherer populations, we cannot establish whether this putative mixed ancestry corresponds to unsampled hunter-gatherer populations or if this composite ancestry resulted from mating between Neolithic farmers with diverse sources of hunter-gatherer ancestries. Intuitively, this latter interpretation appears more parsimonious. Our detection of differences in the timing of the major waves of HG admixture, as well as in the sources of these HG ancestry in the Neolithic farmers migrating in Europe along the continental and Mediterranean routes indicates geographical and temporal differences in the interactions between the migrating Neolithic farmers and the local Mesolithic hunter-gatherers. These differences contributed to the phylogeographical shaping of the early European genome even though the diversity of these admixture patterns does not appear to be the sole generator of genetic differentiation. Indeed, genetic drift during migration is likely playing an important role in this differentiation that cannot be simply correlated to HG ancestry in many cases.

Post-LBK populations of Northern France

After the end of the LBK around 4900 BCE, the expansion of farming and new Neolithic cultures continued to spread within Europe³³. In northern France, this process seems to be accompanied by new waves of admixture with hunter-gatherers. We could detect two main admixture waves in post-LBK western Europeans, one centered around 4900 BCE and one centered around 3800 BCE (Fig. 4b). In parallel, we did not detect Middle Neolithic farmers with little or no HG ancestry, while genomes with elevated levels of HG ancestry can be found in these later periods (Fig. 4a). This argues that individuals with high HG ancestry still existed during this later period even though the Mesolithic lifestyle is not detected anymore at this time in this area. The genomes of the individuals associated with post-LBK cultures in the Aisne Valley, including BVSG (Blicquy-Villeneuve-Saint-Germain), Cerny, Michelsberg and post-Michelsberg, cluster with the preceding LBK farmers, which argues for population continuity (Fig. 2c). The early Neolithic culture of BVSG spreading from the Paris Basin west and north and dating back from 4900/4850 to 4700/4650 BCE, is archaeologically considered as a continuation of the former LBK culture with some distinct features¹. The inferred HG admixture dates for the two contextually-dated BVSG individuals with the highest genome

coverage range between 5100 – 5600 BCE, overlapping with the dates inferred for the preceding LBK individuals (Table S5). In the following periods, a more recent date of HG admixture ranging from 4600-5100 BCE is estimated for the individuals from France Cerny, Michelsberg and post-Michelsberg culture, as well as for the Early Neolithic from England (Fig. 4b). Thus the admixture process between farmers and foragers continued during the westward expansion of farming, in agreement with archaeological evidence that supports interactions with preexisting Mesolithic groups³⁹.

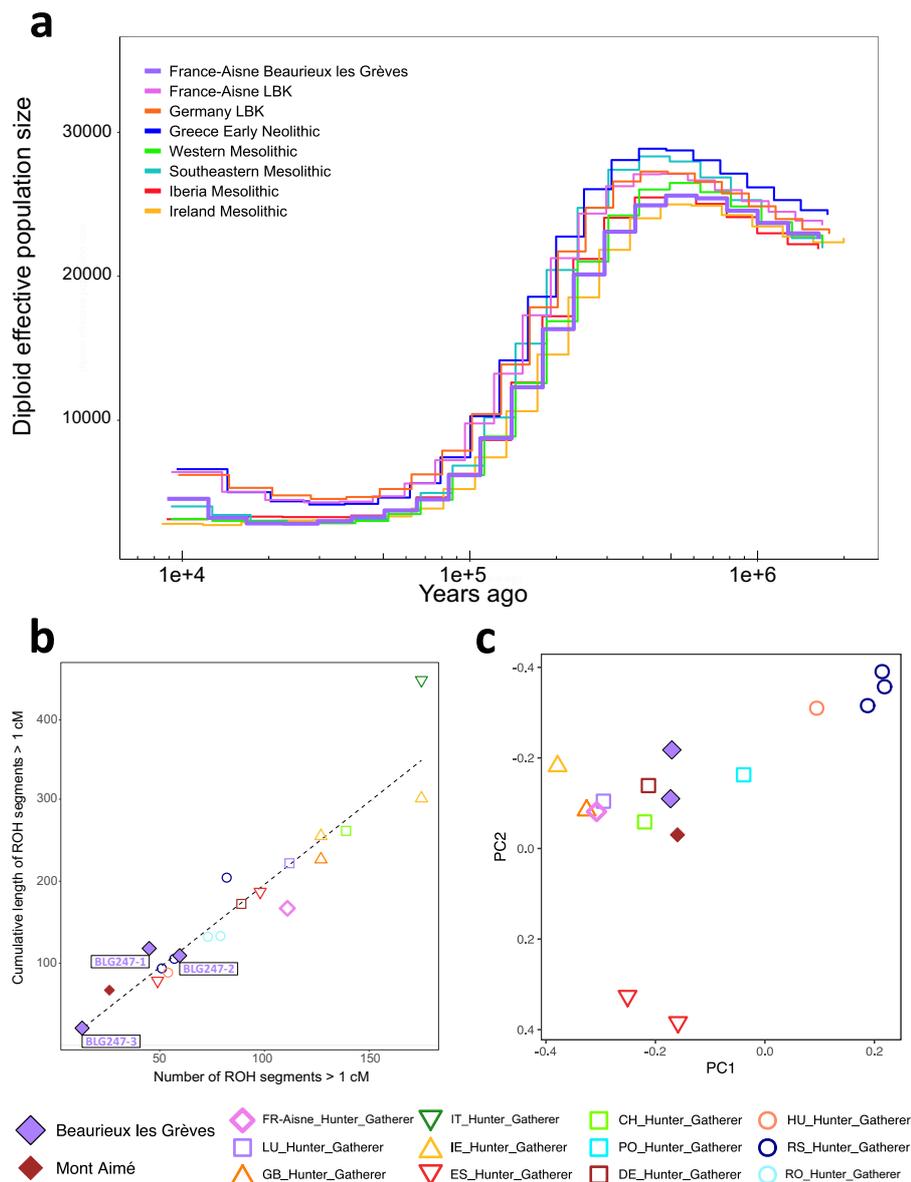


Figure 6. Population history of 4th millennium BCE Paris Basin individuals with high hunter-gatherer ancestry

a, Past effective population size changes inferred with MSMC2. Four haplotypes per population were used for Beaurieux les Grèves (BLG) as well as for the reference Mesolithic and Early Neolithic groups. **b**, ROH computed for BLG, Mont-Aimé and Mesolithic individuals. The ratio of cumulative length and number of ROH segments longer than 1 cM was used to estimate the inbreeding and background population relatedness to assess the recent demographic history. **c**, PCA based on ChromoPainter haplotypic coancestry matrix calculated on hunter-gatherer genomes and on extracted hunter-gatherer segments from BLG and Mont-Aimé individuals. The fraction of variance explained by PC1 and PC2 are, respectively, 63,8% and 9,6%.

Hunter-Gatherers of Western Europe in the 4th Millennium BCE

In western Europe the hunter-gatherer ancestry gradually increased up to 30% from early to later Neolithic periods, if one neglects a few “outlier” Early Neolithic individuals like SMP4738 and the Cardial farmers from Southern France⁷, shown to carry traces of recent admixture with forager groups (Fig 4b). There are, however, some special cases of highly elevated levels of forager ancestry detected in the 4th millennium BCE, such as the roughly contemporaneous individuals from the Blätterhöhle and Tangermünde sites in western and northern Germany, respectively^{7,17}. The Blätterhöhle had two occupation periods dating back to roughly 9000 BCE and the 4th millennium BCE, corresponding to the Mesolithic and Neolithic periods, respectively. The genomes of the four individuals from the Neolithic phase had been shown previously to contain elevated hunter-gatherer ancestry around 70% for one individual, and between 50-40% for the remaining three (Lipson et al., 2017). We report four individuals from northern France, three from the Aisne Valley dated between 3400 – 3600 BCE from the site of Beaurieux “les Grèves” (BLG), and one from Mairy in the Ardennes, who display extremely elevated hunter-gatherer ancestries ranging from 48 to 83% (Fig 1,2a-b, 4a, Table S4). A similar ancestry proportion of ~60% was previously reported for individual 2H11, from one of the hypogea of Mont-Aimé in the Paris Basin dated between 3350- 3100 calBCE (Table S4, Seguin-Orlando et al., 2021). Using fastGLOBETROTTER we inferred a strikingly recent admixture date of 4016-3774 BCE and 3945-3714 BCE for the two individuals from BLG247 with the highest genome coverage (Fig 4b, Table S6). We estimated a very similar range of admixture dates, between 4035 – 3822 BCE, for the ~200 year-younger individual 2H11 from the Mont-Aimé that carried a high hunter-gatherer ancestry proportion whereas

individual 2H10 yielded an admixture dating range similar to the post-LBK Neolithic individuals from the Aisne Valley (Fig. 4b). Furthermore, the recently reported admixture dates of one Blätterhöhle individual based on the ancestry covariance pattern are in the range between 3944-3720 BCE³⁶, which coincides with the inferred admixture dates for the individuals from Beaurieux “les Grèves” and Mont-Aimé in the Paris Basin. When the genomic distributions of ancestral segments in contemporary post-Michelsberg individuals are compared, the differences in size and number of hunter-gatherer segments are striking (Fig. 5b). CCF315-1, whose admixture date was estimated to ~5000 BCE within the first post LBK/ICC admixture wave (Fig. 4b, Table S6), has much smaller hunter-gatherer segments compared to BLG247-1 (Fig. 5b) and 2H11 (Fig. S10f). This allows us to hypothesize that, during the early 4th millennium BCE, individuals who had preserved their hunter-gatherer genomes since the Mesolithic period started to receive gene flow from Neolithic farmers. This phenomenon implies interactions between groups with distinct ancestries that have not been detected by archaeological approaches in this region at this period, presumably due to the lack of distinctive cultural features. In the case of the Blätterhöhle individuals, stable isotope analyses showed two types of diet for different individuals, one being a typical Mesolithic diet with a preference for freshwater fish, the other being a typical Neolithic diet with a preference for terrestrial resources including domestic animals and crops⁴⁰. The protein-based diet coming from freshwater fish was preferred by the individual with the highest hunter-gatherer ancestry, while the others had a typical Neolithic diet. These roughly contemporary individuals might therefore be witnesses of the admixture process followed by acculturation of the hunter-gatherers. The special case of Beaurieux “les Grèves”, where all three excavated individuals showed highly elevated hunter-gatherer ancestries, indicates that their hunter-gatherer ancestry was not further diluted. This means that the members of these groups continued their lives as a distinct community with preferential intra-group mating behavior even if admixture events with Neolithic farmers took place several generations before. Interestingly, a similar range of admixture dates were inferred for later Neolithic individuals in Germany and Sweden (Fig. 4b). The Wartberg group from central Germany also shows elevated levels of hunter-gatherer ancestry (Fig. 2a-b, Fig 4a), indicating that the segregation of groups with distinct ancestry and preferential intra-community mating behavior was not uncommon during this period. This segregation behavior may be based on communication barriers, since Mesolithic hunter-gatherers and Neolithic farmers did not speak the same language, as observed also for present-day populations⁴¹, as well as on other cultural attributes that are specific for each community.

To further investigate the population history at Beaurieux “les Grèves” we used MSMC2 with four haplotypes to analyze the two BLG genomes with the highest hunter-gatherer ancestry together with different Mesolithic and Early Neolithic genomes (Fig 6a). The estimation of the dynamics over the last million of years of the ancestral population of the two BLG individuals reveals that it evolved in parallel to other Mesolithic populations. This gives support to our interpretation that these individuals were genetically hunter-gatherers who received recent gene flow from Neolithic farmers. At the post-LGM period starting from ~ 18 kya, the size of the ancestral population of the BLG247 ancestral group increased to the same degree as the southeastern Mesolithic population (Fig 6a). This apparent population increase may simply be the consequence of the admixture with Neolithic farmer ancestry that increases mechanically the heterozygosity.

We calculated the ROH distribution of the BLG247 and Mesolithic individuals to assess their recent demographic history (Fig. 6b). The ROH distribution of cumulative length and the total segment number have similar values in the two BLG individuals with high hunter-gatherer ancestry and some southeastern Mesolithic individuals, the latter showing relatively fewer and smaller ROH segments compared to Western Mesolithic individuals. Interestingly, the calculations for BLG247-3, who carries $\sim 48\%$ hunter-gatherer ancestry, indicate extremely low ROH numbers similar to Central European LBK individuals. Although the admixture dates propose a common origin for BLG individuals, there is a striking difference in the ROH distribution of BLG247-3 compared to the other two BLG individuals, whose ROH values are closer to those of western Mesolithic individuals. This may result from a recent admixture event with farmers with high Neolithic ancestry that abolished most ROH originating from the HG lineages. Lastly, we performed the same local ancestry inference-based method described above on the two high coverage BLG247 individuals together with the Mont-Aimé 2H11 individual by merging the genomic segments from hunter-gatherer ancestry as described previously.. The hunter-gatherer fraction of these genomes show genetic proximity to hunter-gatherers from the west, not to the more western hunter gatherers but rather to the earlier Mesolithic individual from Bad Dürrenberg, eastern Germany, dated to around 6,600 BCE, or to hunter-gatherers further to the east (Fig. 6c). Indeed, the Bad Dürrenberg Mesolithic genome is similar to more

western Mesolithic genomes, represented by the Aisne Valley CCF711 and the Bichon, Loschbour, and Cheddar individuals, but is nevertheless distinct as it shows some proximity with a Polish Mesolithic individual N22 (Fig. 6c and S2b). It is striking that the SMP4738 with high HG ancestry, dated around 5100 BCE, seems to have received a large part of this ancestry from individuals similar to the local Mesolithic individual CCF711, dated around 6,200 BCE, while 1,700 years later, in the same area, the contributing hunter gatherer ancestry is closer to what was found more eastward. This suggests that the populations that have preserved their Mesolithic ancestry for almost 2,000 years after the arrival of Neolithic farmers and persisted with little admixture during the 4th millennium BCE may originate from Central Europe. These populations may as well have contributed to the high hunter gatherer ancestry found in the Blätterhöhle and Tangermünde individuals that show similar timing and extent of admixture as the BLG247 individuals. Since the former originate from Central Europe such an ancestry is less intriguing. In the pigmentation pattern test using HIrisPlex-S, the BLG247 individuals were positive for the genetic markers associated with lighter skin pigmentation and blue eyes while most of the other Neolithic individuals analyzed carried the SNPs associated with brown eyes, dark brown to black hair and intermediate to dark skin (Table S7).

The Late Neolithic in northern France

We analyzed several Late Neolithic genomes from northern France. The study of a family burial from Bréviandes in the Aube region close to Troyes, Champagne, is reported separately since it concerns the arrival of the steppe ancestry in northern France (Parasayan et al., in preparation). We also analyzed burials from two hypogea at Chouilly « la Grifaine » in the Marne Valley south of Reims dating to between 3400-3300 and 2900-2800 BCE, i.e., during the Late Neolithic (Table S1). The cemetery that includes these and other hypogea was in use over ~900 years. The genomes of the slightly older hypogea F10 with burials covering the period between 3400-2800 BCE, cluster with the Neolithic genomes from northern France while the genomes of the slightly younger hypogea F11 with burials dated between 3400-2500 BCE cluster with Neolithic genomes from southern France and the Iberian Peninsula (Fig. 2c, S4b, S11, S12). This suggests that two different populations were inhumated at this location, the more recent one having a different origin. This is in line with our findings from Bréviandes

since we detected southern Neolithic ancestry in northern France during the 3rd millennium BCE as well as patterns of mobility for some individuals. Another line of evidence supporting a different origin for these two hypogea is provided by the dates of admixture with HG inferred by fastGLOBETROTTER. The number of generations estimated for hypogea F10 (39 ± 15 generations; Table S5) is smaller than for hypogea F11 (97 ± 28 generations; Table S5), which also supports different source populations for these two hypogea, in agreement with our aforementioned observation of an earlier date of admixture with HGs for Neolithic farmers who migrated along the Mediterranean route than for those who migrated along the Danubian route.

The genetics behind the spread of farming into the British Isles

Farming spread into the British Isles with a delay of ~ 1000 years. The mechanisms of the spread are still elusive despite being frequently addressed by both archaeology and genetics. Possible migration routes were proposed based on archaeological evidence^{42,43} as well as possible source populations of the migrating groups^{7,16,30,44,45}. Most of these genetic studies concluded that the Neolithic individuals from the British Isles are descendants of Mediterranean farmers. Coming from Iberia, these farmers would have reached northern France and the British Isles via the Atlantic seaboard. These conclusions were reached through various methods based on the formal analyses of both independent SNPs and haplotypes from Neolithic farmer genomes from Germany, as the representatives of the central European expansion route, and those from the Iberian Peninsula and/or southern France. The strength of these conclusions, however, was limited owing to the lack of Neolithic genomes from northern France. The present study fills this gap since it provides both a large number of Neolithic genomes from northern France and the high resolution of haplotype-based methods enabling us to differentiate with an unprecedented precision northern and southern Neolithic groups. In addition to the Aisne Valley, we analyzed a population from the Middle Neolithic site of Escalles “Mont d’Hubert”, Pas-de-Calais, the Flanders region of Northern France, close to the “Cap blanc Nez” southwest of Calais, dated to 4230-3879 BCE. Not only is this the closest continental site to the British coast analyzed paleogenomically, just in view of present-day Dover, but the corresponding genomes also form a distinct and tight cluster with the Neolithic populations from the British

Isles to the exclusion of all other Neolithic genomes (Fig 2, Fig S4). This pattern is striking in both PCAs whether they include hunter-gatherers or not (Fig. 2a and c). The latter PCA has a higher resolution of Neolithic genomes broadening only slightly the Escalles-British cluster (Fig. 2c and Fig. S4b). Thus, the genomes from Escalles and the British Isles cluster together regardless of the sample size. To explore further the possibility of a continental European origin for the cluster formed by Escalles and British Isles individuals, we successively excluded from the dataset either one of these two related groups to minimize the effect of affinity induced by a large number of individuals with similar ancestry. Both analyses yielded the same result: the two groups cluster with the Aisne Valley Neolithic genomes when only one of them is present (Fig. S11). In all these analyses, the genetic affinity between the Neolithic individuals of the British Isles and the descendants of the Mediterranean farmers is lower than with the individuals from northeastern France. This demonstrates both the genetic affinity between the British group and the continental Neolithic farmers in northern France and the affinity between the Neolithic groups in northern France. Therefore, we propose the Middle Neolithic population in the Flanders region to be the source population of the Neolithic farmers who colonized England. This hypothesis is also consistent with the close geographic proximity between the coast of Flanders and the coast of southeast England. Nevertheless, the possibility remains that an unsampled group from more northern areas could be the source population, in particular in light of the strong sea currents of the Channel between the “Cap Blanc Nez” (Calais) and Dover that may constitute an obstacle for navigation.

To extend our analyses and include the much larger published dataset generated via in-solution hybridization capture methods using the 1240k SNP array, we had to use less powerful analysis methods than the haplotype-based methods on phased shotgun genomes used above. We performed MDS analyses calculated on distance matrices coming from the inverted pairwise outgroup f_3 -statistics. To minimize the effect of allelic biases in the 1240k enrichment dataset and enable a better, less biased co-analysis of shotgun and 1240k capture data, we called the allelic positions in capture and shotgun genomes using a restricted SNP set of ~460,000 positions with the least capture-induced reference biases⁴⁶. Moreover, to prevent any bias that could result from excessive relatedness caused by known methodological artefacts of the capture protocol⁴⁶, we used a single captured population at a time to compare it with our well-covered shotgun genome set. Although, expectedly, a method based on unlinked genetic loci cannot provide a resolution as high as the haplotype-based method, our MDS results still enable us to

identify successfully the clusters formed by populations from the British Isles and Escalles (Fig. S12-14). In this way we could test whether there are any other capture-genotyped populations that behave in a similar way as the Escalles group. First, we demonstrate that Early Neolithic capture-genotyped genomes from Britain and Germany cluster together with the corresponding populations generated via shotgun sequencing (Fig. S12a, b). We also included into these analyses capture-genotyped populations from France (Fig. S13), Iberia and Italy (Fig. S14). All these capture-genotyped populations cluster with their corresponding shotgun genotyped populations. These different MDS analyses thus reveal that the well-covered shotgun genomes in our set are representative of the Neolithic European genome diversity known from the sequence capture analyses. Furthermore, our analyses reveal that no other ancient population, whose genomic data are so far published, shows a level of genetic affinity with the Neolithic genomes of the British Isles as important as the Neolithic population from Escalles. We cannot exclude completely, however, a scenario where the formation of the genetic structure detected in Escalles would also involve some low-level gene flow from a Mediterranean-related Neolithic population into a northern Neolithic population, even though the affinity of the Escalles population with the Aisne valley population is clearly visible in the absence of the closer British Isles populations and the affinity of this British Isles population with the Mediterranean population is always much lower than it is with the Northern French Neolithic population, even when the Escalles population is not considered (Fig S11). Thus the previously proposed strong contribution from Mediterranean populations to the Neolithic population on the British Isles was due to the poor representation of the Neolithic population from northern France. Our analysis reveals the progressive evolution of the genome structure of the Neolithic population that occurred during the westward trans-European migration and further evolved in northern France. Moreover, it uncovers the combined effects on the shaping of the source population structure of the early Neolithic colonizers of the British Isles: these are (i) migration-induced genetic drift, (ii) admixture with local hunter-gatherer populations at various time periods and geographic locations and (iii) interactions of the Danubian/LBK and the Mediterranean/ICC Neolithic populations that met in northern France. Thus, we propose the Escalles population or a genetically close population living in this area to have crossed the Channel and spread in the British Isles.

Material & Methods

DNA extraction and Library Construction

Ancient DNA work was carried out in a high containment laboratory as described⁴⁷. The dense pyramidal parts of the petrous bones were isolated using a flame-sterilized diamond-disc of a Dremel. Bone pieces were then ground to fine powder in a 6775 Freezer/Mill®Spex SamplePrep in liquid nitrogen as described previously⁴⁷. About 100 mg of bone powder were washed three times for 15 min. each with 1 ml 0.5 M sodium phosphate buffer at pH 8.0 Each sample was rinsed with 1 ml TT buffer made of 10 mM Tris HCl pH 8.0 and 0.1% Tween 20 One ml extraction buffer (0.5 M EDTA pH 8.0, 0.05% Tween 20, 250 µg/ml Proteinase K, and 0.14 M β-mercaptoethanol) was added to each sample and incubated on a rotating wheel for ~30 hours at 37°C. After this first incubation, the samples were centrifuged at 6000 × g for 1 minute. The supernatant was removed and kept for purification. Another 1 ml extraction buffer was added to the pellet and after resuspension, incubated for another 30 hours on a rotating wheel at 37°C. After this second incubation step, all samples were centrifuged at 6000×g for 1 minute. The supernatants from both incubation periods were pooled together in 2 ml tubes, mixed well and then centrifuged at maximum speed for 2 minutes. One ml of the supernatant was used for the following purification step using 5M40 binding buffer (5M Guanidine HCl, 40% isopropanol, 120 mM Sodium Acetate (NaAc) pH 5.2, 0.05% Tween 20;. 1 ml of each sample extract was purified with ~10 ml 5M40 binding buffer through Qiagen columns under vacuum, rinsed twice with 1 ml Buffer PE (Qiagen) and eluted with twice 30 µl of TET buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.05% Tween 20) following l Libraries were constructed using 6 µl of extract and the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, Ipswich, MA, USA) after a pretreatment with USER enzyme mix (NEB, Ipswich, MA, USA) as described by the manufacturer. An unindexed Y-shaped adapter was ligated, the ligated libraries were first amplified for 8 cycles with the Illumina Rd1 and Rd2 sequencing primers, then 20% of the library was further amplified for 12 cycles with Illumina unique dual indexed adapter primers. The libraries were purified using 1.3 volume of SPRI beads (Macherey-Nagel™ NucleoMag™ NGS Clean-up and Size Select) for two successive rounds of purification.

Processing of Sequencing Data

All data processing steps were performed using shell instructions and in-house written scripts. The raw output of Illumina sequencers was in the BCL format. Because of the high number of in-house barcodes used, we performed the initial demultiplexing and fastq conversion using `bcl2fastq` (Illumina) on the laboratory server. Sequence quality was assessed using `FastQC` ⁴⁸. Paired-end reads were merged and trimmed for adapters using `leeHom`⁴⁹. Reads larger than 28 bp were selected using `awk` and mapped onto the Human Reference Genome `hs37d5` using `BWA` (Burrows-Wheeler Aligner) `aln` version 0.7.17⁵⁰ with a seed length of 18. Aligned sequences in `BAM` format were processed with `Samtools` v.1.9 ⁵¹ to create index files and `flagstat` reports, and to select and sort mapped reads. `Picard` tools (<http://broadinstitute.github.io/picard/>) was used to remove duplicates. Multiple libraries were merged by `Samtools` v.1.9 before, followed by filtering the mapping quality (`-q 20`). Damage patterns were analyzed using `mapDamage` v.2.2.1 followed by end-clipping using `BamUtil` (https://genome.sph.umich.edu/wiki/BamUtil:_trimBam). According to the damage pattern, three to four bases at the ends of each read were clipped for the UDG-treated libraries and 10 bases for non-UDG treated libraries. Five Mbuti genome sequences (SRR1157019, SRR1157021, SRR1157023, SRR1157024, and SRR1157055) were mapped onto the `hs37d5` human genome using `bwa mem` deduplicated and filtered for a mapping quality higher than 30.

Sex Identification, Uniparental Haplogroups and Contamination Estimation

To determine the sex of each individual, chromosome counts were performed using index statistics (`idxstats`) of `Samtools` on `bam` files by using exclusively the reads with mapping quality higher than 20 to prevent confusing effects of repetitive regions. Then the ratio of the proportion of read counts normalized to chromosome length for both the X and Y, normalized for each to the average of the autosomal proportion was used. (Table S1). Mitochondrial variants were called using `bcftools mpileup` (options `-B -q30 -Q30`) and `call` (option `-ploidy GRCh37`) for all individuals. The output `mtDNA vcf` file was used as input to run `HaploGrouper` ⁵² following the example provided online (https://gitlab.com/bio_anth_decode/haploGrouper) (Table S1). The clipped `bam` files of male individuals were used as inputs to determine Y haplogroups by

placing alleles onto reference phylogenies using pathPhynder¹¹ following the example provided online (<https://github.com/ruidlpm/pathPhynder/tree/master/tutorial>) (Table S1). The clipped bam files were used as the inputs to run Haplocheck⁵³ and ANGSD⁵⁴ software to estimate contamination on the basis of mtDNA for all individuals and X chromosome for male individuals, respectively (Table S1).

Genotype Calling and Dataset Preparation

We followed two different genotyping approaches, which are genotype imputation and pseudo-haploid calling, in order to prevent the possible false positive calls that could come from direct diploid genotyping due to low coverage of the majority of ancient samples. To generate an imputed and phased dataset, GLIMPSE (v.1.1.1)⁵⁵ was used for imputing the ancient genomes, namely 78 individuals from the current study and 206 publicly available ancient individuals (Table S3). We followed the procedure described in the online Glimpse tutorial (https://odelaneau.github.io/GLIMPSE/tutorial_hg19.html), using Genotype Likelihoods (GL) computed with Bcftools (mpileup -I -E, call -Aim parameters) using the 1000 Genomes Project phase 3 reference panel excluding the indels. We filtered out from the imputed genomes the positions that have a MAF lower than 5% in the 1000G reference panel. Then, we applied a further filtration of Genotype Posterior Probability (GP) lower than 0.99 and genotype missingness rate of 0 for downstream analyses. We use the autosomal sites in the 1240k SNP capture⁵⁶ to select 1240k SNPs that were used for calling, with Samtools mpileup and the pileupCaller pipeline (<https://github.com/stschiff/sequenceTools>), the bam files resulting from either shotgun sequencing or in-solution target capture. To minimize the effect of allelic biases in the 1240k enrichment dataset and enable a better, less biased co-analysis of shotgun and 1240k capture data, we called the allelic positions in capture and shotgun genomes on a restricted SNP set of ~460,000 positions (460k) with the least bias toward enriched positions⁴⁶.

Estimation of Kinship, Pigmentation Profile and Runs of Homozygosity (ROH)

We inferred the relatedness between individuals coming from different Neolithic sites described in the current study using NgsRelate¹². We used a mixed group of 250 ancient European individuals covering the periods of the Early to Late Neolithic to provide a distribution of ancestral proportions in the background allele frequency similar to that found in Neolithic sites analyzed in this study. For hunter-gatherers including admixed individuals with the hunter-gatherer ancestry $> \sim 50\%$ we prepared another set of 30 individuals to prevent the bias that can be introduced by elevated levels of hunter-gatherer ancestry. The NgsRelate analysis was performed using as input vcf files containing the filtered imputed genotypes of these 250 and 30 individuals. The estimated kinship (theta) and Cotterman (k0, k1, k2) coefficients of Neolithic individuals reported in this study (Fig. 1c, Table S2) were compared to the ideal values obtained using simulations by Yaka et al., 2021. The kinship relationships described in the text were determined by these analyses, the uniparental haplogroups (mitogenomes and Y chromosomes), the genetic sex and the estimation of the anthropological age at death (Table S1).

In order to predict the pigmentation profile of the newly reported individuals, we used Hirisplex-S²⁹ and the imputed genotypes (Table S7). After filtration of GP and missing variants, described above, plink v1.9⁵⁷ was used in order to identify the proportions of runs of homozygosity in ancient genomes (Fig. 3a, S3) with the following parameters;

```
--homozyg --homozyg-density 50 --homozyg-gap 100 --homozyg-kb 500 --homozyg-snp 50  
--homozyg-window-het 1 --homozyg-window-snp 50 --homozyg-window-threshold 0.05
```

The physical positions estimated by plink was converted to genetic positions using the genetic maps built for hg19.

Population Genetic Analyses

We ran ChromoPainter/fineSTRUCTURE¹⁹ on the imputed and phased data using different sets of all individuals (Fig. 2a), excluding hunter-gatherer rich genomes (Fig. 2c), as described in the text, or using only hunter-gatherers (Fig S2)⁵⁸. After estimation of switch and mutation rates using 10 expectation-maximization (EM) iterations, each haplotype was painted using all the other haplotypes (-a 0 0). We applied fineSTRUCTURE's Markov chain Monte Carlo (MCMC) model on the resulting coancestry matrices with 3,000,000 burn-in and 2,000,000 sampling iterations. After the extraction of the maximum observed posterior probability state and 100,000 hill-climbing iterations, the tree was inferred by using the maximum concordance state as it is used in the PoBI study⁵⁸. To obtain input files for SOURCEFIND v2⁵⁹ and fastGLOBETROTTER³⁴ we ran ChromoPainter v2 with 10 iterations of EM to infer parameters on chromosomes 1, 8, 15 and 22. Then, the copy vector and painting samples files were generated for the 22 autosomes by using the inferred parameters and the ancestral groups, EEF (Bar31, Bar8, LEPE48-2, LEPE52-1, Nea2, Nea3, Klein7, I0707, I0745, I0709, I0708, I2521) and HG (CCF711, Loschbour, Cheddar, Bichon, SRA62, Chan, LaBrana, KO1, I5236, VLASA32, VLASA7, N22, BDB001), as donor populations to paint recipient individuals. An individual-based dataset was prepared including only test individual together with EEF and HG groups. This allowed us to keep only shared genomic positions between tested individuals and ancestral groups after GP and missingness filtrations without any loss of data due to unshared positions between tested individuals. We ran fastGLOBETROTTER to infer the admixture dates on each individual dataset by using HG and EEF ancestral populations as both donor and surrogate populations followed by a jackknife resampling step as implemented in the software in order to calculate the confidence intervals (Fig. 4b, Table S5-6). These CIs were calculated using the delete-mj jackknife variance estimator formula of Busing et al.⁶⁰. Density curves were computed using admixture dates calculated from the number of inferred generations per autosome of individuals that were estimated by the jackknife resampling step of fastGLOBETROTTER.

We inferred effective population sizes of Neolithic and Mesolithic populations using MSMC2³¹. The input files for two or three diploid individuals were generated from quality filtered imputed and phased genomes using the generate_multihetsep.py script from MSMC-tools (<https://github.com/stschiff/msmc-tools>) and masked for high mappability regions per chromosome downloaded from (<https://github.com/wangke16/MSMC-IM/tree/master/masks>). MSMC2 was run on multihetsep.txt input files with four and six

haplotypes using (msmc2 -t11 -I 0,1,2,3 -s -o pop.4haps.msmc2.txt pop.4haps.chr{1..22}.multihetsep.txt) and msmc2 -t11 -I 0,1,2,3,4,5 -s -o pop.6haps.msmc2.txt pop.6haps.chr{1..22}.multihetsep.txt), respectively (Fig. 3b, 6a).

Populations represented with 6 haplotypes are

France-Aisne LBK : BLF148-9, MDV233, MDV93.2 ; Germany LBK : LBK, Dil16, Ess7 ; Austria LBK : Asp6, Klein7, I5069 ; Italian ICC : R9, R3, R2 ; Serbia Starcevo : I4918, STAR1, VC3-2 ; Western Mesolithic : CCF711, Loschbour, Cheddar ; Southeastern Mesolithic : VLASA32, VLASA7, I5236

Populations represented with 4 haplotypes are

France-Aisne LBK : BLG247-1, BLG247-2 ; France-Aisne LBK : BLF148-9, MDV93.2 ; Germany LBK : Herx, LBK ; Greece Early Neolithic : Nea2, Nea3 ; Western Mesolithic : CCF711, Loschbour ; Southeastern Mesolithic: VLASA32, VLASA7 ; Ireland Mesolithic: SRA62, KGH6 ; Iberia Mesolithic: Chan, LaBrana

Local ancestry inference was performed using ELAI³⁷ on the individual datasets as described above for fastGLOBETROTTER. Admixed target individual and source populations, EEf and HG, were extracted from each dataset as vcf files and then manually converted into phased BIMBAM format that is described in the ELAI user manual (<https://github.com/haplotype/elai>). ELAI was run on each admixed individual per chromosome as follows;

```
elai -g EEf.chr1.geno -p 10 -g HG.chr1.geno -p 11 -g admixed.chr1.geno -p 1 -pos admixed.chr1.pos -C 2 -c 10 -mg #generation.
```

Output mean local ancestry files (ps21) from ELAI were used in order to plot ancestry karyograms for admixed individuals (Fig. 5, Fig S9-10). Both HG and EEf segments of admixed individuals that are defined by local ancestry dosages > 90% were manually converted into breakpoint file format using awk for haptools described in (<https://haptools.readthedocs.io/en/latest/formats/breakpoints.html>). Then, we used haptools karyogram to plot these breakpoints files as ancestry karyograms (<https://haptools.readthedocs.io/en/latest/commands/karyogram.html>). The regions < 90% mean local ancestry dosages are indicated as NA.

The same mean local ancestry files (ps21) were filtered keeping the positions with HG ancestry dosage > 90% per haplotype. These HG positions were extracted from each test individual as both diploid and haploid segments and these haploid segments were manually pseudo-diploidized using awk to reconstruct a “semi-diploid” genome from the hunter-gatherer segments of admixed individuals. These reconstructed segments were merged with hunter-gatherer genomes and only shared positions were kept. Each resulting dataset were individually analyzed with ChromoPainterv2/fineSTRUCTURE pipeline as described above (Fig S7-8). Only the hunter-gatherer-segments coming from individuals with elevated ancestries for BLG247-1, BLG247-3 and 2H11 were merged in a single dataset as their high ancestry proportions (See Fig 5b, 9e ,10f) resulting in many shared regions allowed us to analyze them simultaneously (Fig. 6c).

Admixture modelling and outgroup-f₃ statistics were computed using qpAdm⁶¹ and qp3Pop programs from the ADMIXTOOLS package (version 5.0)⁶² for individuals coming from both shotgun sequencing and in-solution target capture data. We used Mbuti, Ust'-Ishim⁶³, Kostenki14⁶⁴, Sunghir3⁶⁵, Karelia⁶⁶, Latvian hunter-gatherers⁶⁷, and Iranian Neolithic²⁰ genomes as the reference groups to calculate admixture proportions for the same EEF and HG source populations in the 1240k SNP set (Fig. 4a, Table S4). The outgroup-f₃ statistics were calculated in a pairwise manner in the form of f₃(Mbuti; ind1, ind2) with 460k SNP. Distance matrices coming from inverted pairwise values (1-f₃) were used to compute Multi-Dimensional Scaling (MDS) analysis (cmdscale package in R) (Fig. S12-14).

Supplementary Materials

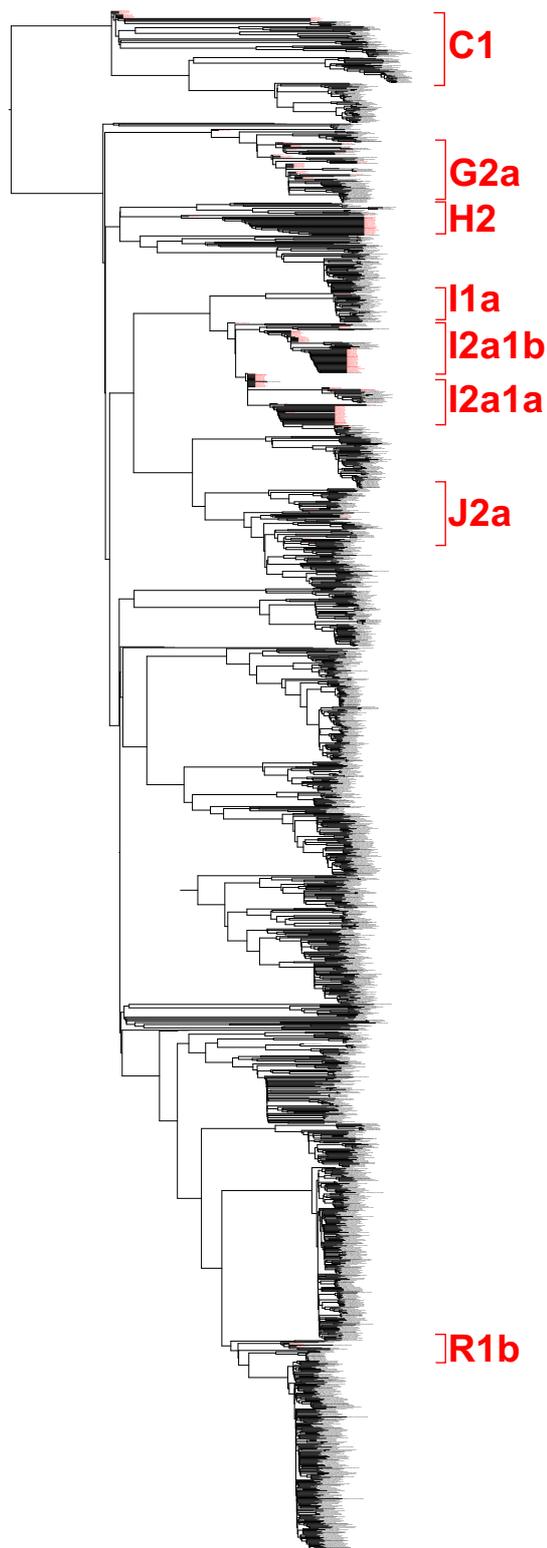


Figure S1. Y chromosome phylogeny of modern and ancient individuals.

Ancient sequences of newly reported and previously published individuals were placed into reference Y chromosome phylogenies with pathPhynder using the best-path workflow. The major haplogroups of male individuals analyzed herein are indicated.

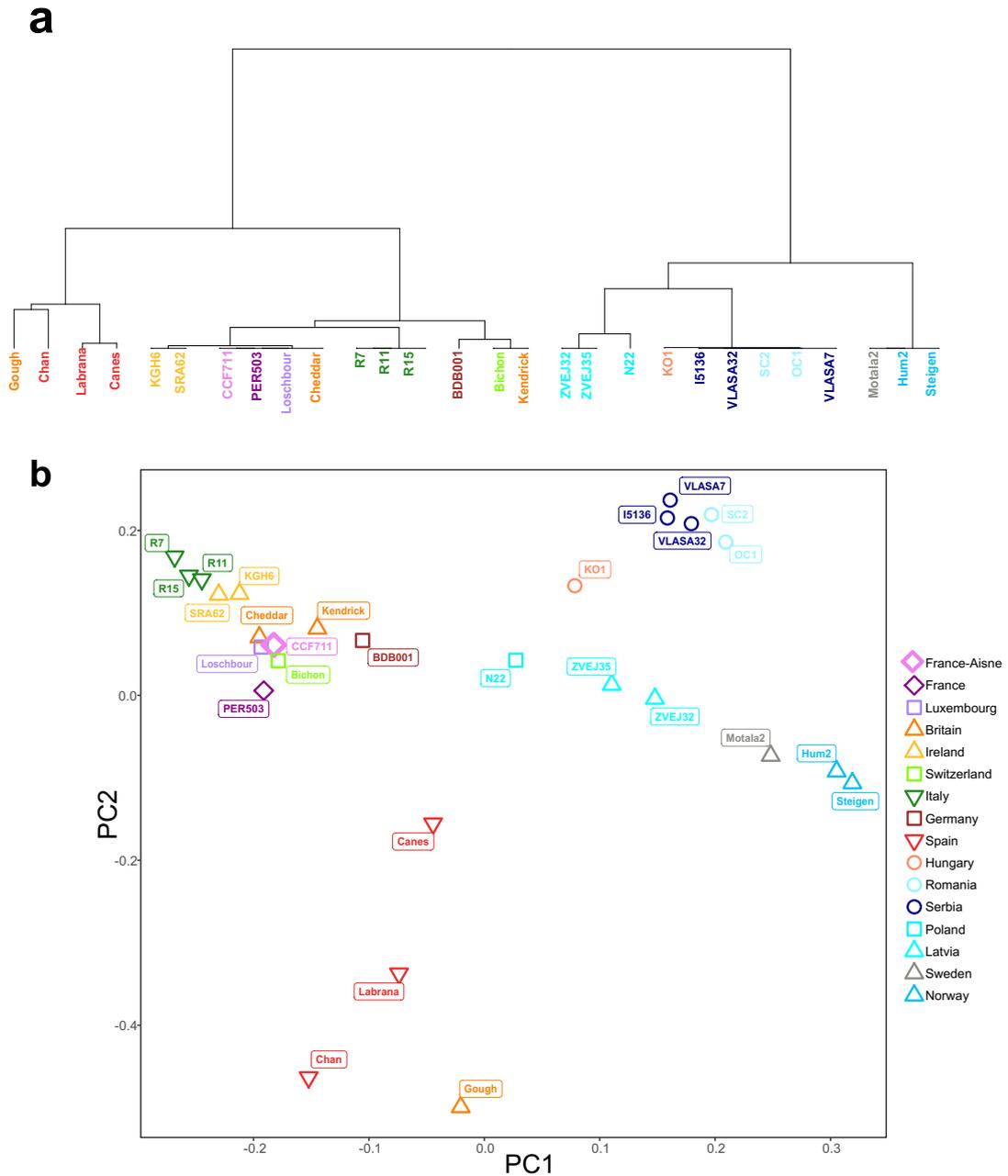


Figure S2. Genetic affinities of hunter-gatherer groups.

a, fineSTRUCTURE clustering dendrogram based on the same coancestry matrix. When compared to Fig. 2a, Late Paleolithic hunter-gatherers from Britain, Gough and Kendrick, were added. In order to better define various ancestral source populations found during the Mesolithic, we also included two recently reported into our analysis. **b**, PCA based on the ChromoPainter haplotypic coancestry matrix calculated on the dataset including only hunter-gatherer genomes with the addition of Late Glacial Paleolithic genomes from Gough and Kendrick's caves in Britain dating to approximately 14,8 and 13,7 kya, respectively. The fraction of variance explained by PC1 and PC2 is, respectively, 55.2% and 13.6%.

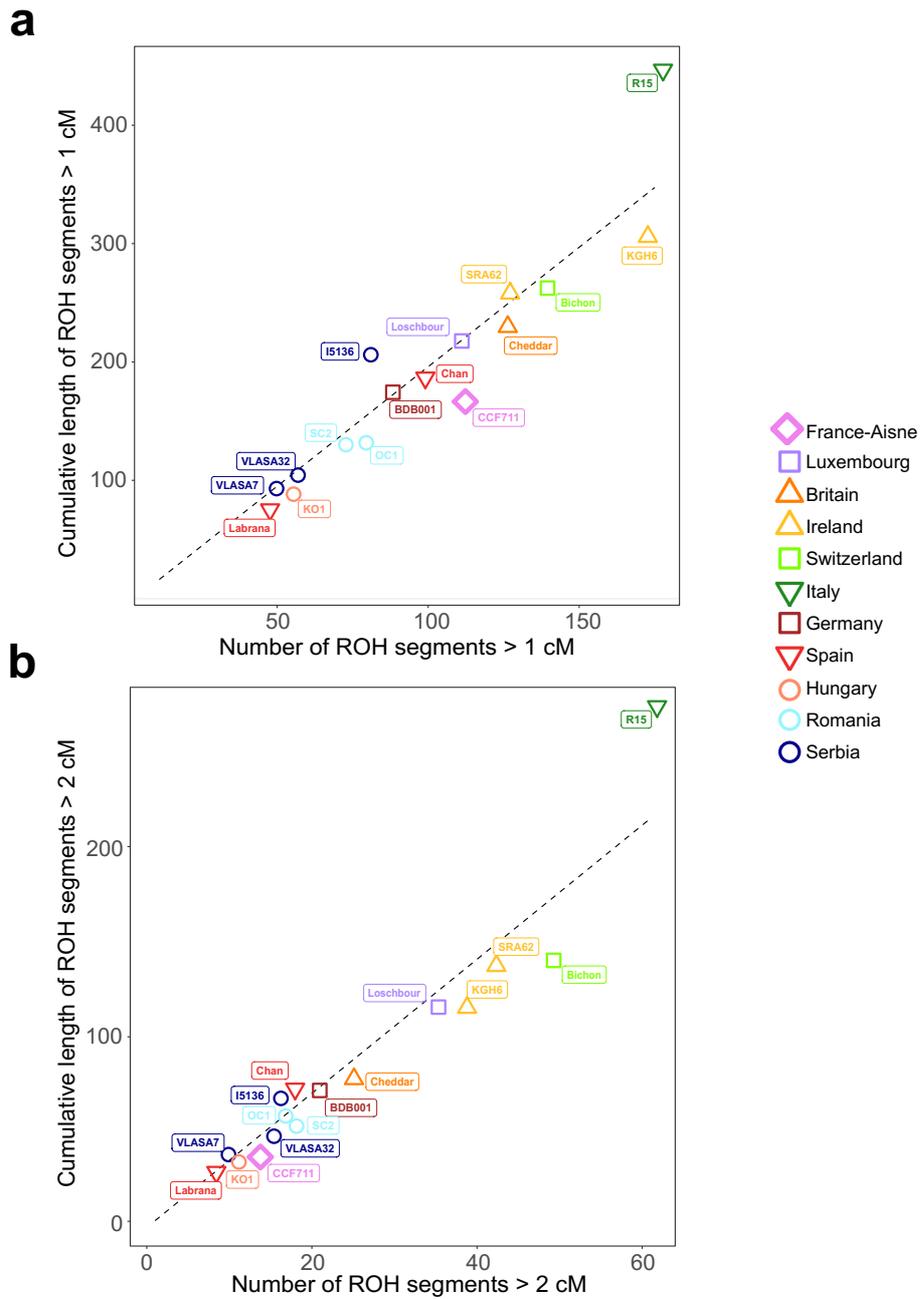


Figure S3. ROH computed for hunter-gatherer individuals. The ratio of cumulative length and number of ROH segments longer than 1 cM (a) and 2 cM (b).

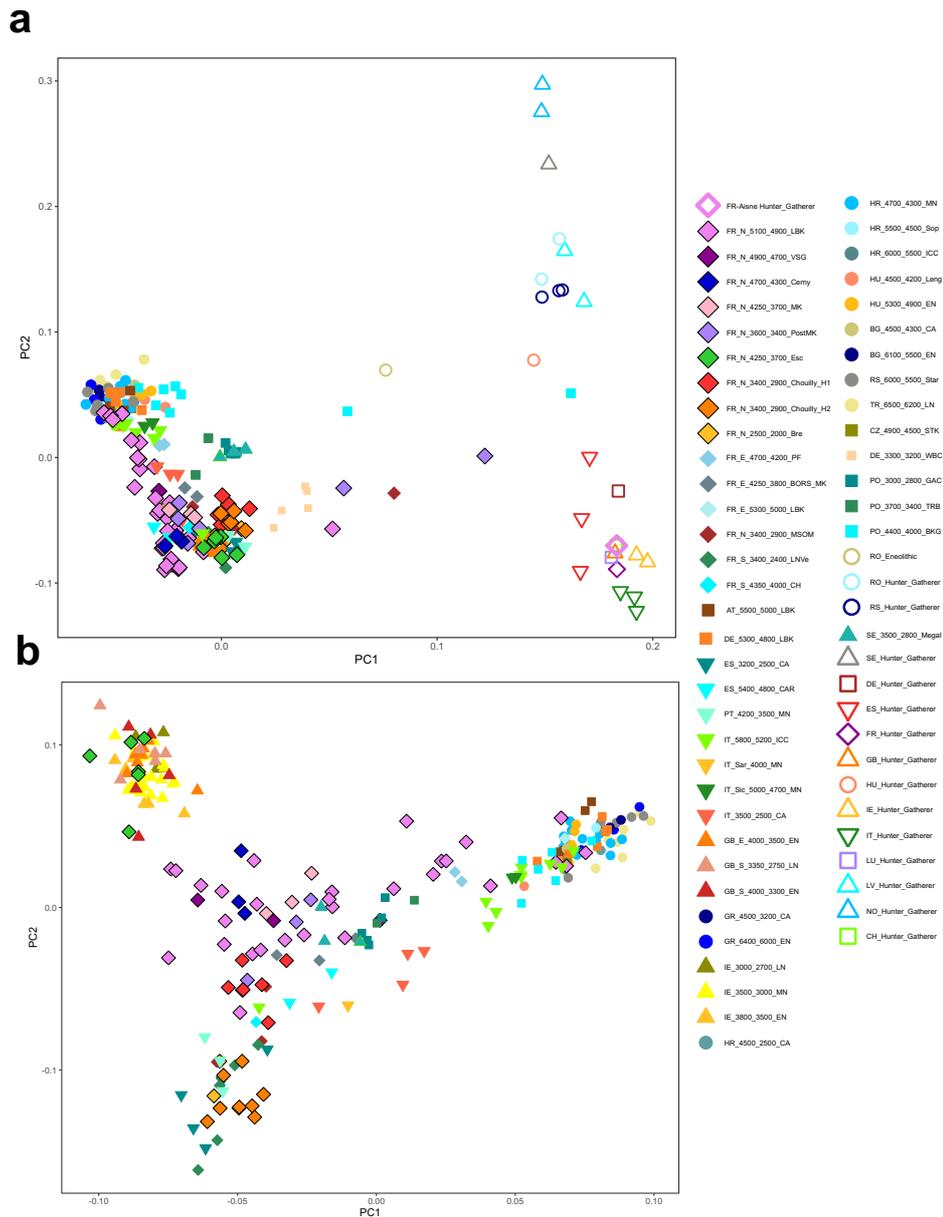


Figure S4. The impact of sample size change of the British Isles Neolithic populations on the fine-scale genetic affinities.

a, PCA based on ChromoPainter haplotypic coancestry matrix calculated on the dataset used in the analysis shown in Figure 2a without the later Neolithic individuals from the British Isles. The fraction of variance explained by PC1 and PC2 is, respectively, 70.4% and 6.6%. **b**, PCA based on ChromoPainter haplotypic coancestry matrix calculated on a subset of individuals excluding hunter-gatherer and outlier farmer genomes with high hunter-gatherer ancestry similar to that shown in Figure 2c but including the later Neolithic individuals from the British Isles. The fraction of variance explained by PC1 and PC2 is, respectively, 39.3% and 5.2%.

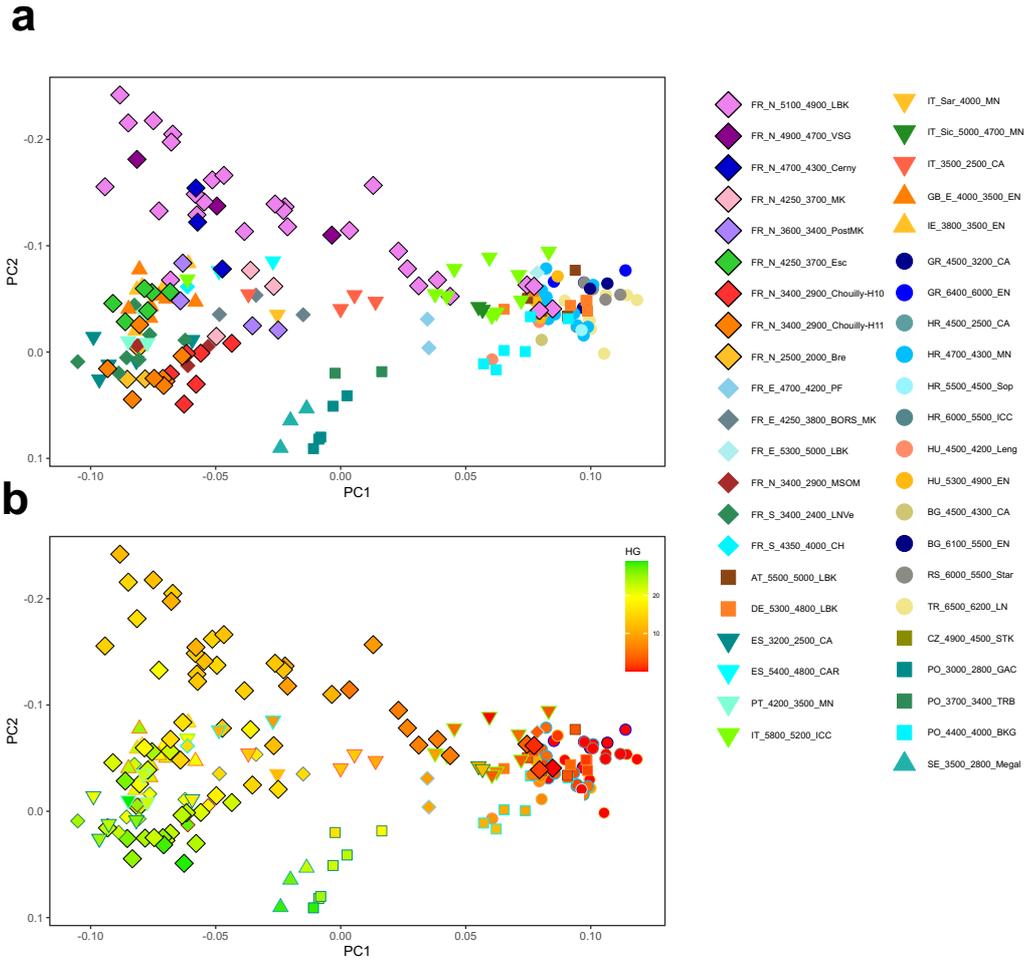


Figure S5. PCA based on ChromoPainter haplotypic coancestry matrix calculated on the same dataset as Figure 2D. The individuals are colored by group labels (a) or hunter-gatherer ancestry proportions (b). The fraction of variance explained by PC1 and PC2 is, respectively, 38.9% and 5.5%.

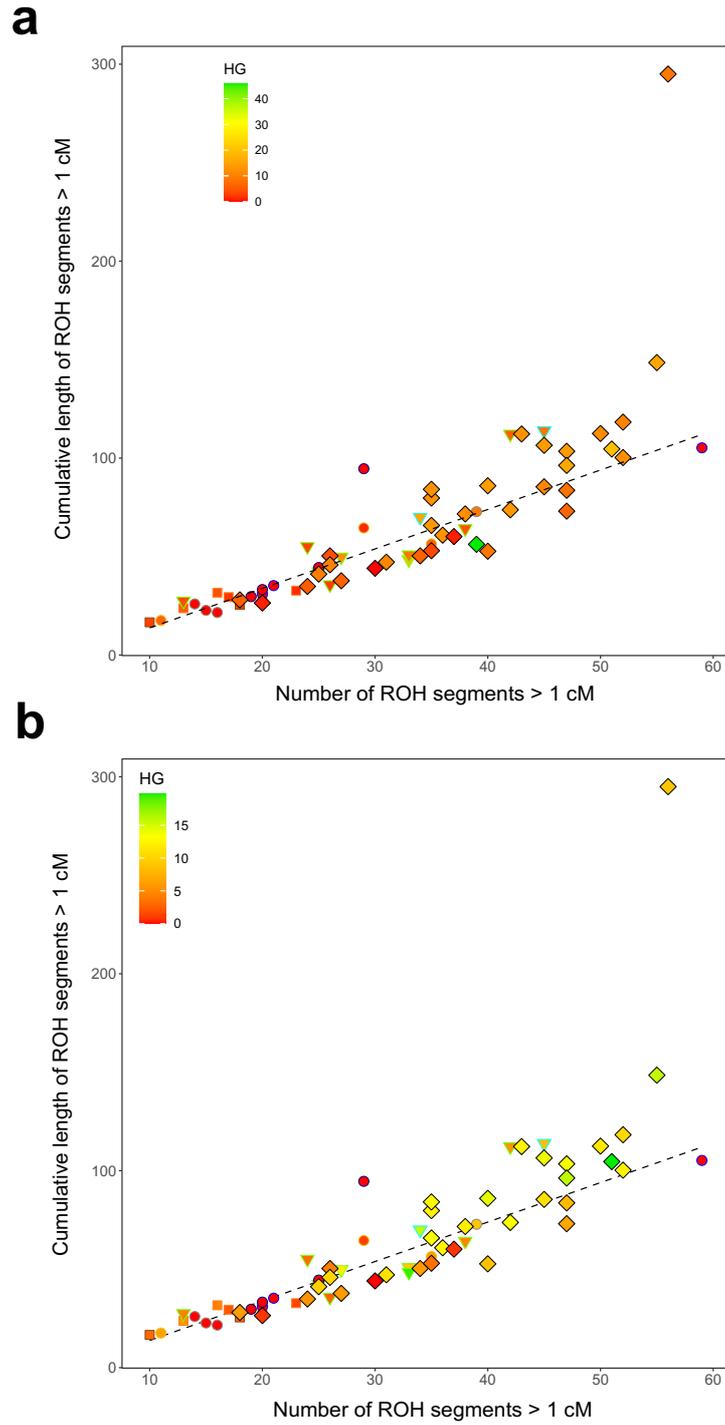


Figure S6. ROH computed for Early Neolithic individuals.

a, ROH plot as in Fig. 3a but showing individuals colored by their hunter-gatherer ancestry proportions estimated by SOURCEFIND. **b**, ROH plot as in Fig. S6a but excluding the individual with highest hunter-gatherer ancestry, SMP4738, to better visualize the ancestry proportions of other individuals.

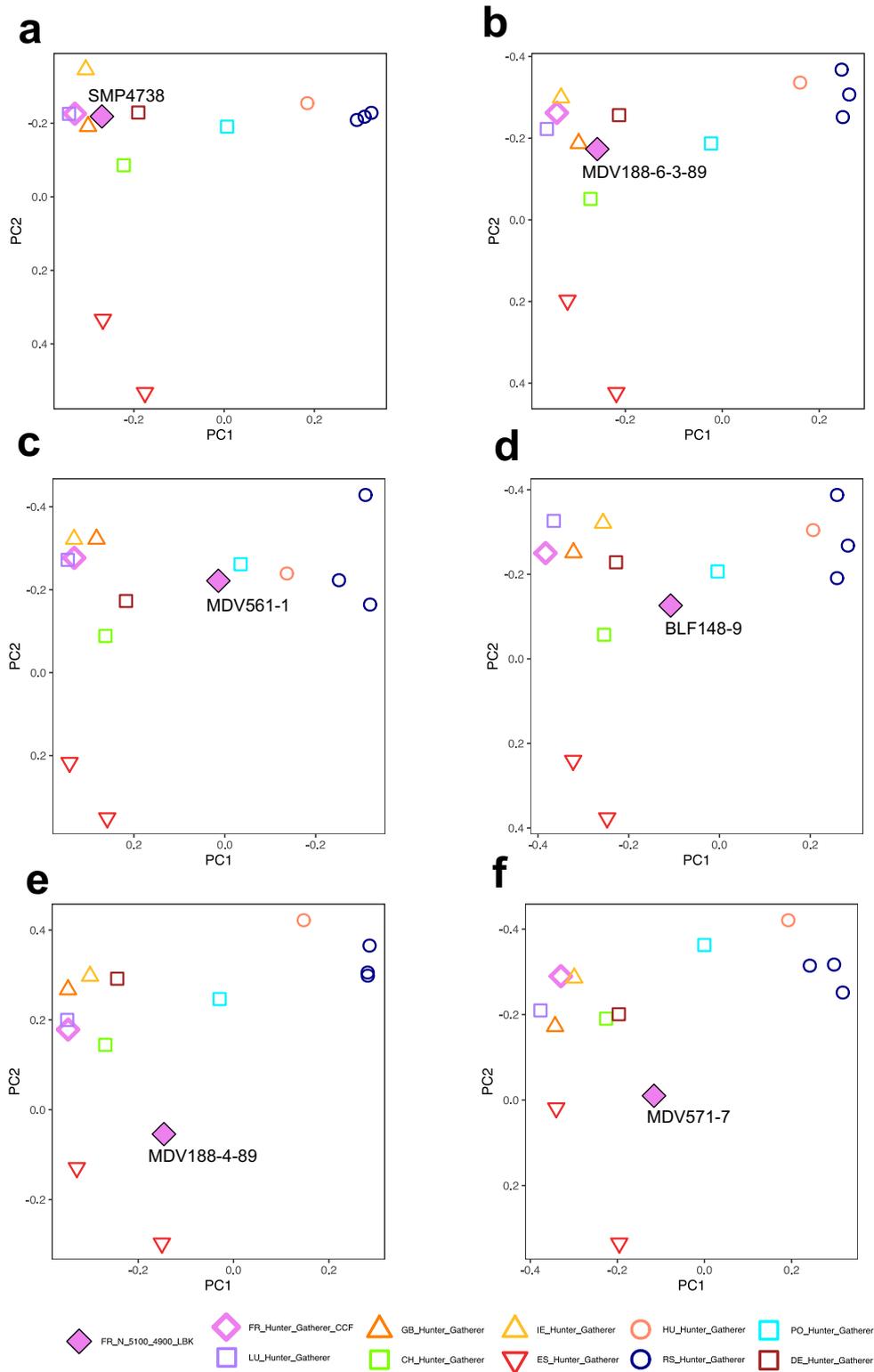


Figure S7. Diversity of hunter-gatherer ancestries in LBK individuals from the Paris Basin. PCA based on different ChromoPainter haplotypic coancestry matrices on hunter-gatherer genomes and on extracted hunter-gatherer segments from Neolithic individuals. Each matrix includes only one Neolithic individual due to the uniqueness of the genomic positions of hunter-gatherer segments found in each Neolithic genome. The fractions of variance explained by PC1 and PC2 are, respectively, 60-65% and 10-15%.

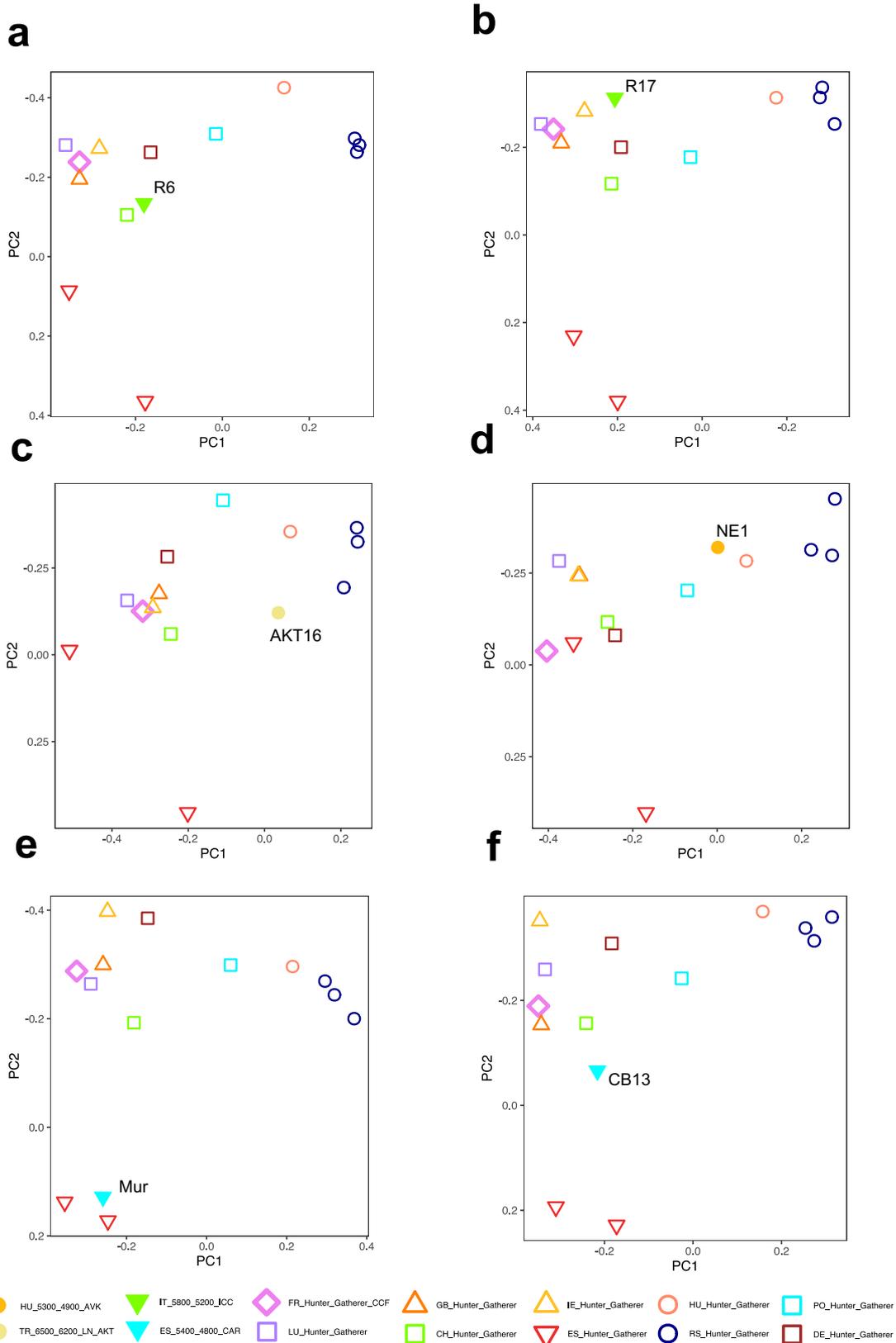


Figure S8. Diversity of hunter-gatherer ancestries in Early Neolithic individuals from different parts of Europe.

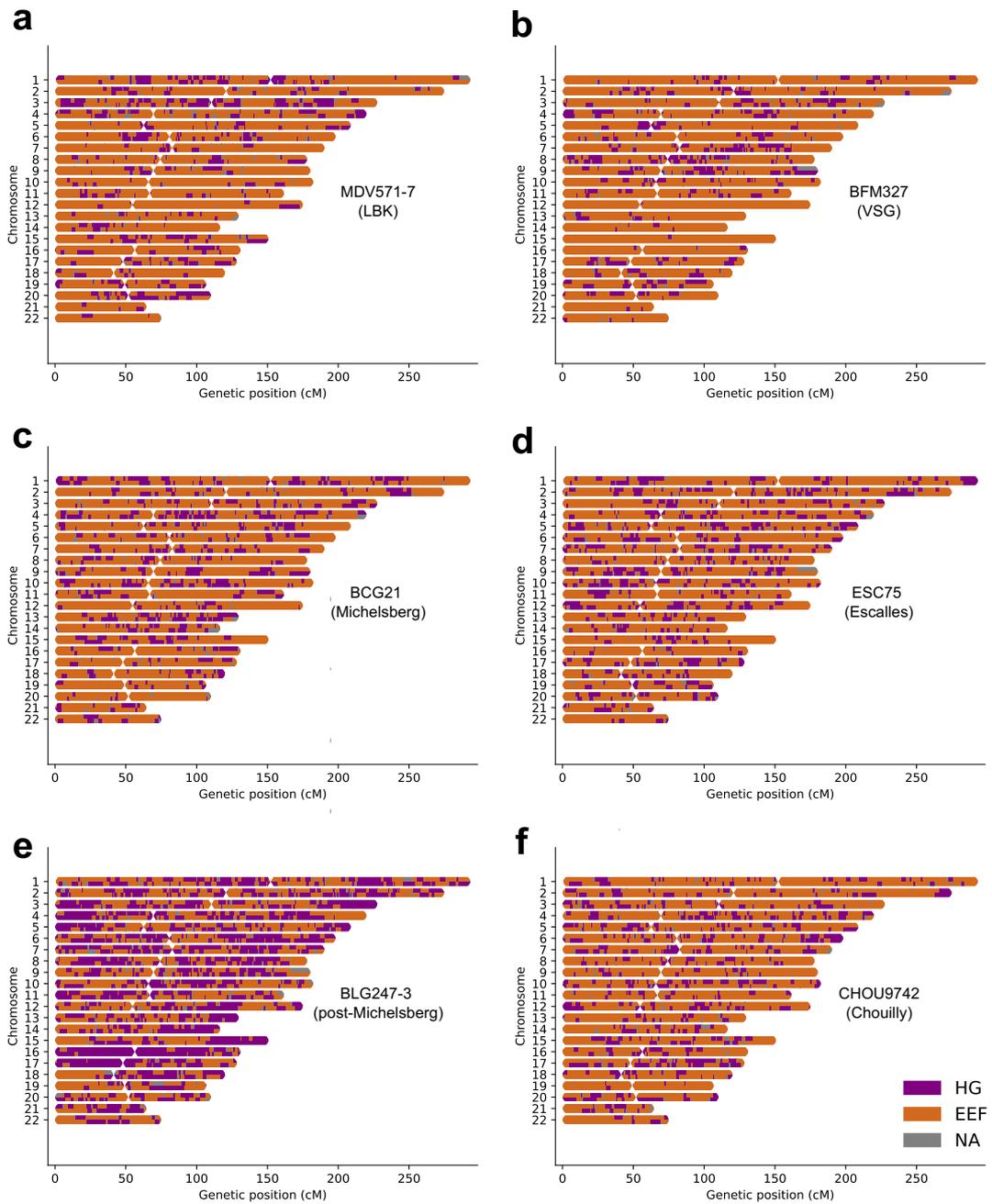


Figure S9. Genomic distribution of hunter-gatherer segments in Neolithic farmer genomes from the Paris Basin. Legend is as in Figure 5.

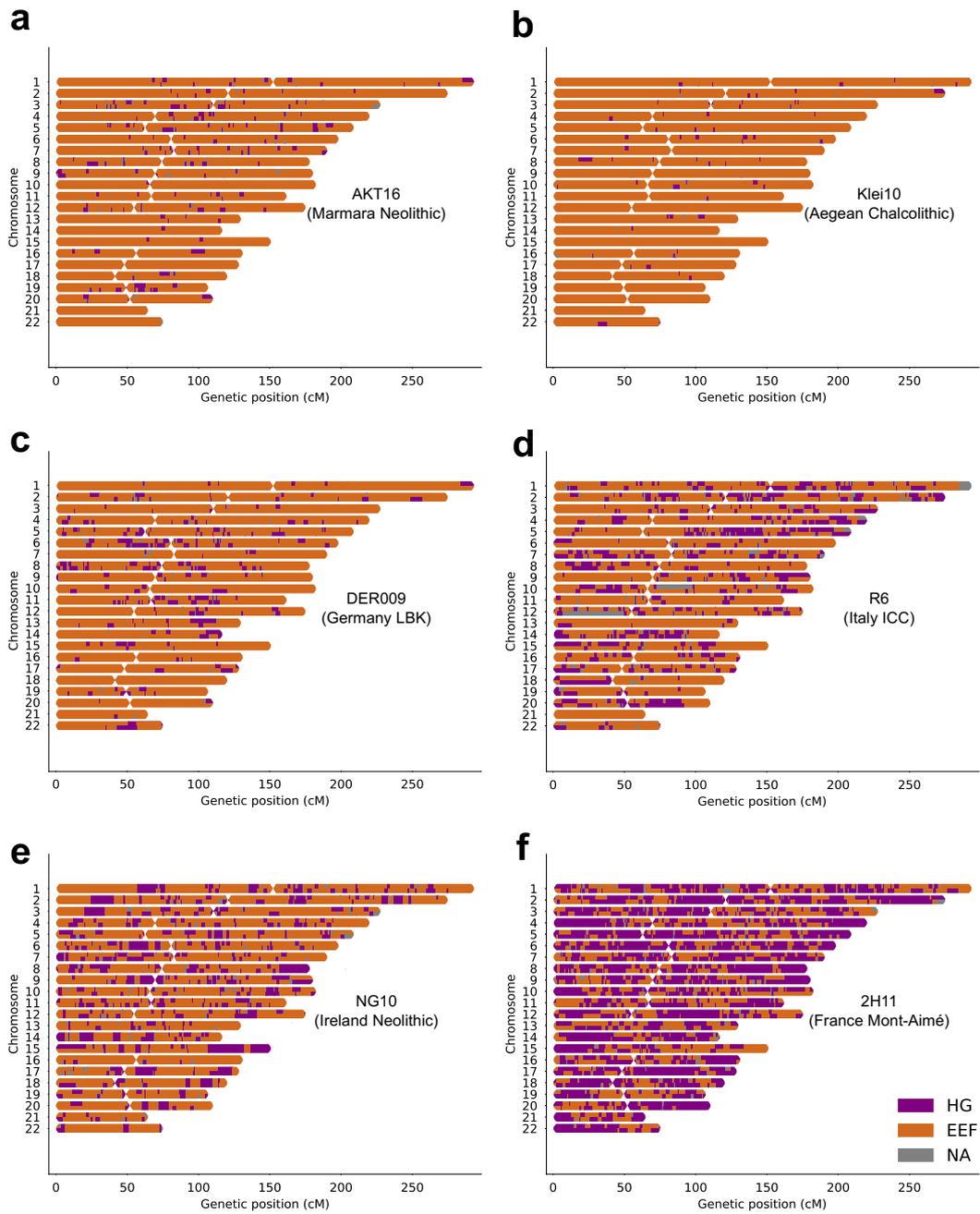


Figure S10. Genomic distribution of hunter-gatherer segments in previously published Neolithic farmer genomes from Europe. Legend is as in Figure 5.

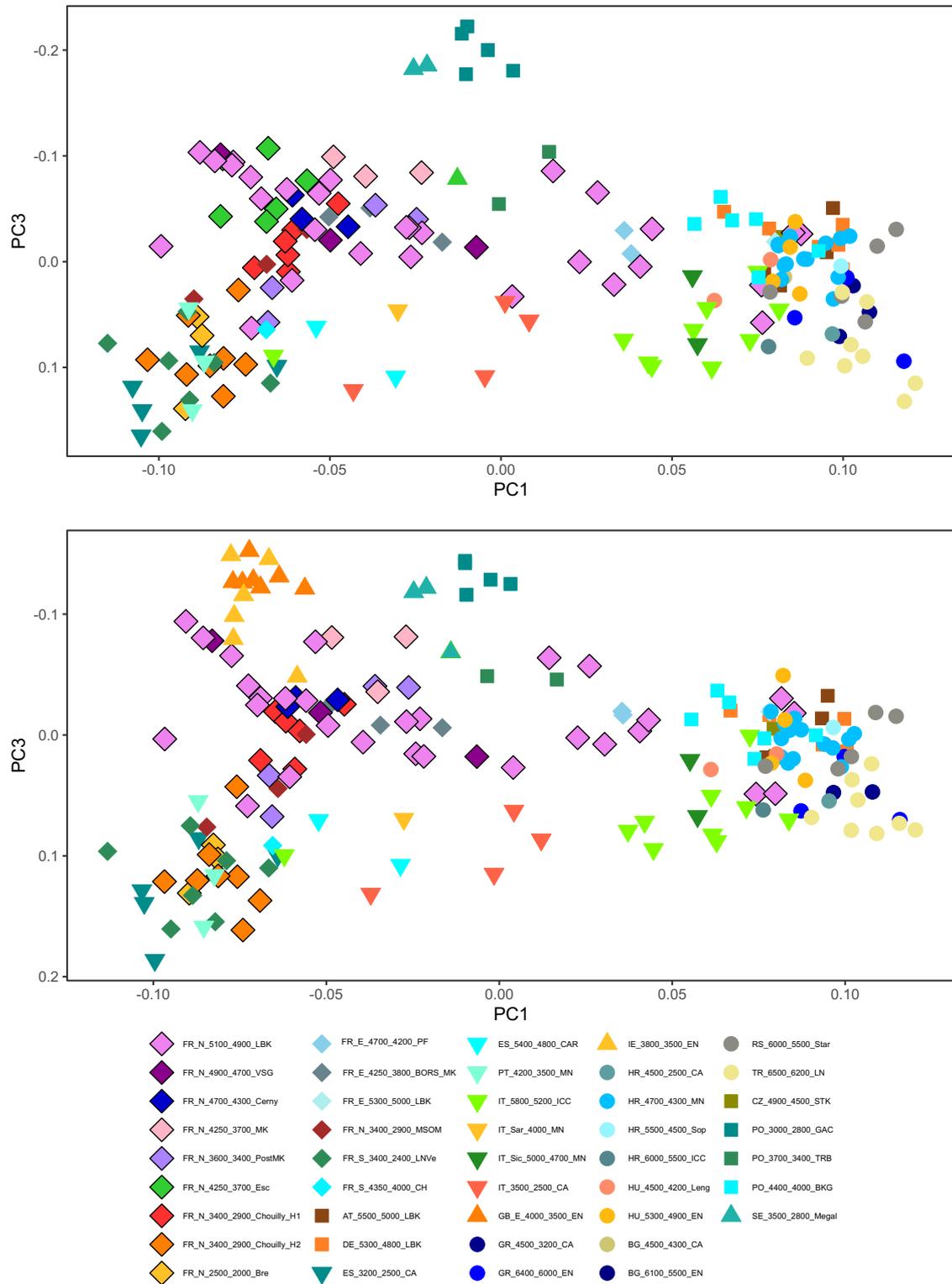


Figure S11. Origins of the British Isles/Escalles population in continental Europe.

PCA based on ChromoPainter haplotypic coancestry matrices calculated on the Neolithic dataset excluding either the British Isles (a) or the Escalles (b) genomes. See Figure 2d for comparison. The fractions of variance explained by PC1 and PC2 are, respectively, $\sim 40\%$ and $\sim 6\%$ in both PCA.

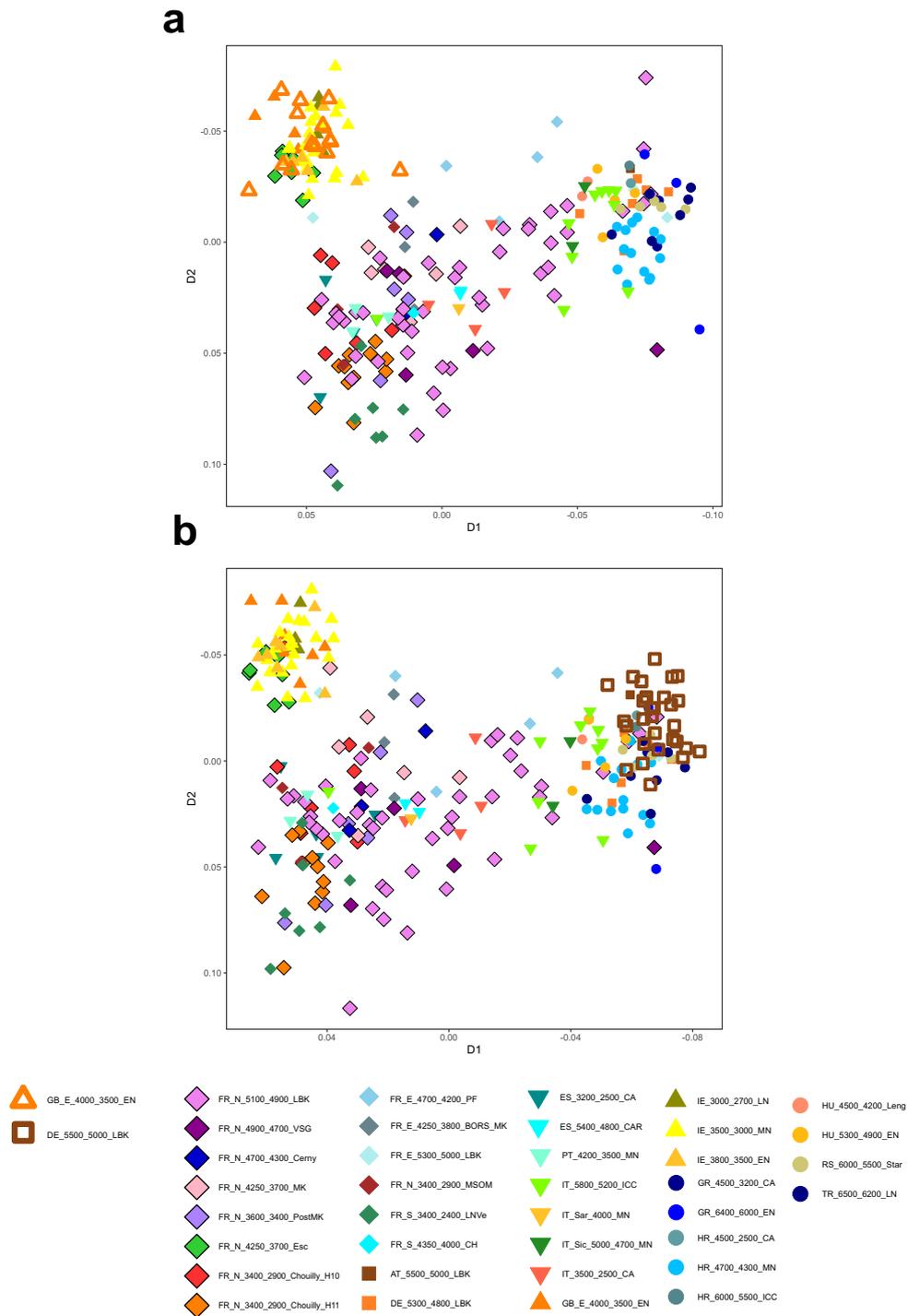


Figure S12. Multidimensional scaling (MDS) including both shotgun and capture genomes using a pseudo-haploidized 1240k SNP set. MDS calculated using the pairwise genetic distance matrix coming from inverted outgroup- f_3 values ($1-f_3$) in the form of $f_3(\text{Mbuti}; \text{Neolithic1}, \text{Neolithic2})$. Each distance matrix includes only one representative of the population whose genome segments were obtained using hybridization capture. See Figures S12-S13. Neolithic populations from England (a) and Germany (b) are clustering with contemporary Neolithic genomes obtained via shotgun sequencing.

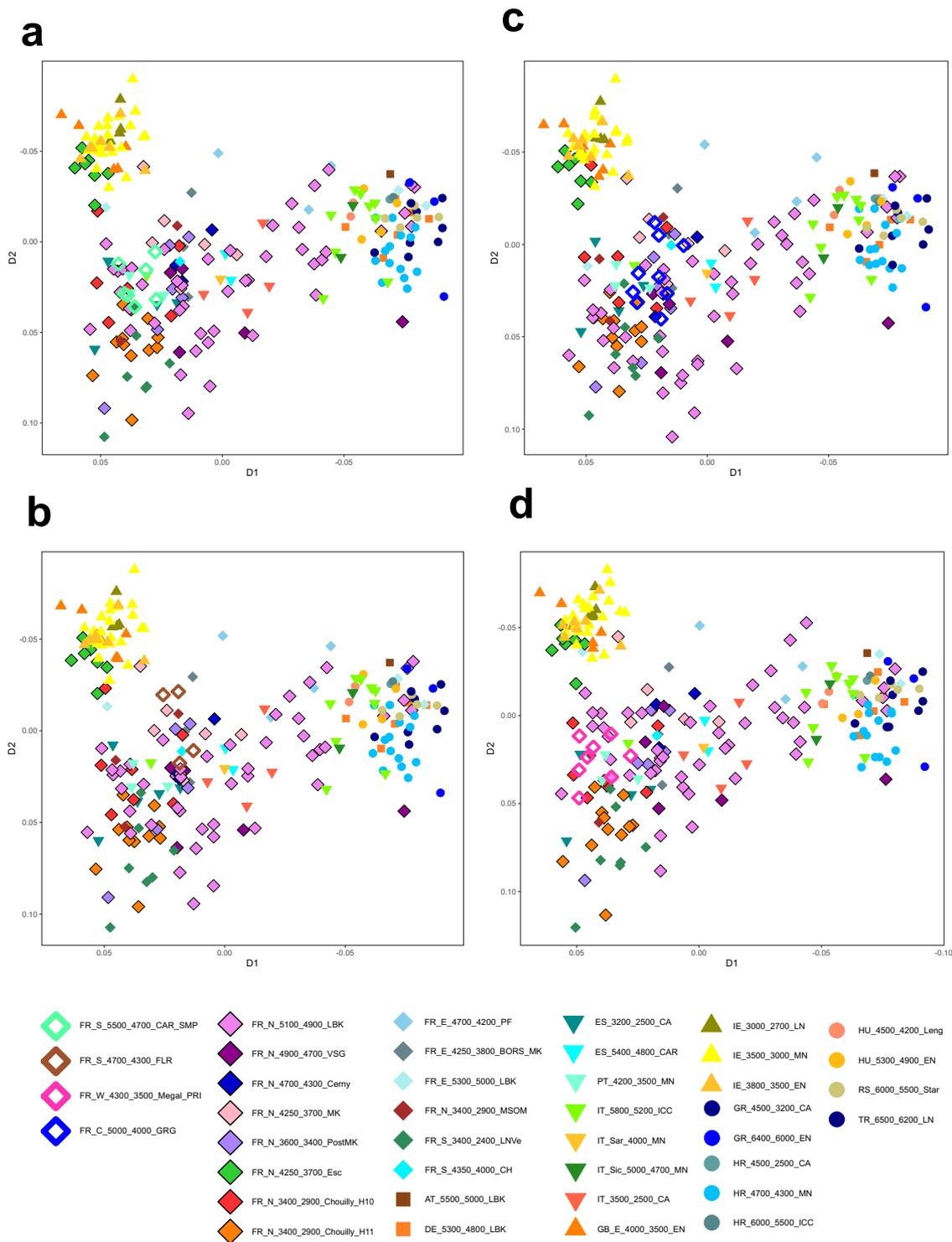


Figure S13. Multidimensional scaling (MDS) including both shotgun and capture genomes from France using a pseudo-haploidized 1240k SNP set.
Legend as in Fig. S12.

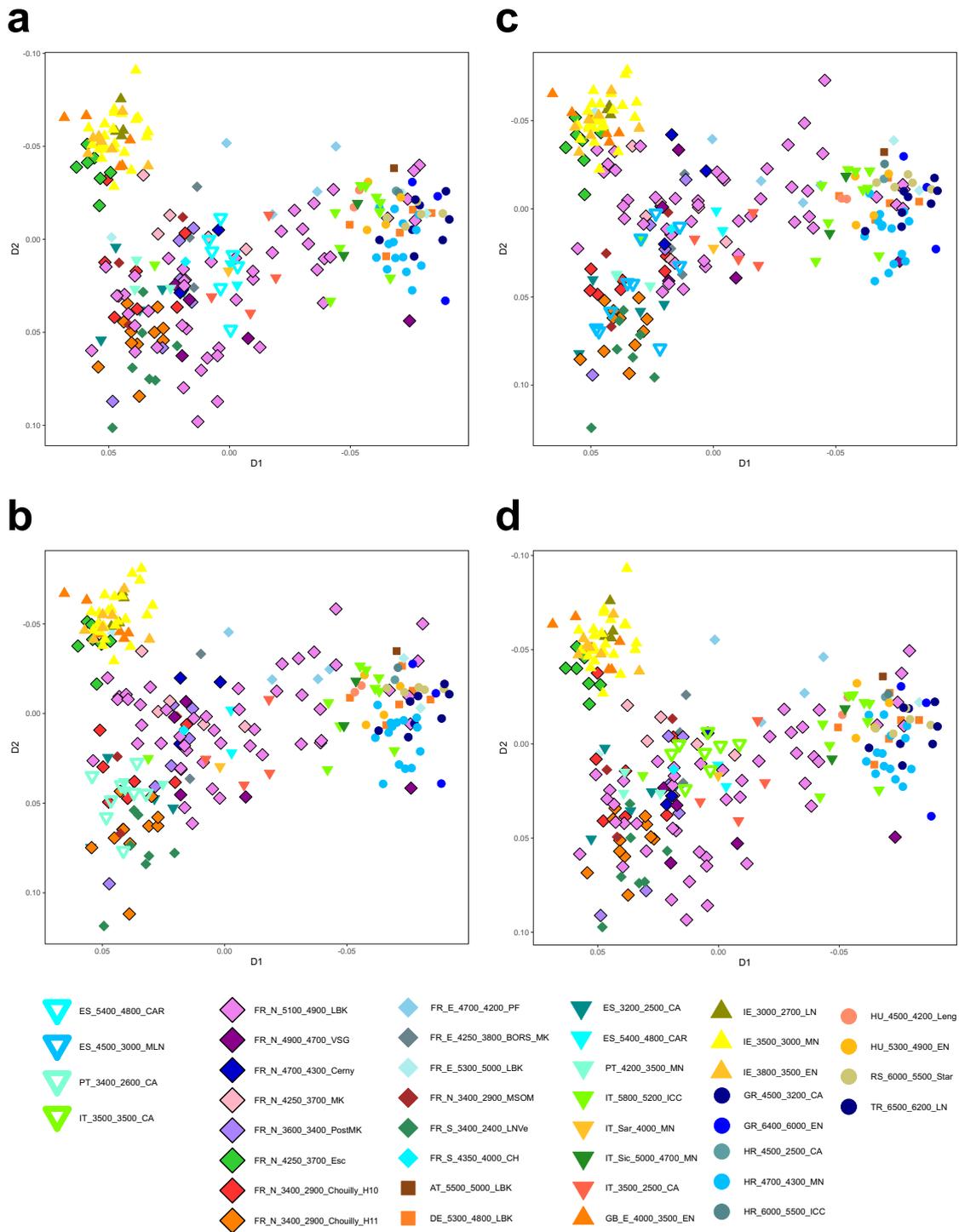


Figure S14. Multidimensional scaling (MDS) including both shotgun and capture genomes from Mediterranean Europe using a pseudo-haploidized 1240k SNP set. Legend as in Fig. S12.

Table S2. Estimation of kinship coefficients by NgsRelate for newly reported data

Ind1	Ind2	Autosomes					X chr	
		# Sites	θ	k2	k1	k0	# Sites	θ
BLF148	BLF149	665488	0,50	1	0	0	127139	0,50
BVT609	BVT610	665488	0,50	1	0	0	127139	0,50
MDV208	MDV233	665488	0,25	0,01	0,99	0,00	127139	0,26
MDV208	MDV213	665488	0,25	0,00	1,00	0,00	127139	0,50
BLF91	MGA46	665488	0,26	0,27	0,50	0,23	127139	0,38
MDV213	MDV233	665488	0,26	0,26	0,50	0,24	127139	0,03
MDV93.2	MDV93.3	665488	0,25	0,29	0,40	0,30	127139	0,32
MDV195	MDV208	665488	0,17	0,00	0,64	0,35	127139	0,00
BLG247-1	BLG247-2	1984995	0,12	0,00	0,47	0,52	137443	0,38
BLF70	MDV248	665488	0,14	0,00	0,54	0,46	127139	0,06
BLF53	BLF54	665488	0,15	0,01	0,59	0,40	127139	0,06
SMP451	SMP471	665488	0,15	0,03	0,54	0,43	127139	0,21
MDV195	MDV233	665488	0,08	0,00	0,31	0,68	127139	0,02
MDV195	MDV213	665488	0,08	0,00	0,29	0,69	127139	0,00
BLF53	BLF70	665488	0,10	0,00	0,42	0,58	127139	0,02
BLF54	BLF70	665488	0,06	0,01	0,23	0,76	127139	0,00
Esc637	Esc791	665488	0,07	0,02	0,24	0,75	127139	0,11
CHOU8047	CHOU3935	665488	0,05	0,02	0,16	0,81	127139	0,11
CHOU9741	CHOUF11P7B122	665488	0,04	0,00	0,15	0,85	127139	0,02
BLP10	BLP69	665488	0,04	0,00	0,15	0,85	127139	0,04

Table S4. qpAdm admixture modeling of each individual										
Group label (in this study)	Sample ID	Data type	Source	ANF proportion	HG proportion	ANF proportion (cc)	HG proportion (cc)	p-value	qpAdm SNP	
AT 5500 5000_LBK	Asp5	Shotgun	MarchCell2022	0.95	0.05	0.02	0.02	0.61	975825	
AT 5500 5000_LBK	IS069	Shotgun	MathiesonNature2018	0.96	0.04	0.02	0.02	0.12	942111	
AT 5500 5000_LBK	IS078	Shotgun	MathiesonNature2018	0.99	0.01	0.03	0.02	0.92	945006	
AT 5500 5000_LBK	IS204	Capture 1240x	MathiesonNature2018	0.96	0.04	0.02	0.02	0.59	773233	
AT 5500 5000_LBK	IS205	Capture 1240x	MathiesonNature2018	0.95	0.05	0.02	0.02	0.32	772051	
AT 5500 5000_LBK	IS206	Capture 1240x	MathiesonNature2018	0.96	0.04	0.02	0.02	0.76	754394	
AT 5500 5000_LBK	IS207	Capture 1240x	MathiesonNature2018	0.94	0.06	0.02	0.39	820741		
AT 5500 5000_LBK	IS208	Capture 1240x	MathiesonNature2018	0.95	0.05	0.02	0.02	0.02	816930	
CH 3500 2800_LNIH	Aesch10	Capture 1240x	FurtwanglerNatureCommunications2020	0.77	0.23	0.02	0.02	0.24	447948	
CH 3500 2800_LNIH	Aesch11	Capture 1240x	FurtwanglerNatureCommunications2020	0.81	0.19	0.02	0.02	0.48	526759	
CH 3500 2800_LNIH	Aesch12	Capture 1240x	FurtwanglerNatureCommunications2020	0.84	0.16	0.02	0.02	0.85	510057	
CH 3500 2800_LNIH	Aesch13	Capture 1240x	FurtwanglerNatureCommunications2020	0.81	0.19	0.02	0.02	0.23	563093	
CH 3500 2800_LNIH	Aesch14	Capture 1240x	FurtwanglerNatureCommunications2020	0.81	0.19	0.02	0.02	0.96	527723	
CH 3500 2800_LNIH	Aesch15	Capture 1240x	FurtwanglerNatureCommunications2020	0.79	0.21	0.02	0.02	0.20	437418	
CH 3500 2800_LNIH	Aesch17	Capture 1240x	FurtwanglerNatureCommunications2020	0.81	0.19	0.02	0.02	0.06	572080	
CH 3500 2800_LNIH	Aesch18	Capture 1240x	FurtwanglerNatureCommunications2020	0.80	0.20	0.02	0.02	0.91	484284	
CH 3500 2800_LNIH	Aesch19	Capture 1240x	FurtwanglerNatureCommunications2020	0.84	0.16	0.02	0.02	0.35	574764	
CH 3500 2800_LNIH	Aesch1	Capture 1240x	FurtwanglerNatureCommunications2020	0.81	0.19	0.02	0.02	0.12	517652	
CH 3500 2800_LNIH	Aesch20	Capture 1240x	FurtwanglerNatureCommunications2020	0.76	0.24	0.02	0.02	0.59	962115	
CH 3500 2800_LNIH	Aesch21	Capture 1240x	FurtwanglerNatureCommunications2020	0.81	0.19	0.02	0.02	0.02	903693	
CH 3500 2800_LNIH	Aesch22	Capture 1240x	FurtwanglerNatureCommunications2020	0.80	0.20	0.02	0.02	0.06	518949	
CH 3500 2800_LNIH	Aesch23	Capture 1240x	FurtwanglerNatureCommunications2020	0.82	0.18	0.02	0.02	0.85	600219	
CH 3500 2800_LNIH	Aesch24	Capture 1240x	FurtwanglerNatureCommunications2020	0.79	0.21	0.02	0.02	0.88	729867	
CH 3500 2800_LNIH	Aesch2	Capture 1240x	FurtwanglerNatureCommunications2020	0.83	0.17	0.02	0.02	0.11	430412	
CH 3500 2800_LNIH	Aesch3	Capture 1240x	FurtwanglerNatureCommunications2020	0.84	0.16	0.02	0.02	0.24	584443	
CH 3500 2800_LNIH	Aesch4	Capture 1240x	FurtwanglerNatureCommunications2020	0.81	0.19	0.02	0.02	0.14	550037	
CH 3500 2800_LNIH	Aesch5	Capture 1240x	FurtwanglerNatureCommunications2020	0.84	0.16	0.02	0.02	0.48	519940	
CH 3500 2800_LNIH	Aesch7	Capture 1240x	FurtwanglerNatureCommunications2020	0.81	0.19	0.02	0.02	0.32	465324	
CH 3500 2800_LNIH	Aesch8	Capture 1240x	FurtwanglerNatureCommunications2020	0.81	0.19	0.02	0.02	0.58	437773	
CH 3500 2800_LNIH	Aesch9	Capture 1240x	FurtwanglerNatureCommunications2020	0.78	0.22	0.02	0.02	0.32	454560	
CZ 3800 3500_TRB	II4167	Capture 1240x	PapaScienceAdvances2021	0.79	0.21	0.02	0.02	0.30	717481	
CZ 3800 3500_TRB	II4168	Capture 1240x	PapaScienceAdvances2021	0.81	0.19	0.02	0.02	0.59	698997	
CZ 3800 3500_TRB	II4169	Capture 1240x	PapaScienceAdvances2021	0.81	0.19	0.02	0.02	0.71	748711	
CZ 3800 3500_TRB	II4170	Capture 1240x	PapaScienceAdvances2021	0.85	0.15	0.02	0.02	0.34	740373	
CZ 3800 3500_TRB	II4171	Capture 1240x	PapaScienceAdvances2021	0.81	0.19	0.02	0.02	0.39	734843	
CZ 3800 3500_TRB	II4172	Capture 1240x	PapaScienceAdvances2021	0.80	0.20	0.02	0.02	0.84	745454	
CZ 3800 3500_TRB	II4173	Capture 1240x	PapaScienceAdvances2021	0.79	0.22	0.02	0.02	0.83	758495	
CZ 3800 3500_TRB	II4174	Capture 1240x	PapaScienceAdvances2021	0.82	0.18	0.02	0.02	0.12	731588	
CZ 3800 3500_TRB	II4175	Capture 1240x	PapaScienceAdvances2021	0.84	0.16	0.02	0.02	0.32	749211	
CZ 3800 3500_TRB	II4176	Capture 1240x	PapaScienceAdvances2021	0.81	0.19	0.02	0.02	0.10	725988	
CZ 3800 3500_TRB	II6121	Capture 1240x	PapaScienceAdvances2021	0.79	0.21	0.02	0.02	0.11	592747	
CZ 3800 3500_TRB	II6122	Capture 1240x	PapaScienceAdvances2021	0.80	0.20	0.02	0.02	0.68	582277	
CZ 3800 3500_TRB	II7185	Capture 1240x	PapaScienceAdvances2021	0.80	0.20	0.02	0.02	0.41	794762	
CZ 3800 3500_TRB	II7188	Capture 1240x	PapaScienceAdvances2021	0.80	0.20	0.02	0.02	0.80	796739	
CZ 3800 3500_TRB	II7191	Capture 1240x	PapaScienceAdvances2021	0.79	0.21	0.02	0.02	0.81	797328	
CZ 3800 3500_TRB	II7191	Capture 1240x	PapaScienceAdvances2021	0.81	0.19	0.02	0.02	0.91	801279	
CZ 3800 3500_TRB	II7192	Capture 1240x	PapaScienceAdvances2021	0.84	0.17	0.02	0.02	0.29	794540	
CZ 3800 3500_TRB	II7193	Capture 1240x	PapaScienceAdvances2021	0.56	0.44	0.02	0.02	0.23	778076	
CZ 3800 3500_TRB	II7194	Capture 1240x	PapaScienceAdvances2021	0.81	0.19	0.02	0.02	0.81	797328	
DE 4000 3000_MLN	Tangermünde (TG4000)	Capture 1240x	RivallatScienceAdvances2020	0.30	0.70	0.02	0.02	0.00	664332	
DE 4000 3000_MLN	II5560	Shotgun	MathiesonNature2015	0.79	0.21	0.02	0.02	0.97	453332	
DE 4000 3000_MLN	KHI50610	Shotgun	ImmelNatureCommunications2021	0.61	0.39	0.02	0.02	0.30	974249	
DE 4000 3000_MLN	KHI50615	Shotgun	ImmelNatureCommunications2021	0.60	0.41	0.02	0.02	0.03	959913	
DE 4000 3000_MLN	KHI50619	Shotgun	ImmelNatureCommunications2021	0.62	0.38	0.02	0.02	0.03	959913	
DE 4000 3000_MLN	KHI50620	Shotgun	ImmelNatureCommunications2021	0.66	0.34	0.02	0.02	0.21	975608	
DE 4000 3000_MLN	KHI50635	Shotgun	ImmelNatureCommunications2021	0.66	0.34	0.02	0.02	0.01	975561	
DE 4000 3000_MLN	II0172	Capture 1240x	MathiesonNature2015	0.83	0.17	0.02	0.02	0.25	944296	
DE 4000 3000_MLN	Blattenerhöhe (II565)	Capture 1240x	LipsonNature2017	0.60	0.40	0.02	0.02	0.02	945696	
DE 4000 3000_MLN	Blattenerhöhe (II565)	Capture 1240x	LipsonNature2017	0.62	0.38	0.02	0.02	0.81	945696	
DE 5500 5000_LBK	II916	Shotgun	MarchCell2022	0.98	0.02	0.02	0.02	0.02	975747	
DE 5500 5000_LBK	Is17	Shotgun	MarchCell2022	0.96	0.04	0.02	0.02	0.12	975778	
DE 5500 5000_LBK	II917	Shotgun	MarchCell2022	0.99	0.01	0.02	0.02	0.39	975747	
DE 5500 5000_LBK	DER002	Capture 1240x	ChildebevaMBe2022	0.99	0.01	0.02	0.02	0.55	882066	
DE 5500 5000_LBK	DER003	Capture 1240x	ChildebevaMBe2022	0.97	0.04	0.02	0.02	0.41	518930	
DE 5500 5000_LBK	DER004	Capture 1240x	ChildebevaMBe2022	0.96	0.04	0.02	0.02	0.81	518930	
DE 5500 5000_LBK	DER005	Capture 1240x	ChildebevaMBe2022	0.98	0.02	0.02	0.02	0.68	489041	
DE 5500 5000_LBK	DER006	Capture 1240x	ChildebevaMBe2022	1.00	0.00	0.02	0.02	0.46	502801	
DE 5500 5000_LBK	DER007	Capture 1240x	ChildebevaMBe2022	0.98	0.02	0.02	0.02	0.81	489041	
DE 5500 5000_LBK	DER009	Shotgun	ChildebevaMBe2022	0.93	0.07	0.02	0.02	0.61	883139	
DE 5500 5000_LBK	DER010	Capture 1240x	ChildebevaMBe2022	0.96	0.04	0.02	0.02	0.96	576749	
DE 5500 5000_LBK	DER011	Capture 1240x	ChildebevaMBe2022	0.98	0.02	0.02	0.02	0.77	502801	
DE 5500 5000_LBK	DER012	Capture 1240x	ChildebevaMBe2022	0.98	0.02	0.02	0.02	0.01	574313	
DE 5500 5000_LBK	DER013	Capture 1240x	ChildebevaMBe2022	1.00	0.00	0.02	0.02	0.85	674049	
DE 5500 5000_LBK	DER014	Capture 1240x	ChildebevaMBe2022	0.95	0.05	0.02	0.02	0.29	487180	
DE 5500 5000_LBK	DER015	Capture 1240x	ChildebevaMBe2022	0.99	0.02	0.02	0.02	0.59	672333	
DE 5500 5000_LBK	DER017	Capture 1240x	ChildebevaMBe2022	0.98	0.02	0.03	0.03	0.22	333127	
DE 5500 5000_LBK	DER018	Capture 1240x	ChildebevaMBe2022	0.97	0.03	0.02	0.02	0.47	479622	
DE 5500 5000_LBK	DER019	Capture 1240x	ChildebevaMBe2022	0.98	0.02	0.02	0.02	0.79	484541	
DE 5500 5000_LBK	DER020	Capture 1240x	ChildebevaMBe2022	0.96	0.04	0.03	0.03	0.75	272605	
DE 5500 5000_LBK	DER021	Capture 1240x	ChildebevaMBe2022	0.97	0.03	0.02	0.02	0.77	487440	
DE 5500 5000_LBK	DER022	Capture 1240x	ChildebevaMBe2022	0.98	0.02	0.02	0.02	0.91	455484	
DE 5500 5000_LBK	DER023	Capture 1240x	ChildebevaMBe2022	0.98	0.02	0.02	0.02	0.93	687534	
DE 5500 5000_LBK	DER024	Capture 1240x	ChildebevaMBe2022	0.96	0.04	0.02	0.02	0.34	581830	
DE 5500 5000_LBK	DER025	Capture 1240x	ChildebevaMBe2022	0.96	0.04	0.02	0.02	0.98	46096	
DE 5500 5000_LBK	DER028	Capture 1240x	ChildebevaMBe2022	0.94	0.06	0.02	0.02	0.50	676666	
DE 5500 5000_LBK	DER029	Capture 1240x	ChildebevaMBe2022	0.97	0.03	0.02	0.02	0.41	447606	
DE 5500 5000_LBK	DER030	Capture 1240x	ChildebevaMBe2022	0.97	0.03	0.02	0.02	0.04	482141	
DE 5500 5000_LBK	DER031	Capture 1240x	ChildebevaMBe2022	0.96	0.04	0.03	0.03	0.33	347610	
DE 5500 5000_LBK	DER032	Capture 1240x	ChildebevaMBe2022	0.95	0.05	0.02	0.02	0.62	630553	
DE 5500 5000_LBK	DER033	Capture 1240x	ChildebevaMBe2022	0.97	0.03	0.02	0.02	0.93	673762	
DE 5500 5000_LBK	DER034	Capture 1240x	ChildebevaMBe2022	0.96	0.04	0.02	0.02	0.81	489041	
DE 5500 5000_LBK	DER035	Capture 1240x	ChildebevaMBe2022	0.95	0.05	0.02	0.02	0.71	460539	
DE 5500 5000_LBK	DER036	Capture 1240x	ChildebevaMBe2022	0.92	0.08	0.02	0.02	0.84	663886	
DE 5500 5000_LBK	DER037	Capture 1240x	ChildebevaMBe2022	0.96	0.04	0.02	0.02	0.93	208863	
DE 5500 5000_LBK	Stuttgart	Shotgun	LazaridNature2014	0.95	0.05	0.02	0.02	0.21	975699	
DE 5500 5000_LBK	II0046	Capture 1240x	LipsonNature2017	0.96	0.04	0.02	0.02	0.10	820274	
DE 5500 5000_LBK	II0048	Capture 1240x	LipsonNature2017	0.98	0.02	0.02	0.02	0.81	821216	
DE 5500 5000_LBK	II0054	Capture 1240x	MathiesonNature2015	0.96	0.04	0.02	0.02	0.47	880139	
DE 5500 5000_LBK	II0056	Capture 1240x	LipsonNature2017	0.94	0.07	0.02	0.02	0.20	442434	
DE 5500 5000_LBK	II0057	Capture 1240x	LipsonNature2017	0.92	0.08	0.03	0.03	0.25	364413	
DE 5500 5000_LBK	II0100	Capture 1240x	LipsonNature2017	0.95	0.05	0.02	0.02	0.68	491706	
DE 5500 5000_LBK	II0659	Capture 1240x	LipsonNature2017	0.98	0.02	0.02	0.02	0.35	595312	
DE 5500 5000_LBK	II0821	Capture 1240x	LipsonNature2017	0.94	0.06	0.02	0.02	0.77	460495	
DE 5500 5000_LBK	II0826	Capture 1240x	LipsonNature2017	0.96	0.04	0.02	0.02	0.91	483370	
DE 5500 5000_LBK	II0025	Capture 1240x	MathiesonNature2015	0.97	0.03	0.02	0.02	0.71	794052	
DE 5500 5000_LBK	II0026	Capture 1240x	MathiesonNature2015	0.95	0.05	0.02	0.02	0.20	808622	
DE 5500 5000_LBK	II1165	Capture 1240x	RivallatScienceAdvances2020	0.32	0.68	0.02	0.02	0.03	292982	
ES 4300 3000_MLN	II0406	Capture 1240x	MathiesonNature2015							

ES 4300 3000 MLN	17646	Capture 12406	OldaleScience2019	0.81	0.19	0.03	0.03	0.52	294019
ES 4300 3000 MLN	17647	Capture 12406	OldaleScience2019	0.80	0.17	0.07	0.07	0.12	221061
ES 5400 4800 CAR	OH4002	Capture 12406	VillalbaMascocCurrentBiology2020	0.93	0.10	0.03	0.03	0.31	297978
ES 5400 4800 CAR	OH4003	Capture 12406	VillalbaMascocCurrentBiology2021	0.88	0.13	0.03	0.03	0.37	451819
ES 5400 4800 CAR	04009	Capture 12406	MathesonNature2015	0.90	0.10	0.02	0.02	0.53	457450
ES 5400 4800 CAR	04010	Capture 12406	MathesonNature2015	0.91	0.09	0.02	0.02	0.01	585391
ES 5400 4800 CAR	04011	Capture 12406	MathesonNature2015	0.94	0.06	0.03	0.03	0.54	875601
ES 5400 4800 CAR	04012	Capture 12406	MathesonNature2015	0.89	0.12	0.02	0.02	0.19	732453
ES 5400 4800 CAR	110942	Capture 12406	OldaleScience2019	0.84	0.16	0.02	0.02	0.06	446186
ES 5400 4800 CAR	110971	Capture 12406	LipsonNature2017	0.93	0.08	0.02	0.02	0.10	449314
ES 4300 3000 MLN	112248	Capture 12406	OldaleScience2019	0.72	0.28	0.02	0.02	0.08	598139
ES 4300 3000 MLN	17602	Capture 12406	OldaleScience2019	0.81	0.19	0.03	0.03	0.64	319100
ES 4300 3000 MLN	17603	Capture 12406	OldaleScience2019	0.76	0.20	0.02	0.02	0.04	613121
ES 4300 3000 MLN	17604	Capture 12406	OldaleScience2019	0.74	0.24	0.02	0.02	0.65	800766
ES 4300 3000 MLN	17605	Capture 12406	OldaleScience2019	0.72	0.29	0.02	0.02	0.16	688165
ES 4300 3000 MLN	17606	Capture 12406	OldaleScience2019	0.75	0.25	0.02	0.02	0.65	454955
ES 5400 4800 CAR	03111	Shogun	ValderramaPNA2018	0.89	0.11	0.02	0.02	0.31	598146
ES 5400 4800 CAR	Mur	Shogun	OldaleMBE2015	0.90	0.10	0.02	0.02	0.16	908161
FR E 4700 4200 PF	1046	Shogun	FragePNA2018	0.80	0.20	0.03	0.03	0.64	327740
FR E 4700 4200 PF	3X631	Capture 12406	FurtwanglerNatureCommunications2020	0.97	0.03	0.02	0.02	0.13	626245
FR E 4700 4200 PF	3X29	Capture 12406	FurtwanglerNatureCommunications2020	0.88	0.12	0.02	0.02	0.60	434261
FR E 4700 4200 PF	3X30	Capture 12406	FurtwanglerNatureCommunications2020	0.90	0.10	0.02	0.02	0.69	502139
FR S 5500 4600 CAR SMP	LR8001	Capture 12406	RivoltaScienceAdvances2020	0.80	0.20	0.02	0.02	0.10	626245
FR S 5500 4600 CAR SMP	LR8002	Capture 12406	RivoltaScienceAdvances2020	0.78	0.22	0.02	0.02	0.94	641370
FR S 5500 4600 CAR SMP	LR8003	Capture 12406	RivoltaScienceAdvances2020	0.76	0.25	0.02	0.02	0.71	651486
FR S 5500 4600 CAR SMP	LR8004	Capture 12406	RivoltaScienceAdvances2020	0.77	0.23	0.02	0.02	0.51	632886
FR S 5500 4600 CAR SMP	PEN0011	Capture 12406	RivoltaScienceAdvances2020	0.76	0.24	0.02	0.02	0.20	831924
FR S 5500 4600 CAR SMP	LR8005	Capture 12406	RivoltaScienceAdvances2020	0.65	0.35	0.02	0.02	0.29	721943
FR S 5500 4600 CAR SMP	PEN0002	Capture 12406	RivoltaScienceAdvances2020	0.76	0.24	0.02	0.02	0.51	657506
FR S 5500 4600 CAR SMP	PEN003	Capture 12406	RivoltaScienceAdvances2020	0.76	0.24	0.02	0.02	0.19	657506
FR E 4700 4200 PF	RO5102	Shogun	BruneiPNA2020	0.85	0.15	0.03	0.03	0.01	322757
FR E 4700 4200 PF	RO5045	Shogun	BruneiPNA2020	0.86	0.14	0.03	0.03	0.17	328345
FR E 4700 4200 PF	RO578	Shogun	BruneiPNA2020	0.87	0.13	0.02	0.02	0.36	475461
FR E 4700 4200 PF	OBN003	Capture 12406	RivoltaScienceAdvances2020	0.68	0.32	0.02	0.02	0.75	603363
FR E 4700 4200 PF	OBN007	Capture 12406	RivoltaScienceAdvances2020	0.89	0.11	0.02	0.02	0.63	465263
FR E 4700 4200 PF	OBN008	Capture 12406	RivoltaScienceAdvances2020	0.78	0.22	0.02	0.02	0.44	761822
FR E 4700 4200 PF	OBN009	Capture 12406	RivoltaScienceAdvances2020	0.89	0.11	0.02	0.02	0.27	616229
FR E 4700 4200 PF	OBN010	Capture 12406	RivoltaScienceAdvances2020	0.79	0.21	0.02	0.02	0.63	464210
FR E 4700 4200 PF	OBN011	Capture 12406	RivoltaScienceAdvances2020	0.89	0.11	0.02	0.02	0.77	798615
FR E 5300 5000 LNK	Morfe	Shogun	BruneiPNA2020	0.88	0.12	0.03	0.03	0.52	324991
FR E 5300 5000 LNK	SO712	Shogun	BruneiPNA2020	0.96	0.04	0.03	0.03	0.49	324991
FR N 3400 2900 LNK	Shew432	Shogun	BruneiPNA2020	0.90	0.10	0.02	0.02	0.62	419773
FR N 3400 2900 LNK	CHOU3923	Shogun	This study	0.74	0.27	0.02	0.02	0.65	452630
FR N 3400 2900 LNK	CHOU3925	Shogun	This study	0.69	0.31	0.02	0.02	0.03	623746
FR N 3400 2900 LNK	CHOU3980	Shogun	This study	0.75	0.25	0.02	0.02	0.91	333254
FR N 3400 2900 LNK	CHOU698	Shogun	This study	0.73	0.27	0.02	0.02	0.36	529971
FR N 3400 2900 LNK	CHOU7237	Shogun	This study	0.79	0.21	0.02	0.02	0.40	633528
FR N 3400 2900 LNK	CHOU8047	Shogun	This study	0.78	0.22	0.02	0.02	0.42	446528
FR N 3400 2900 LNK	CHOU863	Shogun	This study	0.77	0.23	0.02	0.02	0.65	425011
FR N 3400 2900 LNK	CHOU9373	Shogun	This study	0.80	0.20	0.02	0.02	0.87	369373
FR N 3400 2900 LNK	CHOU9740	Shogun	This study	0.78	0.22	0.02	0.02	0.37	425011
FR N 3400 2900 LNK	CHOU9741	Shogun	This study	0.76	0.24	0.03	0.03	0.60	354723
FR N 3400 2900 LNK	CHOU9742	Shogun	This study	0.79	0.21	0.02	0.02	0.05	649533
FR N 3400 2900 LNK	CHOU9743	Shogun	This study	0.77	0.23	0.02	0.02	0.07	649533
FR N 3400 2900 LNK	CHOU9744	Shogun	This study	0.76	0.24	0.03	0.03	0.60	429596
FR N 3400 2900 LNK	CHOU9745	Shogun	This study	0.71	0.29	0.02	0.02	0.73	430491
FR N 3400 2900 LNK	CHOU1185013	Shogun	This study	0.76	0.25	0.02	0.02	0.00	609091
FR N 3400 2900 LNK	CHOU1187812	Shogun	This study	0.74	0.26	0.02	0.02	0.05	658311
FR N 3400 2900 LNK	CHOU1187822	Shogun	This study	0.76	0.24	0.02	0.02	0.94	627139
FR N 3400 2900 LNK	CHOU1189822	Shogun	This study	0.73	0.27	0.02	0.02	0.37	678113
FR N 3400 2900 LNK	1H13	Shogun	SequimOrlandoPNA2021	0.73	0.27	0.02	0.02	0.72	715070
FR N 3400 2900 LNK	2H07	Shogun	SequimOrlandoPNA2021	0.79	0.22	0.02	0.02	0.88	476956
FR N 3400 2900 LNK	3H03	Shogun	SequimOrlandoPNA2021	0.83	0.17	0.02	0.02	0.37	765613
FR N 3400 2900 LNK	3H11	Shogun	SequimOrlandoPNA2021	0.41	0.59	0.02	0.02	0.17	976004
FR N 3600 3400 PostMK	BOC137	Shogun	This study	0.75	0.25	0.03	0.03	0.34	349228
FR N 3600 3400 PostMK	BLG2471	Shogun	This study	0.49	0.51	0.02	0.02	0.01	510704
FR N 3600 3400 PostMK	BLG2472	Shogun	This study	0.17	0.83	0.02	0.02	0.09	602606
FR N 3600 3400 PostMK	BLG2473	Shogun	This study	0.52	0.48	0.02	0.02	0.37	938776
FR N 3600 3400 PostMK	BLP88	Shogun	This study	0.83	0.17	0.02	0.02	0.91	544120
FR N 3600 3400 PostMK	CC1315	Shogun	This study	0.80	0.20	0.02	0.02	0.74	674211
FR N 3600 3400 PostMK	CC1312	Shogun	This study	0.82	0.18	0.02	0.02	0.34	323647
FR N 3600 3400 PostMK	18425	Capture 12406	PattersonNature2021	0.86	0.15	0.03	0.03	0.80	372264
FR N 3600 3400 PostMK	Manv9499	Shogun	This study	0.94	0.06	0.02	0.02	0.91	491905
FR N 3600 3400 PostMK	P551	Shogun	This study	0.83	0.17	0.02	0.02	0.25	395346
FR N 4250 3700 Esc	Esc10388	Shogun	This study	0.81	0.19	0.02	0.02	0.32	507251
FR N 4250 3700 Esc	Esc42	Shogun	This study	0.73	0.27	0.02	0.02	0.12	605940
FR N 4250 3700 Esc	Esc48	Shogun	This study	0.76	0.25	0.02	0.02	0.16	540006
FR N 4250 3700 Esc	Esc437	Shogun	This study	0.81	0.19	0.02	0.02	0.78	472045
FR N 4250 3700 Esc	Esc75	Shogun	This study	0.79	0.21	0.02	0.02	0.41	793933
FR N 4250 3700 Esc	Esc791	Shogun	This study	0.82	0.18	0.02	0.02	0.32	681674
FR N 4250 3700 Esc	Esc97	Shogun	This study	0.77	0.23	0.02	0.02	0.37	727776
FR N 4250 3700 MK	BLP21	Shogun	This study	0.82	0.18	0.02	0.02	0.35	618246
FR N 4250 3700 MK	BLP10	Shogun	This study	0.83	0.17	0.02	0.02	0.35	433959
FR N 4250 3700 MK	BLP27	Shogun	This study	0.86	0.14	0.03	0.03	0.12	336861
FR N 4250 3700 MK	BLP89	Shogun	This study	0.86	0.14	0.03	0.03	0.12	336861
FR N 4250 3700 MK	CF488	Shogun	This study	0.82	0.18	0.02	0.02	0.71	621886
FR N 4250 3700 MK	CT199	Shogun	This study	0.83	0.18	0.03	0.03	0.04	331229
FR N 4700 4300 Cerny	SH002	Shogun	BruneiPNA2020	0.86	0.15	0.02	0.02	0.92	399055
FR N 4700 4300 Cerny	BLP31	Shogun	This study	0.90	0.10	0.02	0.02	0.14	573829
FR N 4700 4300 Cerny	MAC1	Shogun	This study	0.89	0.11	0.02	0.02	0.60	397995
FR N 4700 4300 Fleury	F1R003	Capture 12406	RivoltaScienceAdvances2020	0.82	0.18	0.02	0.02	0.37	694210
FR N 4700 4300 Fleury	F1R004	Capture 12406	RivoltaScienceAdvances2020	0.85	0.15	0.02	0.02	0.08	694717
FR N 4700 4300 Fleury	F1R005	Capture 12406	RivoltaScienceAdvances2020	0.89	0.12	0.02	0.02	0.54	614087
FR N 4700 4300 Fleury	F1R007	Capture 12406	RivoltaScienceAdvances2020	0.86	0.14	0.02	0.02	0.76	698449
FR N 4900 4700 VSG	BFM357	Shogun	This study	0.86	0.14	0.02	0.02	0.47	446147
FR N 4900 4700 VSG	BFM369	Shogun	This study	0.88	0.12	0.02	0.02	0.29	562703
FR N 4900 4700 VSG	BFM370	Shogun	This study	0.93	0.07	0.03	0.03	0.92	334461
FR N 4900 4700 VSG	BFM373	Shogun	This study	0.86	0.14	0.02	0.02	0.31	674611
FR N 4900 4700 VSG	BFT181	Shogun	This study	0.89	0.11	0.02	0.02	0.72	354402
FR N 4800 4200 Gury	GRG003	Capture 12406	RivoltaScienceAdvances2020	0.88	0.12	0.02	0.02	0.61	695962
FR N 4800 4200 Gury	GRG008	Capture 12406	RivoltaScienceAdvances2020	0.85	0.15	0.02	0.02	0.47	449768
FR N 4800 4200 Gury	GRG015	Capture 12406	RivoltaScienceAdvances2020	0.84	0.16	0.02	0.02	0.54	726318
FR N 4800 4200 Gury	GRG016	Capture 12406	RivoltaScienceAdvances2020	0.87	0.14	0.02	0.02	0.38	768789
FR N 4800 4200 Gury	GRG019	Capture 12406	RivoltaScienceAdvances2020	0.87	0.14	0.02	0.02	0.41	749777
FR N 4800 4200 Gury	GRG021	Capture 12406	RivoltaScienceAdvances2020	0.87	0.13	0.02	0.02	0.63	674471
FR N 4800 4200 Gury	GRG022	Capture 12406	RivoltaScienceAdvances2020	0.83	0.17	0.02	0.02	0.06	658288
FR N 4800 4200 Gury	GRG025	Capture 12406	RivoltaScienceAdvances2020	0.86	0.14	0.02	0.02	0.62	644621
FR N 4800 4200 Gury	GRG027	Capture 12406	RivoltaScienceAdvances2020	0.85	0.15	0.02	0.02	0.79	706762
FR N 4800 4200 Gury	GRG028	Capture 12406	RivoltaScienceAdvances2020	0.89	0.11	0.02	0.02	0.22	614132
FR N 4800 4200 Gury	GRG032	Capture 12406	RivoltaScienceAdvances2020	0.87	0.13	0.02	0.02	0.18	695919
FR N 4800 4200 Gury	GRG041	Capture 12406	RivoltaScienceAdvances2020	0.86	0.14	0.02	0.02	0.89	577402
FR N 4800 4200 Gury	GRG043	Capture 12406	RivoltaScienceAdvances2020	0.91	0.10	0.02	0.02	0.68	655085
FR N 4800 4200 Gury	GRG063	Capture 12406	RivoltaScienceAdvances2						

FR_5_5500_4600_CAR_SMP	H305	Capture 1240k	OlaldeNature2018	0.78	0.22	0.02	0.02	0.37	735409
FR_W_4300_4000_MN	I15034	Capture 1240k	PattersonNature2021	0.74	0.26	0.02	0.02	0.99	739578
FR_W_4300_4000_MN	I15035	Capture 1240k	PattersonNature2021	0.77	0.23	0.02	0.02	0.18	771103
FR_W_4300_4000_MN	I15029	Capture 1240k	PattersonNature2021	0.73	0.27	0.02	0.02	0.21	754055
FR_W_4300_4000_MN	I15030	Capture 1240k	PattersonNature2021	0.81	0.19	0.02	0.02	0.35	708913
FR_W_4300_4000_MN	I15041	Capture 1240k	PattersonNature2021	0.75	0.25	0.02	0.02	0.03	750463
FR_W_4300_4000_MN	I15032	Capture 1240k	PattersonNature2021	0.75	0.25	0.02	0.02	0.29	753998
FR_W_4300_4000_MN	PR0001	Capture 1240k	RivollatScienceAdvances2020	0.81	0.19	0.02	0.02	0.42	596157
FR_W_4300_4000_MN	PR0006	Capture 1240k	RivollatScienceAdvances2020	0.79	0.21	0.02	0.02	0.67	453271
GB_E_4000_3500_EN	AveLine_1	Shotgun	BraceNatureEcologyEvolution2019	0.80	0.20	0.02	0.02	0.52	799555
GB_E_4000_3500_EN	Carlington_pasture_1	Shotgun	BraceNatureEcologyEvolution2019	0.79	0.19	0.02	0.02	0.08	974894
GB_E_4000_3500_EN	Jubilee_cave	Shotgun	BraceNatureEcologyEvolution2019	0.81	0.21	0.02	0.02	0.41	775649
GB_E_4000_3500_EN	Kelce_cave	Shotgun	BraceNatureEcologyEvolution2019	0.76	0.24	0.02	0.02	0.34	729195
GB_E_4000_3500_EN	NC7	Capture 1240k	FowlerNature2021	0.81	0.19	0.02	0.02	0.41	778398
GB_E_4000_3500_EN	NC37	Capture 1240k	FowlerNature2021	0.82	0.18	0.02	0.02	0.34	950682
GB_E_4000_3500_EN	NC5m	Capture 1240k	FowlerNature2021	0.80	0.20	0.02	0.02	0.30	719399
GB_E_4000_3500_EN	SC17	Capture 1240k	FowlerNature2021	0.80	0.20	0.02	0.02	0.44	680626
GB_E_4000_3500_EN	SC3m	Capture 1240k	FowlerNature2021	0.77	0.23	0.02	0.02	0.50	805228
GB_E_4000_3500_EN	SC3m	Capture 1240k	FowlerNature2021	0.77	0.23	0.02	0.02	0.02	941330
GB_E_4000_3500_EN	BumGrassland	Shotgun	BraceNatureEcologyEvolution2019	0.79	0.21	0.02	0.02	0.55	698718
GB_E_4000_3500_EN	Coldrum_1	Shotgun	BraceNatureEcologyEvolution2019	0.80	0.20	0.02	0.02	0.09	803340
GB_E_4000_3500_EN	Fussell_ledge_1	Shotgun	BraceNatureEcologyEvolution2019	0.78	0.22	0.02	0.02	0.40	817820
GB_E_4000_3500_EN	I16463	Capture 1240k	PattersonNature2021	0.85	0.15	0.02	0.02	0.13	528478
GB_E_4000_3500_EN	I16068	Capture 1240k	OlaldeNature2018	0.80	0.20	0.02	0.02	0.65	600721
GB_E_4000_3500_EN	I13135	Capture 1240k	OlaldeNature2018	0.85	0.15	0.02	0.02	0.16	733113
GB_E_4000_3500_EN	I1866	Capture 1240k	OlaldeNature2018	0.83	0.17	0.02	0.02	0.25	387404
GB_E_4000_3500_EN	I6624	Capture 1240k	OlaldeNature2018	0.80	0.20	0.02	0.02	0.00	629552
GB_E_4000_3500_EN	I6750	Capture 1240k	OlaldeNature2018	0.79	0.21	0.02	0.02	0.14	479840
GB_E_4000_3500_EN	I6751	Capture 1240k	OlaldeNature2018	0.77	0.23	0.02	0.02	0.50	899559
GB_E_4000_3500_EN	I6750	Capture 1240k	OlaldeNature2018	0.78	0.22	0.02	0.02	0.17	718478
GB_E_4000_3500_EN	I6761	Capture 1240k	OlaldeNature2018	0.78	0.22	0.02	0.02	0.52	567339
GB_E_4000_3500_EN	I6762	Capture 1240k	OlaldeNature2018	0.79	0.21	0.02	0.02	0.10	604211
GB_E_4000_3500_EN	I6035	Capture 1240k	OlaldeNature2018	0.80	0.20	0.02	0.02	0.60	652043
GB_S_3350_2750_MN	I2932	Shotgun	OlaldeNature2018	0.78	0.22	0.02	0.02	0.63	524653
GB_S_3350_2750_MN	I2934	Shotgun	OlaldeNature2018	0.82	0.18	0.02	0.02	0.44	898626
GB_S_3350_2750_MN	I2980	Shotgun	OlaldeNature2018	0.83	0.17	0.02	0.02	0.20	975272
GB_S_3350_2750_MN	I1118	Shotgun	OlaldeNature2018	0.78	0.22	0.02	0.02	0.82	612026
GB_S_3350_2750_MN	I7554	Shotgun	OlaldeNature2018	0.81	0.19	0.02	0.02	0.09	945021
GB_S_3350_2750_MN	I2631	Capture 1240k	OlaldeNature2018	0.81	0.19	0.02	0.02	0.79	781020
GB_S_3350_2750_MN	I1551	Capture 1240k	OlaldeNature2018	0.80	0.20	0.02	0.02	0.00	714509
GB_S_3350_2750_MN	I9313	Capture 1240k	OlaldeNature2018	0.76	0.24	0.02	0.02	0.00	693040
GB_S_3350_2750_MN	I9315	Capture 1240k	OlaldeNature2018	0.82	0.18	0.02	0.02	0.18	749370
GB_S_3350_2750_MN	I2977	Capture 1240k	OlaldeNature2018	0.83	0.17	0.02	0.02	0.17	728969
GB_S_3350_2750_MN	I2978	Capture 1240k	OlaldeNature2018	0.80	0.20	0.02	0.02	0.07	738993
GB_S_3350_2750_MN	I2979	Capture 1240k	OlaldeNature2018	0.82	0.18	0.02	0.02	0.96	737139
GB_S_3350_2750_MN	I3085	Capture 1240k	OlaldeNature2018	0.80	0.20	0.02	0.02	0.72	698818
GB_S_4000_3300_EN	I2431	Capture 1240k	OlaldeNature2018	0.79	0.21	0.02	0.02	0.66	695644
GB_S_4000_3300_EN	I2635	Capture 1240k	OlaldeNature2018	0.85	0.15	0.02	0.02	0.11	712619
GB_S_4000_3300_EN	I2636	Shotgun	OlaldeNature2018	0.79	0.21	0.02	0.02	0.72	938161
GB_S_4000_3300_EN	I16317	Shotgun	OlaldeNature2018	0.82	0.18	0.02	0.02	0.82	682793
GB_S_4000_3300_EN	I0550	Capture 1240k	OlaldeNature2018	0.83	0.17	0.02	0.02	0.51	720160
GB_S_4000_3300_EN	I1660	Capture 1240k	OlaldeNature2018	0.77	0.23	0.02	0.02	0.23	754426
GB_S_4000_3300_EN	I6691	Capture 1240k	OlaldeNature2018	0.72	0.28	0.02	0.02	0.12	727446
GB_S_4000_3300_EN	I2884	Capture 1240k	OlaldeNature2018	0.84	0.16	0.02	0.02	0.63	746133
GB_S_4000_3300_EN	I1313	Shotgun	OlaldeNature2018	0.80	0.20	0.02	0.02	0.17	975293
GB_S_4000_3300_EN	I1314	Shotgun	OlaldeNature2018	0.77	0.23	0.02	0.02	0.70	938919
GB_S_4000_3300_EN	I1316	Shotgun	OlaldeNature2018	0.81	0.19	0.02	0.02	0.17	840515
HR_4700_4300_MN	POP06	Shotgun	FreilichScientificReports2021	0.94	0.07	0.02	0.02	0.52	720625
HR_4700_4300_MN	POP08	Shotgun	FreilichScientificReports2021	0.91	0.09	0.02	0.02	0.87	607342
HR_4700_4300_MN	POP09	Shotgun	FreilichScientificReports2021	0.93	0.08	0.02	0.02	0.91	660398
HR_4700_4300_MN	POP11	Shotgun	FreilichScientificReports2021	0.95	0.05	0.02	0.02	0.78	615094
HR_4700_4300_MN	POP12	Shotgun	FreilichScientificReports2021	0.98	0.02	0.02	0.02	0.19	648198
HR_4700_4300_MN	POP13	Shotgun	FreilichScientificReports2021	0.92	0.08	0.02	0.02	0.00	802537
HR_4700_4300_MN	POP14	Shotgun	FreilichScientificReports2021	0.93	0.07	0.02	0.02	0.60	629017
HR_4700_4300_MN	POP16	Shotgun	FreilichScientificReports2021	1.00	0.00	0.02	0.02	0.15	621379
HR_4700_4300_MN	POP19	Shotgun	FreilichScientificReports2021	0.98	0.02	0.02	0.02	0.53	656979
HR_4700_4300_MN	POP27	Shotgun	FreilichScientificReports2021	1.00	0.00	0.02	0.02	0.00	635023
HR_4700_4300_MN	POP30	Shotgun	FreilichScientificReports2021	0.99	0.02	0.02	0.02	0.50	608066
HR_4700_4300_MN	POP33	Shotgun	FreilichScientificReports2021	0.98	0.03	0.02	0.02	0.64	560471
HR_4700_4300_MN	POP36	Shotgun	FreilichScientificReports2021	1.00	0.00	0.02	0.02	0.17	955125
HR_4700_4300_MN	POP36	Shotgun	FreilichScientificReports2021	0.97	0.04	0.02	0.02	0.15	999501
HR_6000_5500_ICC	I1433	Shotgun	MathiesonNature2018	0.98	0.03	0.02	0.02	0.16	830489
HR_6000_5500_ICC	I1496	Capture 1240k	MathiesonNature2018	1.00	0.00	0.02	0.02	0.91	538789
HR_6000_5500_ICC	I1497	Capture 1240k	MathiesonNature2018	1.00	0.00	0.02	0.02	0.61	716550
HR_6000_5500_ICC	I1418	Capture 1240k	MathiesonNature2018	0.95	0.05	0.03	0.03	0.97	333672
HU_3600_3000_CA	I2366	Capture 1240k	LipsonNature2017	0.88	0.12	0.02	0.02	0.42	739717
HU_3600_3000_CA	I2367	Capture 1240k	LipsonNature2017	0.92	0.08	0.02	0.02	0.64	764209
HU_3600_3000_CA	I2369	Capture 1240k	LipsonNature2017	0.92	0.08	0.02	0.02	0.01	712893
HU_3600_3000_CA	I2753	Capture 1240k	LipsonNature2017	0.86	0.14	0.02	0.02	0.02	784632
HU_3600_3000_CA	I2788	Capture 1240k	LipsonNature2017	0.86	0.14	0.02	0.02	0.17	751121
HU_3600_3000_CA	I2790	Capture 1240k	LipsonNature2017	0.91	0.09	0.02	0.02	0.46	850000
HU_3600_3000_CA	I2791	Capture 1240k	LipsonNature2017	0.87	0.13	0.02	0.02	0.41	760850
HU_5300_4900_LBKT_AVK	I1498	Shotgun	GambaNatureCommunications2014	0.91	0.09	0.02	0.02	0.74	899401
HU_5300_4900_LBKT_AVK	I1499	Capture 1240k	MathiesonNature2015	0.91	0.09	0.02	0.02	0.17	735556
HU_5300_4900_LBKT_AVK	I1505	Capture 1240k	MathiesonNature2015	0.89	0.11	0.02	0.02	0.11	735541
HU_5300_4900_LBKT_AVK	I2381	Capture 1240k	MathiesonNature2015	0.90	0.10	0.02	0.02	0.36	575790
HU_5300_4900_LBKT_AVK	I2382	Capture 1240k	LipsonNature2017	0.92	0.08	0.02	0.02	0.68	668254
HU_5300_4900_LBKT_AVK	NE1	Shotgun	GambaNatureCommunications2014	0.92	0.08	0.02	0.02	0.27	975608
HU_5300_4900_LBKT_AVK	NE5	Shotgun	GambaNatureCommunications2014	0.92	0.09	0.02	0.02	0.82	542862
HU_5300_4900_LBKT_AVK	I1496	Capture 1240k	MathiesonNature2015	0.96	0.04	0.02	0.02	0.55	741566
HU_5300_4900_LBKT_AVK	I1494	Capture 1240k	LipsonNature2017	1.00	0.00	0.02	0.02	0.50	744674
HU_5300_4900_LBKT_AVK	I2739	Capture 1240k	LipsonNature2017	0.97	0.04	0.02	0.02	0.36	707791
HU_5300_4900_LBKT_AVK	I1496	Capture 1240k	LipsonNature2017	0.98	0.02	0.02	0.02	0.91	632046
HU_5300_4900_LBKT_AVK	NE6	Shotgun	GambaNatureCommunications2014	0.97	0.03	0.02	0.02	0.67	626665
HU_5300_4900_LBKT_AVK	I2377	Capture 1240k	LipsonNature2017	0.87	0.14	0.02	0.02	0.25	638069
HU_5300_4900_LBKT_AVK	I2379	Capture 1240k	LipsonNature2017	0.91	0.10	0.02	0.02	0.87	620488
HU_5800_5500_Star_Koros	I1588	Capture 1240k	MathiesonNature2015	0.93	0.07	0.02	0.02	0.17	560775
HU_5800_5500_Star_Koros	I2794	Capture 1240k	LipsonNature2017	0.99	0.01	0.02	0.02	0.24	765371
HU_5800_5500_Star_Koros	I1877	Capture 1240k	LipsonNature2017	1.00	0.00	0.02	0.02	0.75	566129
HU_5800_5500_Star_Koros	I1878	Capture 1240k	LipsonNature2017	0.96	0.04	0.02	0.02	0.85	681445
HU_5800_5500_Star_Koros	I1880	Capture 1240k	LipsonNature2017	1.00	0.00	0.02	0.02	0.36	524617
IE_3500_3000_MM	ANN1	Shotgun	CassidyScience2020	0.80	0.20	0.02	0.02	0.36	764652
IE_3500_3000_MM	ANN2	Shotgun	CassidyScience2020	0.77	0.23	0.02	0.02	1.00	469646
IE_3500_3000_MM	AS11	Shotgun	CassidyScience2020	0.85	0.15	0.02	0.02	0.12	657172
IE_3500_3000_MM	BA64	Shotgun	CassidyPNAS2016	0.82	0.18	0.02	0.02	0.03	975050
IE_3500_3000_MM	BA72	Shotgun	CassidyScience2020	0.78	0.22	0.02	0.02	0.92	778552
IE_3500_3000_MM	CA448	Shotgun	CassidyScience2020	0.78	0.22	0.02	0.02	0.27	784652
IE_3500_3000_MM	GNM1007	Shotgun	CassidyScience2020	0.77	0.23	0.02	0.02	0.77	635586
IE_3500_3000_MM	MB6	Shotgun	CassidyScience2020	0.83	0.17	0.02	0.02	0.48	742579
IE_3500_3000_MM	NC10	Shotgun	CassidyScience2020	0.78	0.22	0.02	0.02	0.80	975947
IE_3500_3000_MM	NC27	Shotgun	CassidyScience2020	0.78	0.22	0.02	0.02	0.88	938997
IE_3500_3000_MM									

Table S5. fastGLOBETROTTER inference for admixture between Early European Farmers and Hunter-Gatherer surrogate groups for ancient individuals from this study

Group label	Sample ID	fastGT inferred generation	95% CI fastGT jackknifing	Sourcefind estimated HG ancestry
FR N 5100 4900_LBK	BCP196	20	4	12%
FR N 5100 4900_LBK	BLF81	30	9	7%
FR N 5100 4900_LBK	BLF1489	11	3	6%
FR N 5100 4900_LBK	BVT60910	24	6	11%
FR N 5100 4900_LBK	CCF606	18	3	10%
FR N 5100 4900_LBK	CGH12	28	5	12%
FR N 5100 4900_LBK	MDV186	15	4	7%
FR N 5100 4900_LBK	MDV13-1	18	8	8%
FR N 5100 4900_LBK	MDV131	25	3	10%
FR N 5100 4900_LBK	MDV144	24	10	6%
FR N 5100 4900_LBK	MDV189-2.2	18	4	12%
FR N 5100 4900_LBK	MDV195	29	13	8%
FR N 5100 4900_LBK	MDV208	24	8	4%
FR N 5100 4900_LBK	MDV214	16	4	9%
FR N 5100 4900_LBK	MDV242	17	7	11%
FR N 5100 4900_LBK	MDV272	19	4	20%
FR N 5100 4900_LBK	MDV629	28	7	10%
FR N 5100 4900_LBK	MDV93.2	22	3	15%
FR N 5100 4900_LBK	MDV188-4	34	7	15%
FR N 5100 4900_LBK	MDV188-6.3	20	3	10%
FR N 5100 4900_LBK	MDV561-1	10	4	5%
FR N 5100 4900_LBK	MDV571-7	32	5	14%
FR N 5100 4900_LBK	MGA45	20	5	10%
FR N 5100 4900_LBK	SMP4738	9	1	46%
FR N 5100 4900_LBK	SMP451	23	5	12%
FR N 5100 4900_LBK	SMP471	19	3	10%
FR N 4900 4700_VSG	BFM327	13	4	11%
FR N 4900 4700_VSG	BFM369	33	5	14%
FR N 4250 3700_MK	BLP10	21	12	17%
FR N 4250 3700_MK	BLP69	49	10	15%
FR N 4250 3700_MK	BCG21	32	7	19%
FR N 4250 3700_MK	CCF488	39	6	18%
FR N 4700 4300_Gerny	BLP31	10	3	13%
FR N 4250 3700_Esc	Esc42	19	3	24%
FR N 4250 3700_Esc	Esc75	35	8	21%
FR N 4250 3700_Esc	Esc97	36	4	26%
FR N 4250 3700_Esc	Esc791	50	6	20%
FR N 4250 3700_Esc	Esc10188	35	4	22%
FR N 3600 3400_Post_MK	BLP68	53	10	16%
FR N 3600 3400_Post_MK	CCF315-1	53	9	17%
FR N 3600 3400_Post_MK	BLG247-1	16	4	81%
FR N 3600 3400_Post_MK	BLG247-3	14	4	51%
FR N 3400 2900_Chouilly_H10	CHOU3935	57	13	31%
FR N 3400 2900_Chouilly_H10	CHOU5958	41	9	25%
FR N 3400 2900_Chouilly_H10	CHOU6908	37	12	26%
FR N 3400 2900_Chouilly_H10	CHOU7237	21	7	24%
FR N 3400 2900_Chouilly_H11	CHOU9742	90	23	23%
FR N 3400 2900_Chouilly_H11	CHOU9743	136	45	25%
FR N 3400 2900_Chouilly_H11	CHOUF11P6D13	91	14	25%
FR N 3400 2900_Chouilly_H11	CHOUF11P7B22	71	24	23%

Table S6. fastGLOBETROTTER inference for admixture between Early European Farmer and Hunter-Gatherer surrogate groups for radiocarbon dated ancient individuals

Target Group	Target Individual	Generations	95% CI	Date calBCE (OxCal mean date ± Standard deviation)	Admixture date BCE	Upper limit of admixture date BCE	Lower limit of admixture date BCE
Italy ICC	R16	12	7	5301-5214	5602	5795	5408
Italy ICC	R17	11	3	5324-5223	5583	5654	5511
Italy ICC	R19	11	12	5345-5221	5594	5935	5253
Italy ICC	R6	26	7	5242-5177	5924	6132	5716
Spain Cardial	Mur	9	5	5249-5107	5416	5567	5266
Spain Cardial	CB13	13	4	5445-5347	5766	5875	5657
Germany LBK	DER009	19	3	5152-5006	5599	5672	5527
Germany LBK	Herx	13	5	5164-4993	5441	5569	5314
Germany LBK	Dil16	15	5	5235-4998	5526	5656	5397
Hungary ALP	NE2	18	4	5297-5048	5679	5802	5556
Hungary ALP	NE5	21	7	5198-5020	5692	5900	5484
France-Aisne LBK	MDV188-6.3	20	3	5038-4916	5545	5633	5457
France-Aisne LBK	MGA45	20	5	5171-5043	5673	5807	5538
France-Aisne LBK	MDV93.2	22	2	5130-4976	5668	5725	5610
France-Aisne LBK	CCF606	18	3	4912-4778	5359	5453	5264
France-Aisne LBK	MDV561-1	10	4	5173-5055	5407	5532	5282
France-Aisne LBK	MDV571-7	32	5	5154-5000	5975	6102	5849
France-Champagne LBK	SMP471	19	3	4984-4872	5463	5550	5376
France-Champagne LBK	SMP4738	9	1	5177-5045	5364	5404	5325
Italy-Sicily Neolithic	I4062	44	8	4984-4784	6108	6330	5886
Italy-Sicily Neolithic	I4063	30	8	4987-4794	5727	5947	5507
France-Aisne Michelsberg	CCF488	39	6	3851-3689	4866	5036	4696
France-Aisne Michelsberg	BCG21	32	7	3888-3728	4706	4895	4517
France-Aisne Michelsberg	BLP10	21	12	4334-4056	4779	4977	4807
France-Aisne Cerny	BLP31	10	3	4684-4564	4892	5118	4439
Italy-Sardinia Neolithic	I16168	75	26	3949-3714	5931	6651	5211
France-Aisne Post-Michelsberg	BLG247-1	16	4	3494-3398	3896	4017	3775
France-Aisne Post-Michelsberg	BLG247-3	14	4	3494-3398	3830	3945	3715
France-Aisne Post-Michelsberg	CCF315-1	53	9	3627-3515	5044	5307	4780
France-Aisne Post-Michelsberg	BLP68	53	10	3563-3409	4967	5243	4691
France-Champagne Post-Michelsberg	PSS4693	58	10	3605-3349	5094	5383	4806
France Mont-Aimé	2H11	26	4	3288-3138	3929	4036	3823
France Mont-Aimé	2H10	55	9	3341-3098	4744	5004	4485
England Early Neolithic	Aveline1	52	7	3631-3523	5035	5221	4850
England Early Neolithic	Carsington pasture1	47	7	3621-3537	4891	5095	4687
Italy Chalcolithic	R24	55	9	3626-3374	5034	5292	4776
Sweden Megalithic	Ansarve8	28	7	3340-3030	3972	4162	3781
Sweden Megalithic	Ansarve17	27	5	3330-2930	3882	4018	3746
Sweden Megalithic	Gökhem	34	9	3050-2750	3857	4123	3591
Germany Wartberg	KH150620	18	6	3334-3094	3729	3892	3566
Germany Wartberg	KH150615	15	5	3334-3025	3624	3750	3497
Germany Wartberg	KH150610	27	7	3365-3102	3977	4178	3776
Germany Wartberg	KH150635	20	2	3322-2924	3653	3723	3584

Table S7. Phenotypic inference with the HirisPlex-S system

Group Label	Individual	Eye Color			Hair Color				Hair Shade		Skin Color				
		Blue	Intermediate	Brown	Blond	Brown	Red	Black	Light	Dark	Very Pale	Pale	Intermediate	Dark	Dark to Black
FR_N_3600_3400_Post_MK	BLG2471	0.77	0.09	0.14	0.02	0.71	0.00	0.27	0.08	0.92	0.04	0.13	0.80	0.03	0.00
FR_N_3600_3400_Post_MK	BLG2472	0.84	0.09	0.07	0.13	0.67	0.00	0.20	0.41	0.59	0.03	0.30	0.67	0.01	0.00
FR_N_3600_3400_Post_MK	BLG2473	0.90	0.06	0.04	0.12	0.67	0.00	0.21	0.55	0.45	0.05	0.49	0.46	0.00	0.00
FR-Aisne Mesolithic	CCF711	0.56	0.19	0.25	0.00	0.58	0.00	0.42	0.01	0.99	0.00	0.00	0.61	0.31	0.09
FR_N_5100_4900_LBK	BLF54	0.65	0.12	0.23	0.27	0.53	0.00	0.19	0.73	0.27	0.00	0.00	0.67	0.04	0.28
FR_N_5100_4900_LBK	BCP196	0.00	0.00	1.00	0.00	0.30	0.00	0.70	0.01	0.99	0.00	0.00	0.31	0.67	0.01
FR_N_5100_4900_LBK	BLF1489	0.00	0.00	1.00	0.01	0.38	0.00	0.61	0.01	0.99	0.00	0.00	0.09	0.83	0.08
FR_N_5100_4900_LBK	BLF81	0.00	0.03	0.97	0.03	0.69	0.00	0.28	0.05	0.95	0.01	0.26	0.72	0.01	0.00
FR_N_5100_4900_LBK	BLH417	0.01	0.03	0.96	0.02	0.51	0.00	0.47	0.03	0.97	0.00	0.01	0.12	0.70	0.16
FR_N_5100_4900_LBK	BVT60910	0.03	0.07	0.90	0.10	0.56	0.00	0.34	0.20	0.80	0.04	0.16	0.55	0.25	0.00
FR_N_5100_4900_LBK	CCF606	0.00	0.00	1.00	0.01	0.29	0.00	0.71	0.02	0.98	0.00	0.02	0.50	0.48	0.00
FR_N_5100_4900_LBK	CGH12	0.01	0.03	0.97	0.02	0.42	0.00	0.56	0.05	0.95	s	0.01	0.54	0.44	0.00
FR_N_5100_4900_LBK	MDV186	0.01	0.05	0.94	0.12	0.51	0.00	0.37	0.42	0.58	0.00	0.12	0.70	0.07	0.11
FR_N_5100_4900_LBK	MDV131	0.03	0.06	0.91	0.09	0.54	0.00	0.37	0.34	0.66	0.00	0.14	0.75	0.10	0.00
FR_N_5100_4900_LBK	MDV131	0.00	0.00	1.00	0.00	0.32	0.00	0.68	0.00	1.00	0.00	0.00	0.06	0.92	0.02
FR_N_5100_4900_LBK	MDV144	0.65	0.12	0.23	0.30	0.54	0.00	0.15	0.64	0.36	0.01	0.14	0.63	0.18	0.04
FR_N_5100_4900_LBK	MDV188-4	0.87	0.06	0.07	0.29	0.61	0.00	0.10	0.61	0.39	0.03	0.25	0.69	0.02	0.00
FR_N_5100_4900_LBK	MDV188-6.3	0.00	0.01	0.99	0.00	0.34	0.00	0.66	0.01	0.99	0.00	0.09	0.82	0.08	0.01
FR_N_5100_4900_LBK	MDV189-2.2	0.00	0.02	0.98	0.02	0.56	0.00	0.42	0.11	0.89	0.00	0.00	0.14	0.30	0.56
FR_N_5100_4900_LBK	MDV195	0.00	0.01	0.99	0.02	0.55	0.00	0.43	0.03	0.97	0.00	0.02	0.36	0.61	0.01
FR_N_5100_4900_LBK	MDV208	0.00	0.00	1.00	0.00	0.25	0.00	0.74	0.01	0.99	0.00	0.00	0.04	0.70	0.25
FR_N_5100_4900_LBK	MDV213	0.00	0.01	0.99	0.06	0.64	0.01	0.29	1.00	0.00	0.01	0.06	0.92	0.00	0.01
FR_N_5100_4900_LBK	MDV214	0.00	0.01	0.99	0.10	0.56	0.00	0.34	0.21	0.79	0.00	0.05	0.59	0.22	0.14
FR_N_5100_4900_LBK	MDV233	0.00	0.01	0.99	0.02	0.48	0.00	0.50	0.07	0.93	0.00	0.05	0.61	0.27	0.07
FR_N_5100_4900_LBK	MDV234	0.03	0.06	0.91	0.09	0.57	0.00	0.34	0.20	0.80	0.00	0.10	0.49	0.36	0.05
FR_N_5100_4900_LBK	MDV242	0.00	0.01	0.99	0.00	0.27	0.00	0.73	0.00	1.00	0.00	0.00	0.11	0.89	0.00
FR_N_5100_4900_LBK	MDV272	0.03	0.06	0.91	0.06	0.74	0.00	0.20	0.21	0.79	0.00	0.08	0.65	0.25	0.01
FR_N_5100_4900_LBK	MDV5611	0.05	0.08	0.87	0.11	0.58	0.00	0.31	0.25	0.75	0.00	0.06	0.67	0.22	0.05
FR_N_5100_4900_LBK	MDV571-7	0.00	0.01	0.99	0.03	0.47	0.00	0.49	0.06	0.94	0.01	0.06	0.55	0.39	0.00
FR_N_5100_4900_LBK	MDV629	0.00	0.00	1.00	0.00	0.38	0.00	0.62	0.00	1.00	0.00	0.00	0.13	0.86	0.00
FR_N_5100_4900_LBK	MDV93-2	0.01	0.05	0.94	0.13	0.61	0.00	0.26	0.51	0.49	0.00	0.10	0.87	0.02	0.01
FR_N_5100_4900_LBK	MDV93-3	0.01	0.05	0.94	0.09	0.69	0.00	0.22	0.30	0.70	0.00	0.06	0.66	0.26	0.01
FR_N_5100_4900_LBK	MG445	0.00	0.00	1.00	0.00	0.51	0.00	0.49	0.01	0.99	0.00	0.00	0.04	0.58	0.38
FR_N_5100_4900_LBK	SMP4738	0.00	0.00	1.00	0.00	0.18	0.00	0.82	0.00	1.00	0.00	0.02	0.68	0.15	0.15
FR_N_5100_4900_LBK	SMP45	0.01	0.05	0.94	0.09	0.65	0.00	0.26	0.16	0.84	0.01	0.16	0.83	0.00	0.00
FR_N_5100_4900_LBK	SMP471	0.02	0.09	0.89	0.02	0.62	0.00	0.36	0.07	0.93	0.03	0.21	0.76	0.01	0.00
FR_N_4900_4700_VSG	BFM327	0.00	0.00	1.00	0.01	0.29	0.00	0.70	0.02	0.98	0.00	0.02	0.34	0.64	0.00
FR_N_4900_4700_VSG	BFM369	0.00	0.01	0.99	0.00	0.31	0.00	0.69	0.01	0.99	0.00	0.00	0.01	0.54	0.45
FR_N_4250_3700_MK	BCG21	0.34	0.11	0.55	0.08	0.69	0.00	0.23	0.19	0.81	0.02	0.06	0.76	0.17	0.00
FR_N_4250_3700_MK	BLP69	0.01	0.05	0.94	0.09	0.57	0.00	0.34	0.33	0.67	0.00	0.11	0.81	0.03	0.04
FR_N_4250_3700_MK	CCF488	0.00	0.00	1.00	0.00	0.33	0.00	0.67	0.01	0.99	0.00	0.00	0.14	0.83	0.02
FR_N_4250_3700_Esc	Esc1018B	0.00	0.00	1.00	0.00	0.31	0.00	0.69	0.00	1.00	0.00	0.04	0.64	0.31	0.00
FR_N_4250_3700_Esc	Esc42	0.04	0.08	0.89	0.13	0.58	0.00	0.29	0.31	0.69	0.01	0.13	0.66	0.20	0.01
FR_N_4250_3700_Esc	Esc48	0.01	0.03	0.97	0.03	0.38	0.00	0.59	0.21	0.79	0.00	0.08	0.83	0.05	0.04
FR_N_4250_3700_Esc	Esc637	0.03	0.06	0.91	0.08	0.47	0.00	0.45	0.36	0.64	0.01	0.22	0.77	0.00	0.00
FR_N_4250_3700_Esc	Esc75	0.43	0.12	0.45	0.08	0.63	0.00	0.28	0.18	0.82	0.02	0.07	0.53	0.37	0.01
FR_N_4250_3700_Esc	Esc791	0.02	0.09	0.89	0.02	0.58	0.00	0.40	0.04	0.96	0.01	0.13	0.78	0.07	0.02
FR_N_4250_3700_Esc	Esc97	0.00	0.03	0.97	0.02	0.46	0.00	0.52	0.08	0.92	0.01	0.07	0.72	0.20	0.00
FR_N_3600_3400_Post_MK	BLP31	0.00	0.00	1.00	0.00	0.34	0.00	0.66	0.00	1.00	0.00	0.00	0.06	0.91	0.03
FR_N_3600_3400_Post_MK	BLP68	0.00	0.01	0.99	0.03	0.39	0.00	0.59	0.09	0.91	0.00	0.15	0.80	0.04	0.01
FR_N_3600_3400_Post_MK	CCF3151	0.03	0.06	0.91	0.14	0.58	0.00	0.28	0.52	0.48	0.01	0.16	0.82	0.01	0.00
FR_N_3400_2900_Chouilly_H10	CHOU3935	0.01	0.04	0.96	0.01	0.46	0.00	0.54	0.01	0.99	0.01	0.04	0.49	0.45	0.01
FR_N_3400_2900_Chouilly_H10	CHOU5980	0.00	0.00	1.00	0.00	0.34	0.00	0.66	0.00	1.00	0.00	0.00	0.05	0.90	0.05
FR_N_3400_2900_Chouilly_H10	CHOU6908	0.00	0.03	0.97	0.03	0.56	0.00	0.42	0.03	0.97	0.01	0.02	0.28	0.69	0.00
FR_N_3400_2900_Chouilly_H10	CHOU7237	0.03	0.06	0.91	0.06	0.64	0.00	0.29	0.12	0.88	0.00	0.08	0.47	0.42	0.03
FR_N_3400_2900_Chouilly_H11	CHOU9742	0.65	0.12	0.23	0.26	0.60	0.00	0.13	0.62	0.38	0.01	0.14	0.78	0.07	0.00
FR_N_3400_2900_Chouilly_H11	CHOU9743	0.01	0.03	0.97	0.02	0.42	0.00	0.55	0.08	0.92	0.00	0.02	0.55	0.27	0.16
FR_N_3400_2900_Chouilly_H11	CHOUF11P6D13	0.00	0.00	1.00	0.00	0.31	0.00	0.69	0.00	1.00	0.00	0.01	0.09	0.73	0.18
FR_N_3400_2900_Chouilly_H11	CHOUF11P7B122	0.00	0.01	0.99	0.02	0.55	0.00	0.43	0.03	0.97	0.00	0.04	0.58	0.36	0.03
FR_N_3400_2900_Chouilly_H11	CHOUF11P7B22	0.00	0.05	0.95	0.01	0.59	0.00	0.40	0.02	0.98	0.08	0.36	0.56	0.00	0.00
ES_3200_2500_CA	CYH1	0.00	0.00	1.00	0.00	0.36	0.00	0.63	0.01	0.99	0.00	0.00	0.04	0.95	0.01

References

1. Ilett M. Linear Pottery and Blicquy/Villeneuve-Saint-Germain settlement in the Aisne valley and its environs: an overview. In: Wolfram, Sabine, Stäuble, Harald, eds. *Siedlungsstruktur Und Kulturwandel in Der Bandkeramik: Beiträge Der Internationalen Tagung „Neue Fragen Zur Bandkeramik Oder Alles Beim Alten?“ Leipzig 23. Bis 24. September 2010.* ; 2012.
2. Colas C, Allard P, Chartier M, et al. LES ENSEMBLES MONUMENTAUX DU NÉOLITHIQUE MOYEN DE BEAURIEUX “LA PLAINE.” *Dir Régionale Antiq Hist Picardie Cent Rech Archéologiques Val Oise*. Published online January 1, 2018.
3. Dubouloz J. Lecture multiscale des enceintes du néolithique moyen (4500-3800 bc) en France du nord : hypothèses sur la structuration économique, sociale et politique au néolithique. In: *Les Sites Ceinturés de La Préhistoire Récente : Nouvelles Données, Nouvelles Approches, Nouvelles Hypothèses*. Éditions Archives d'Écologie Préhistorique; 2018:197-213.
4. Thevenet C, Baillieu M, Dubouloz J. Une sépulture post-Michelsberg (?) à Berry-au-Bac “la Croix Maigret” (Aisne). In: *La Revue Archéologique de Picardie*; 2015.
5. Rivollat M, Mendisco F, Pemonge MH, et al. When the Waves of European Neolithization Met: First Paleogenetic Evidence from Early Farmers in the Southern Paris Basin. *PLOS ONE*. 2015;10(4):e0125521. doi:10.1371/journal.pone.0125521
6. Rivollat M, Réveillas H, Mendisco F, et al. Ancient mitochondrial DNA from the middle neolithic necropolis of Obernai extends the genetic influence of the LBK to west of the Rhine. *Am J Phys Anthropol*. 2016;161(3):522-529. doi:10.1002/ajpa.23055
7. Rivollat M, Jeong C, Schiffels S, et al. Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers. *Sci Adv*. 2020;6(22):eaaz5344. doi:10.1126/sciadv.aaz5344
8. Brunel S, Bennett EA, Cardin L, et al. Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proc Natl Acad Sci*. 2020;117(23):12791-12798. doi:10.1073/pnas.1918034117
9. Rootsi S, Kivisild T, Benuzzi G, et al. Phylogeography of Y-Chromosome Haplogroup I Reveals Distinct Domains of Prehistoric Gene Flow in Europe. *Am J Hum Genet*. 2004;75(1):128-137. doi:10.1086/422196
10. Szécsényi-Nagy A, Brandt G, Haak W, et al. Tracing the genetic origin of Europe’s first farmers reveals insights into their social organization. *Proc R Soc B Biol Sci*. 2015;282(1805):20150339. doi:10.1098/rspb.2015.0339
11. Martiniano R, De Sanctis B, Hallast P, Durbin R. Placing Ancient DNA Sequences into Reference Phylogenies. *Mol Biol Evol*. 2022;39(2):msac017. doi:10.1093/molbev/msac017
12. Hanghøj K, Moltke I, Andersen PA, Manica A, Korneliussen TS. Fast and accurate relatedness estimation from high-throughput sequencing data in the presence of inbreeding. *GigaScience*. 2019;8(5):giz034. doi:10.1093/gigascience/giz034
13. Yaka R, Mapelli I, Kaptan D, et al. Variable kinship patterns in Neolithic Anatolia revealed by ancient genomes. *Curr Biol*. 2021;31(11):2455-2468.e18. doi:10.1016/j.cub.2021.03.050
14. Karruggia JP, Guichard Y, Hachem L. Les ensembles funéraires rubanés de Menneville « Derrière le Village » (Aisne). In: ARTEHIS Éditions; 1996:119-174. doi:10.4000/books.artehis.571
15. Thevenet C. L’Enceinte rubanée de Menneville Derrière le Village et les structures associées (Aisne, France): De la diversité du traitement des défunts à la cohérence d’un système. *Gall Préhistoire*. Published online September 1, 2016:32-92. doi:10.4000/galliap.341
16. Cassidy LM, Maoldúin RÓ, Kador T, et al. A dynastic elite in monumental Neolithic society. *Nature*. 2020;582(7812):384-388. doi:10.1038/s41586-020-2378-6

17. Lipson M, Szécsényi-Nagy A, Mallick S, et al. Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature*. 2017;551(7680):368-372. doi:10.1038/nature24476
18. Arzelier A, Rivollat M, De Belvalet H, et al. Neolithic genomic data from Southern France showcase intensified interactions with hunter-gatherer communities. *iScience*. Published online October 19, 2022:105387. doi:10.1016/j.isci.2022.105387
19. Lawson DJ, Hellenthal G, Myers S, Falush D. Inference of Population Structure using Dense Haplotype Data. *PLOS Genet*. 2012;8(1):e1002453. doi:10.1371/journal.pgen.1002453
20. Fu Q, Posth C, Hajdinjak M, et al. The genetic history of Ice Age Europe. *Nature*. 2016;534(7606):200-205. doi:10.1038/nature17993
21. Bortolini E, Pagani L, Oxilia G, et al. Early Alpine occupation backdates westward human migration in Late Glacial Europe. *Curr Biol*. 2021;31(11):2484-2493.e7. doi:10.1016/j.cub.2021.03.078
22. Jones ER, Gonzalez-Forbes G, Connell S, et al. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat Commun*. 2015;6(1):8912. doi:10.1038/ncomms9912
23. Charlton S, Brace S, Hajdinjak M, et al. Dual ancestries and ecologies of the Late Glacial Palaeolithic in Britain. *Nat Ecol Evol*. 2022;6(11):1658-1668. doi:10.1038/s41559-022-01883-z
24. Villalba-Mouco V, van de Loosdrecht MS, Posth C, et al. Survival of Late Pleistocene Hunter-Gatherer Ancestry in the Iberian Peninsula. *Curr Biol*. 2019;29(7):1169-1177.e7. doi:10.1016/j.cub.2019.02.006
25. Kılınç GM, Omrak A, Özer F, et al. The Demographic Development of the First Farmers in Anatolia. *Curr Biol*. 2016;26(19):2659-2666. doi:10.1016/j.cub.2016.07.057
26. Ringbauer H, Novembre J, Steinrücken M. Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nat Commun*. 2021;12. doi:10.1038/s41467-021-25289-w
27. Svensson E, Günther T, Hoischen A, et al. Genome of Peștera Muierii skull shows high diversity and low mutational load in pre-glacial Europe. *Curr Biol*. 2021;31(14):2973-2983.e9. doi:10.1016/j.cub.2021.04.045
28. Marchi N, Winkelbach L, Schulz I, et al. The genomic origins of the world's first farmers. *Cell*. 2022;185(11):1842-1859.e18. doi:10.1016/j.cell.2022.04.008
29. Chaitanya L, Breslin K, Zuñiga S, et al. The HIRISplex-S system for eye, hair and skin colour prediction from DNA: Introduction and forensic developmental validation. *Forensic Sci Int Genet*. 2018;35:123-135. doi:10.1016/j.fsigen.2018.04.004
30. Brace S, Diekmann Y, Booth TJ, et al. Ancient genomes indicate population replacement in Early Neolithic Britain. *Nat Ecol Evol*. 2019;3(5):765-771. doi:10.1038/s41559-019-0871-9
31. Schiffels S, Wang K. MSMC and MSMC2: The Multiple Sequentially Markovian Coalescent. In: Dutheil JY, ed. *Statistical Population Genomics*. Springer US; 2020:147-166. doi:10.1007/978-1-0716-0199-0_7
32. Bergström A, McCarthy SA, Hui R, et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science*. 2020;367(6484):eaay5012. doi:10.1126/science.aay5012
33. Shennan S. *The First Farmers of Europe: An Evolutionary Perspective*. Cambridge University Press; 2018. doi:10.1017/9781108386029
34. Wangkumhang P, Greenfield M, Hellenthal G. An efficient method to identify, date, and describe admixture events using haplotype information. *Genome Res*. Published online July 6, 2022. doi:10.1101/gr.275994.121

35. Antonio ML, Gao Z, Moots HM, et al. Ancient Rome: A genetic crossroads of Europe and the Mediterranean. *Science*. 2019;366(6466):708-714. doi:10.1126/science.aay6826
36. Chintalapati M, Patterson N, Moorjani P. The spatiotemporal patterns of major human admixture events during the European Holocene. Perry GH, ed. *eLife*. 2022;11:e77625. doi:10.7554/eLife.77625
37. Guan Y. Detecting Structure of Haplotypes and Local Ancestry. *Genetics*. 2014;196(3):625-642. doi:10.1534/genetics.113.160697
38. Martiniano R, Cassidy LM, Ó'Maoldúin R, et al. The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genet*. 2017;13(7):e1006852. doi:10.1371/journal.pgen.1006852
39. Allard P. The Mesolithic-Neolithic transition in the Paris Basin: a review. In: Whittle A, Cummings V, eds. *Going Over: The Mesolithic-Neolithic Transition in North-West Europe*. British Academy; 2007:0. doi:10.5871/bacad/9780197264140.003.0011
40. Bollongino R, Nehlich O, Richards MP, et al. 2000 Years of Parallel Societies in Stone Age Central Europe. *Science*. 2013;342(6157):479-481. doi:10.1126/science.1245049
41. López S, Tarekegn A, Band G, et al. Evidence of the interplay of genetics and culture in Ethiopia. *Nat Commun*. 2021;12(1):3581. doi:10.1038/s41467-021-23712-w
42. Sheridan A. The Neolithization of Britain and Ireland: The 'Big Picture.' In: Finlayson, Bill, Warren, Graeme, eds. *Landscapes in Transition*. Oxbow Books; 2010.
43. Whittle AWR, Healy F, Bayliss A, Allen MJ. *Gathering Time: Dating the Early Neolithic Enclosures of Southern Britain and Ireland*. Oxbow Books; 2011. <https://books.google.fr/books?id=OV5ApwAACAAJ>
44. Cassidy LM, Martiniano R, Murphy EM, et al. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proc Natl Acad Sci*. 2016;113(2):368-373. doi:10.1073/pnas.1518445113
45. Olalde I, Brace S, Allentoft ME, et al. The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature*. 2018;555(7695):190-196. doi:10.1038/nature25738
46. Rohland N, Mallick S, Mah M, Maier R, Patterson N, Reich D. Three Reagents for in-Solution Enrichment of Ancient Human DNA at More than a Million SNPs. *bioRxiv*. Published online January 1, 2022:2022.01.13.476259. doi:10.1101/2022.01.13.476259
47. Bennett EA, Weber J, Bendhafer W, et al. The genetic identity of the earliest human-made hybrid animals, the kungas of Syro-Mesopotamia. *Sci Adv*. 2022;8(2):eabm0218. doi:10.1126/sciadv.abm0218
48. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Published online 2010. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
49. Renaud G, Stenzel U, Kelso J. leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res*. 2014;42(18):e141-e141. doi:10.1093/nar/gku699
50. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760. doi:10.1093/bioinformatics/btp324
51. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079. doi:10.1093/bioinformatics/btp352
52. Jagadeesan A, Ebenesersdóttir SS, Guðmundsdóttir VB, Thordardottir EL, Moore KHS, Helgason A. HaploGrouper: a generalized approach to haplogroup classification. *Bioinformatics*. 2021;37(4):570-572. doi:10.1093/bioinformatics/btaa729
53. Weissensteiner H, Forer L, Fendt L, et al. Contamination detection in sequencing studies using the mitochondrial phylogeny. *Genome Res*. 2021;31(2):309-316. doi:10.1101/gr.256545.119

54. Korneliussen TS, Albrechtsen A, Nielsen R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics*. 2014;15. doi:10.1186/s12859-014-0356-4
55. Rubinacci S, Ribeiro DM, Hofmeister RJ, Delaneau O. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nat Genet*. 2021;53(1):120-126. doi:10.1038/s41588-020-00756-0
56. Mathieson I, Lazaridis I, Rohland N, et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*. 2015;528(7583):499-503. doi:10.1038/nature16152
57. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet*. 2007;81(3):559-575. doi:10.1086/519795
58. Leslie S, Winney B, Hellenthal G, et al. The fine-scale genetic structure of the British population. *Nature*. 2015;519(7543):309-314. doi:10.1038/nature14230
59. Chacón-Duque JC, Adhikari K, Fuentes-Guajardo M, et al. Latin Americans show wide-spread *Converso* ancestry and imprint of local Native ancestry on physical appearance. *Nat Commun*. 2018;9(1):5388. doi:10.1038/s41467-018-07748-z
60. Busing FMTA, Meijer E, Leeden RVD. Delete-m Jackknife for Unequal m. *Stat Comput*. 1999;9(1):3-8. doi:10.1023/A:1008800423698
61. Harney É, Patterson N, Reich D, Wakeley J. Assessing the performance of qpAdm: a statistical tool for studying population admixture. *Genetics*. 2021;217(4):iyaa045. doi:10.1093/genetics/iyaa045
62. Patterson N. Ancient admixture in human history. *Genetics*. 2012;192. doi:10.1534/genetics.112.145037
63. Fu Q, Li H, Moorjani P, et al. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature*. 2014;514(7523):445-449. doi:10.1038/nature13810
64. Seguin-Orlando A, Korneliussen TS, Sikora M, et al. Genomic structure in Europeans dating back at least 36,200 years. *Science*. 2014;346(6213):1113-1118. doi:10.1126/science.aaa0114
65. Sikora M, Seguin-Orlando A, Sousa VC, et al. Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science*. 2017;358(6363):659-662. doi:10.1126/science.aao1807
66. Haak W, Lazaridis I, Patterson N, et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*. 2015;522(7555):207-211. doi:10.1038/nature14317
67. Jones ER, Zarina G, Moiseyev V, et al. The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers. *Curr Biol*. 2017;27(4):576-582. doi:10.1016/j.cub.2016.12.060

Chapter III

Late Neolithic family burial reveals admixture dynamics during the third millennium BCE and the shaping of the European genome

Oğuzhan Parasayan¹, Christophe Laurelut², Christine Bôle³, Lola Bonnabel², Alois Corona⁴, Cynthia Domenech-Jaulneau⁴, Cécile Paresys⁵, Isabelle Richard⁶, Thierry Grange^{1*}, Eva-Maria Geigl^{1*}

¹Paris Cité University, CNRS, Institut Jacques Monod, Paris, France

²INRAP Grand Est - UMR 8215 Trajectoires (CNRS-University Paris I), Reims, France
Inrap, UMR 8215 Trajectoires

³Genomics Core Facility, Institut Imagine-Structure Fédérative de Recherche Necker, INSERM U1163 et INSERM US24/CNRS UAR3633, Paris Descartes Sorbonne Paris Cité University, Paris, France

⁴Service Régional, Direction Régionale des Affaires culturelles d'Île-de-France, UMR 8215 Trajectoires (CNRS-University Paris I), Paris, France

⁵Ministère de la Culture, Service régional de l'archéologie d'Île-de-France - UMR 8215 Trajectoires, Paris, France

⁶Inrap Grand-Est, Châlons-en-Champagne, UMR 6472 CEPAM (CNRS- Nice University), France

***co-senior authors**

Abstract

The 3rd millennium BCE was a pivotal period characterized by profound cultural and genomic transformations in Europe associated with migration of populations from the Pontic-Caspian steppe that shaped the ancestry patterns in the present-day European genome. By means of high-resolution analysis of whole genomes of a multiple family burial in France dating to about 2500 cal BCE, haplotype phasing of the genomes and genome reconstruction of an unsampled individual through its relative's genomes, we shed light on the hitherto undetected early-stage admixture patterns, dynamics and propagation of steppe-ancestry in the Late Neolithic context in Europe. We identified two main Neolithic/steppe-ancestry admixture pulses dating to the arrival of steppe populations in Eastern Europe around 2900 BCE and to their progression into Western Europe around 2500 BCE. These pulses reveal different expansion dynamics throughout Europe with striking links to different cultural phenomena, Corded Ware and Bell Beaker.

Introduction

Past population migrations followed by admixture between migrating and autochthonous populations were discovered to leave traces in the genomes of the descendants, such as the first anatomically modern humans (AMHs) who migrated out of Africa and interbred with Neanderthals (1) or Upper Pleistocene AMH populations in Europe who were replaced by populations originating from southwest Asia (2). During the Holocene, two major population transformations occurred in Europe that shaped the present-day genomic landscape, the first one occurring in the course of the transition from the Mesolithic to the Neolithic and the second during the transition from the Late Neolithic to the Bronze Age. On the territory of present-day France the first one occurred at the end of the 6th millennium BCE (e.g., (3-5)), the second during the 3rd millennium BCE (e.g., (3, 4, 6)). While the first one was clearly associated with the introduction of the Neolithic cultures into Europe, the cultural association of the second one shows regional differences. The first genomic transformation was the consequence of Neolithic farmers (NFs), who descended from populations in western Anatolia and/or the Aegean (e.g., (7, 8)), migrating into Europe that had been occupied since the end of the last Ice Age by

Mesolithic hunter-gatherers (HG) (2). This led to admixture between the two populations (e.g., (3-5, 9-12)). The second transformation event was the consequence of gene flow from pastoralist populations of the European Pontic-Caspian steppe moving northwards and visible in burials associated first with the Corded Ware Culture (CWC) in eastern, Central and northern Europe and later in Bell Beaker and Bronze Age contexts (e.g., (3, 6, 13, 14)). On the territory of present-day France (further called “France”) this steppe ancestry has been found in Bell Beaker and Bronze Age burials dating to the mid 3rd millennium BCE (4, 6, 15). Similar observations were previously made in Central and northern Europe and Great Britain in archaeological contexts associated with the Corded Ware Culture (CWC) and the Bell Beaker Complex (BBC), respectively (3, 6, 13) (14). Thus, in contrast to the Neolithic transformation of Europe that introduced both a new lifestyle and a new genomic component, the Late Neolithic/Copper Age-Bronze Age transformation followed a more complex pattern involving 600 years of migration and spread of different populations and cultures in Europe and formation of a new social and economic order (16, 17).

Based on archaeological evidence, three successive Europe-wide migration phenomena have been postulated for the period between ca. 3100 to ca. 2450 BCE, namely (i) migrations of various steppe peoples with pit-grave (kurgan/yamna) burial traditions (currently called “Yamnaya”), (ii) of peoples in Central Europe associated with the CWC and (iii) of peoples using Bell Beaker funerary rituals (e.g., for review (17) and citations therein). The first two phenomena seemingly followed a (south)east-(north)west dispersal direction in Europe, i.e., around 3000 BCE from the Pontic Steppes (Dniester-Dnieper area), to Central Europe, (Hungary, Bohemia and Germany), where the CWC with its characteristic cord-decorated pottery formed out of the encounter of steppe and Late Neolithic traditions (16). The CWC also spread to the North Sea arriving in the 29th c. BCE (e.g., (16, 17) and references therein).

The origin of the third phenomenon, the BBC, is still unclear and controversially discussed. The major controversy is caused by the uncertainty of the dates associated with sites in which typical items of the Beaker package have been found, mostly individual items and rarely the entire BBC set, the most emblematic being a bell-shaped vessel, but also stone wrist-guards or flint knives of the type Grand Pressigny. Two hypotheses confront each other concerning the origin of the BBC: the first hypothesis proposes an origin of the Maritime beakers on the Iberian Peninsula where some of the earliest radiocarbon dates, around 2823-2658 cal BCE (4,220±45 BP), testify to the presence of the BBC (18-20). During a subsequent early phase, the BBC would have spread from the Mediterranean shores quickly northwards along the Atlantic coast and

along the Rhône-Saône valleys (21). The early BBC is proposed to have then developed into the Europe-wide Bell Beaker civilization with four geographically distinct Beaker cultures in the southwest, south, north and east of Europe (22, 23). The second hypothesis proposes an origin in the Lower Rhine, more specifically the Rhine-Meuse-Delta, where the culturally ambiguous maritime All-Over-Corded (AOC) and All-Over-Ornamented (AOO) beakers were found in the context of single graves dating to ca. 2,600 BCE, i.e., preceding the Maritime Bell Beakers. This led to the formulation of the so-called Dutch Model of BBC origins (24).

The problem with these hypotheses is that difficulties in the chronological resolution render dates imprecise owing to numerous radiocarbon plateaus throughout the 3rd millennium BCE (25). The resulting wide calibration spans have so far hampered thus far the identification of the cradle of the BBC. The absence of a clear answer to this question, however, does not obscure certain fundamental lines of evidence for large-scale processes. On the one hand, the Bell Beaker societies in the Iberian Peninsula, as successors of the sophisticated Chalcolithic Iberian societies with their fortified mega-sites, seemed to have catalyzed a truly Pan-European movement characterized by a vast trading network of goods like salt, copper, ceramics, ivory and amber within Europe as well as with Asia and Africa based on pre-existing Chalcolithic networks (26). On the other hand, a change in funerary rituals from the collective graves and megaliths, which had prevailed for 2000 years, to individual burials with the introduction of gender differentiation and warrior display, is interpreted as the consequence of CWC influence (23, 27). Indeed, the AOO beakers developed in the Rhine-Meuse delta in single grave contexts of the Vlaardingen culture (3400-2500 cal BCE), roughly contemporary to Bell Beaker pottery (28). In the 26th century BCE in northern and west-central France, individual graves with AOC/AOO beakers, maybe predating the Maritime beakers, argue in favor of a fast north-south “leap-frog” dispersal along preexisting Neolithic flint networks (17). In northern France, the AOC burials of Blignicourt dating to 2857-2488 cal BCE (29) and Ciry-Salsogne « La Bouche à Vesle » dating to 2,574-2,452 cal BCE (30) could be witnesses of this southward movement along the Rhine and Meuse lending support to parts of the *Rückstrom* hypothesis (31). Such a move might have been achieved in a lifetime as individual mobility was shown, based on ⁸⁷Sr/⁸⁶Sr-isotope analyses, to be high during this period, at least in some areas (32). Moreover, these two graves belong to a series of burials in France, the Meuse-Moselle areas and the eastern part of the Netherlands containing Grand Pressigny flint knives, hinting at a vast flint exchange network (28, 30).

The river Rhine, a natural environmental boundary, seems to have been the western limit of CWC distribution and a contact zone where CWC and BBC met, blended and spread further ((17) and references therein). The Middle Rhine (Meuse-Moselle region) may have seen the blending of the AOC Beakers from the Lower Rhine and the Maritime Beakers from the Upper Rhine, the region that has the largest Maritime beaker collection known in central Europe, dated to shortly after 2600 BCE (17). The origin of the latter, however, is unclear. While it cannot be excluded that they came from the Lower Rhine, it has been hypothesized that they rather traveled with Iberian pioneers moving north via the Rhône-Saône rivers spreading the BBC (17). In both scenarios, the Upper Rhine region would have been what V. Heyd calls “the spring board” for the Maritime Beaker radiation to the East.

These European-wide cultural transformations were seemingly accompanied by the transformation of the composition of populations. Indeed, paleogenomic analyses revealed large genomic shifts over short time periods that hint to migrations of people with steppe ancestry related to the Yamnaya from the Samara steppe in present-day Russia to northwest Europe and admixture with local Neolithic individuals, such as those associated with the Globular Amphora culture (GAC), a Late Neolithic culture (3400–2800 BCE) from the northern and eastern part of Central Europe involved in the admixture events with the people from the steppes (e.g., (3, 4, 6, 13-15, 33, 34)). This led to an east-west cline of steppe-ancestry in the CWC-associated population (35), with varying proportions of steppe-ancestry translating regional differences of interactions with groups coming from the steppe, which is also reflected in the material culture (34). Concomitantly with the dispersal throughout Europe of steppe ancestry, a shift in Y-chromosome haplotypes was detected with rapid expansion of the R1b M269 (R1b1a1b) lineage, similar to the one found in Samara Yamnaya (R1b Z2103, R1b1a1b1b) (3, 34), whereas in north-eastern Europe, the major Y haplogroup spreading during the CWC was the R1a M417 (R1a1a1) and its derivative Z465 (R1a1a1b) (36), indicating more complex steppe ancestry diffusion patterns than one just involving the Yamnaya.

The earliest European arrival of the steppe ancestry has hitherto been detected in Central Europe, in Bohemia, in individuals associated with the early CWC dating to about 3000-2800 BCE (14). The early Bohemian CWC individuals are generally characterized by high steppe ancestry, reaching 80-90%, and a Y-haplotype belonging to the R1b L151 lineage (R1b1a1b1a1a) and its following derivative U106 (R1b1a1b1a1a1). The later CWC-associated population is characterized by the additional presence of Latvian-like ancestry and the dominance of another Y-haplotype R1a-M417 (R1a1a1) that is more characteristic of north-

eastern Europe suggesting complex population dynamics in Central Europe during the CWC period (14). This complexity, as seen through the male Y-lineage, is also observed in Bohemia during the subsequent BBC expansion, between 2400 and 2000 BCE, where the R1b-P312 (R1b1a1b1a1a2), another derivative of the L151 lineage, became dominant alongside progressive erosion of the percentage of steppe-ancestry (14). Based on the pattern of the geographic distribution of the Y-haplogroups of the L151 lineage, Papac and colleagues suggested that the R1b-P312 lineage originated around the Lower Rhine valley followed by an expansion associated with the spread of the BBC, and thus that the BBC expansion into Bohemia was concomitant with an eastward migration from western steppe-ancestry groups (14). These replacement waves of Y-lineages hint to complex and changing male dominance patterns in groups whose movements are associated with cultural shifts.

The expansion of steppe-ancestry in western Europe is less precisely documented since the distribution in time and space of genetically analyzed ancient individuals is patchy and uneven. Nevertheless, thanks to a denser dataset, the timing of the arrival of the steppe-ancestry in present-day Switzerland could be estimated to have taken place around 2700 BCE, between 2860-2460 BCE, with a rapid increase upon arrival to ~60% followed by a decrease to 25-35% over the following 1000 years (37). If not due to a sampling bias, this arrival time may be slightly earlier than in Germany (3) (38), where it is detectable about 100 years later (37). The carriers of the steppe ancestry in Switzerland were found to be genetically similar to CWC-associated individuals from the Middle-Elbe-Saale region in Germany with a high steppe ancestry component suggesting that the population associated with this cultural horizon was relatively homogeneous across large parts of Central Europe (37). The Neolithic source population of the carriers of steppe-ancestry who arrived in Switzerland was closer to individuals associated with the GAC (37). Yet, not all individuals in Switzerland carried steppe-ancestry, and even 1000 years after its arrival, there were still individuals around without the corresponding signature. This pattern suggests limited admixture between carriers of steppe-ancestry and local Late Neolithic farmers.

Steppe-ancestry associated with Y haplogroup R1b P312 arrived later in the Iberian Peninsula, and is detected in northern Spain around 2400 BCE (39) and in southern Spain around 2200 BCE (40). Here, individuals with and without steppe-ancestry lived together for a few hundred years but ultimately steppe-ancestry penetrated into all genomes (6) (39, 40). Steppe-ancestry arrived in the British Isles at a similar period, around 2450 BCE, and this

arrival is associated with the spread of the BBC (6) and the introduction of Y-haplogroup R1b L21 (R1b1a1b1a1a2c1), derived from the P312 haplogroup (41).

Since its initial observation, a male bias in the steppe ancestry propagation in Europe has been contentious (42, 43). The sensitivity and specificity of the sex bias detected depend both on the methods used and the quality and richness of the X-chromosome data, with particular variations linked to the SNP density of the capture array and of the shotgun genome coverage. In CWC individuals from Germany, a sex-bias was detected in two individuals dated around 2600 BCE but was not detectable in the later Bronze age individuals (44). A sex bias has also been observed in Switzerland where the increase of Neolithic ancestry in steppe ancestry individuals between 2700 and 2000 BCE was driven preferentially by Neolithic females, during a period of cohabitation of these populations with two distinct ancestries (37). In Estonia, during the CWC between 2800 and 2000 BCE, the steppe ancestry propagation was also male-biased even though it involved a population distinct from the Yamnaya that carried Y-haplogroup R1a Z645 (36).

To better understand the spread of steppe ancestry in Western Europe, a clear view of the phenomenon in France is required due to its geographic position not only between eastern and Central Europe and the far west (British Isles), but also between the northwest and southwest of Europe (Iberian Peninsula). The 3rd millennium BCE transformation of the genomic landscape in France has, however, been sketched only in broad strokes. The genomes of roughly a dozen individuals in Bell Beaker and Bronze Age contexts in both northern and southern France are characterized by a substantial steppe ancestry-related genomic portion, around 50% (38-68%), and by males who were carriers of the steppe-related Y-chromosome haplotype group R1b-M269, with most individuals typed at higher resolution belonging to the P312 group (4, 6, 15). The aforementioned AOC burial CBV95 (Ciry-Salsogne « La Bouche à Vesle » in northern France), a single burial dated to 2574-2452 cal BCE, was a carrier of both a substantial genomic portion of steppe ancestry and the Y-chromosome R1bM269 lineage, while the BBC burial PEI2 (Dolmen des Peirières¹⁸, Villedubert, Aude, southern France) from a collective grave of men, women and children and dated to 2564-2302 BCE (3935 +/- 30BP), had no steppe ancestry and carried the Neolithic Y-chromosome lineage G2a2b2b1a1~ (4) akin to individual I1392 from Hegenheim, Alsace, dated to ,835-2472 calBCE/ 4047 +/- 29 BP (6). This heterogeneous situation was also encountered in the Iberian Peninsula where only a part of the BBC-associated burials carry steppe ancestry and individuals without any steppe-related ancestry were still present at least until 1950 BCE (6, 40).

From these various lines of evidence arises a picture of 3rd millennium BCE European societies of very diverse genetic origins facing each other with regionally different outcomes. The CWC seems to emerge after the encounter between Late Neolithic populations in Central Europe and various migrating steppe people from eastern Europe (14). In contrast, it is not clear whether the BBC results from the encounter between local Neolithic people and migrants with steppe-ancestry or if it developed fully prior to the arrival of steppe-ancestry. Indeed, most of the early Iberian BBC-associated individuals were found to be unadmixed descendants of the previous Neolithic populations (6). The propagation of the BBC from the southwest/west to Europe's north and east, such as southern Poland, involved carriers of both southern Neolithic ancestry and steppe ancestry of the R1b M269 and its derivative P312 (14, 34). Little is known, however, about the fine-scale processes at the individual level that led to these profound genomic transformations accompanying the cultural transformations hampered by the dating uncertainties. Genetic data from individuals involved in these processes could compensate this limitation and fill in the information gap.

Here we present the paleogenomic study of a Late Neolithic collective burial from the Paris Basin in northern France demonstrating a “real-time” observation of the particular circumstances under which the steppe ancestry component was first introduced into the Neolithic gene pool. Moreover, the genomes enabled us to model the origins of the different genomic components hence shedding light on the admixture dynamics and on questions of the origins, directions, timing and mechanisms of the expansion of steppe-ancestry at the end of the first half of the 3rd millennium BCE in western Europe as well as their initial features.

Results

Archaeological context of a Late Neolithic collective burial

The LBK settlement of “Bréviandes les Pointes et les Grèvottes » (hereafter “Bréviandes”), near Troyes in the southern Paris Basin, was also repeatedly used as a burial ground from the Early Neolithic (~5200 cal BCE) to the beginning of the Final Bronze Age (~1200 cal BCE). Several graves with few or no associated artifacts were radiocarbon-dated to the later Neolithic. Among them, a multiple burial of seven individuals in a subcircular pit named BRE445 (Fig. 1A) produced no other finds apart from a circular bone bead and a dog paw (Fig.S3). The bones of the individuals (BRE445 A-E, FK, HI, hereafter abbreviated as A, B, C, D, E, FK, HI) are rather well preserved and unfragmented. The burial comprises three adult women, an adult man, two young children and a neonate, a newborn or yet unborn baby. According to the anthropological analysis estimated that the three women A, B and E were respectively 20-39, 20-30 and around 60 years old at death, while the adult man was 20-39 years old. The two children HI and C, a boy and a girl, were respectively 4-8 and 6-10 years old at death. Direct radiocarbon dates on the bones were obtained from the individuals B, FK, and E and range between 2580-2275 cal BCE (98.9%), 2580-2284 cal BCE (99.6%) and 2706-2287 cal BCE (96.9%), respectively (see Table S1 and Fig. S8), assigning the burials to the Late Neolithic period. Apart from signs of stress and degeneracies the skeletons did not show traces of particular pathologies (see supplementary text). Variation in the bone morphology, as measured through the cnemic indices, discriminate individual FK from the others. Signs of violence were not identified. The temporal succession of the inhumations could only partially be determined based on stratigraphic data alone (see supplementary text). The spatial layout of the skeletons shows that some of the bones of the earlier burials had been disturbed on arrival of a new corpse, although there was no evidence for intentional spatial displacement. The archaeological and anthropological analysis concluded that this was most likely a family burial (see supplementary text).

Genetic and isotopic characterization of the individuals of the family burial at Bréviandes

In search of further information to understand the BRE445 burial, we produced whole genomes from the petrous bones of each of the individual of the burial with a coverage of 0.75X to 4.6 X (median coverage 1X). When we determined the haplogroups of the mitochondrial genomes and of the Y-chromosomes we obtained a first glimpse of the family relationships of the burials (Fig. 1B). Shared mitochondrial haplotypes together with age at death of the individuals suggest individuals FK, HI and D to be first generation descendants of E, A and B, respectively. FK and HI shared the same Y-chromosome haplotype R1b-L151 suggesting, if their age at death is taken into account, that they could be father and son. Only individual C does not share her mitochondrial haplotype with any other individual of the burial. These conclusions were confirmed when the whole genome data were analyzed with NgsRelate (45) and READ (46) and the results combined with the anthropological data allowing us to reconstruct the family relationships (Fig. 1B, Table S2): the burial consists of one nuclear family comprising three generations, together with a mother, her child and a single child who were not related genetically to the individuals of the nuclear family. The analysis of runs of homozygosity of their genomes using the package hapROH (47) showed that none of the individuals was inbred (Fig. S9). The combination of genetic, anthropological and archaeological results also enabled reconstruction of the most likely temporal sequence of the burials (Supplementary Materials). Since the adult man FK and his son HI were the first to be buried, the mother of FK and grandmother of HI in the middle of the burial and at an intermediate time point and the mother of HI at the end, we can conclude that the genetically unrelated individuals who are embedded within the burial structure must have been part of the social community that was buried here. The combined genetic, anthropological and archaeological results identify this collective grave as a singular case of a group with mixed social and biological ties important enough to be conserved in burial.

Through genotyping using the forensic Hirisplex-S assay (48) we reconstructed some phenotypic traits of the BRE445 individuals, such as the color of the eyes, the skin and the hair (Table S3). This analysis suggests that all of the individuals had brown eyes, that the hair color of individuals E and FK was lighter than of individuals A, HI, D, C, and B, and that B had the darkest hair color. It also suggests that B had the darkest skin tone followed by D, while the others had a lighter skin color.

To shed further light on the constitution of this group, we performed a stable strontium isotope analysis ($^{87}\text{Sr}/^{86}\text{Sr}$ ratio) on both teeth and bones of the adult individuals and on faunal remains to reveal potential life time mobility of the individuals (supplementary text and Table S9). The rationale behind this analysis is that the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio in teeth show the incorporation of Sr isotopes during infancy while the content of Sr isotopes in the bones reveals the incorporation during the last 10-20 years before death (49-51). The obtained values were compared with those obtained from contemporaneous faunal remains analyzed to reveal the local signature (Table S9). Combining the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio results and the age at death of the individuals, the results suggest potential lifetime mobility only for individuals E and B and lack of mobility for the other adult individuals. Finally, these values were compared with those mapped in present-day France (52) (supplementary text). This latter comparison proposes geographical regions in agreement with the values obtained. Here it suggests a mobility pattern for individuals E and B comprising but potentially also going beyond the Paris Basin including the north/northeast or east/southeast of France.

Genomic composition of the individuals of the BRE445 burial and haplotype reconstruction of the missing grandfather

To obtain a deeper insight into the potential genetic origins of the Bréviandes individuals, we performed ChromoPainter analysis (53) on a dataset of 116 imputed ancient genomes including only one representative of the related individuals of each “biological family” and excluding all other related individuals. With this dataset, we generated a haplotypic coancestry matrix, on which principal component analysis (PCA) and fineSTRUCTURE (53) were performed. Both the dimensionality reduction through PCA and the dendrogram analysis with fineSTRUCTURE allowed us to identify and visualize different genetic clusters and structures within the dataset. The first two PCs strongly differentiate three ancestral populations, i.e., northwest Anatolian Neolithic farmers (ANF), Western hunter-gatherers (WHG) and individuals associated with the Yamnaya culture of the Pontic steppe (Fig. 2A and Fig S10-11). PC1 separates ANF from HG/Yamnaya ancestries whereas PC2 separates WHG from Yamnaya ancestries. Since Yamnaya-related steppe ancestry is a mix of Eastern European hunter-gatherers (EHG) and Caucasian hunter-gatherers (CHG) (54), and since Western and Eastern HGs are also separated along PC2, PC2 could be considered as approximating a west to east cline of the European Mesolithic ancestries.

The European individuals from Neolithic to Bronze Age contexts, including those from Bréviandes, are positioned between the three ancestry poles in a manner that reflects their ancestry proportions (Fig. 2A and Fig. S10-11). Neolithic individuals without any steppe ancestry are distributed on a line that can be drawn between the ANF and the WHG clusters. The proportion of steppe-ancestry in the genomes of Late Neolithic and Yamnaya-associated individuals, estimated in the individuals grouped in four clusters using SOURCEFIND ((55); Fig. 2A a-d, bottom part), positions the carriers of steppe ancestry along a line of increasing steppe-ancestry. The five female individuals from Bréviandes cluster with Late Neolithic and Chalcolithic genomes from France and Spain, respectively, in cluster “a” and are devoid of steppe ancestry (Fig. 2A and Fig. S10-11). In contrast, the two related male individuals HI and FK, are positioned within clusters “b” and “c”, respectively, because of varying steppe ancestry proportions, that of the son HI being half of that of his father FK, in agreement with the absence of steppe ancestry found in his mother A (Fig.2A and Fig. S10-11).

Since the proportion of steppe ancestry found in FK being about 35-37%, we predicted that this proportion should be twice as high in his father, since the mother of FK, individual E, is devoid of this steppe ancestry (Fig. 2A and Table S4). This high steppe ancestry proportion in FK prompted us to reconstruct the phased genotype of the unsampled father of FK, herein referred to as YY, using the phased genotypes of FK and E. Since only one haploid genome of YY could be inferred, that of the gamete that led to FK, we modeled YY as a pseudo-diploid homozygous at all positions, using only the positions where the phased genotypes of FK and E allow for unambiguous deduction of the allele that was contributed to FK by YY (Fig. S12 and supplementary methods). The date associated with this individual was inferred from the radiocarbon dated individuals FK and E assuming they all died within the same time frame (see supplementary methods and Table S1).

A new genome from a BBC-associated individual in northern France is carrier of steppe ancestry

The identification of steppe ancestry in the FK individual and the scarcity of contemporaneous shotgun genomes from northern France (only a single BBC individual from Ciry-Salsogne « La Bouche à Vesle » was known genomically, CBV95 ((4) and Table S5) prompted us to improve the resolution of our analysis by adding another genome. We produced the genome of an individual from Saint-Martin-la-Garenne (SMGB54), west of Paris in the Yvelines, radiocarbon-dated to 2410-2129 cal BCE (97.8%) and representing a 30-49-year-old

man buried with BBC funerary rituals, in particular with a shale wrist-guard of BBC type (Fig. S4-5; Table S1). This individual is also a carrier of steppe ancestry (Fig. 2A, blue diamond) and was included in the further analyses.

SMGB54 is found in cluster “c” with FK and other BBC and Bronze age individuals from France as well as Bronze Age Hungarian and Croatian individuals. The pseudo-diploid genome of the reconstructed individual YY is found in cluster “d” (Supp Fig. S11) with individuals showing genomic profiles with a similarly high steppe ancestry proportion. It is remarkable that the two most contemporaneous individuals associated with the BBC/AOC in northwestern Europe, CBV95 from northeastern France and I5748 from the Netherlands, are clustering with YY.

Spatio-temporal high-resolution analysis of Late Neolithic genomes

Our fineSTRUCTURE analysis reveals similar relationships between individuals as seen with the PC1-PC2 plot, yet it provides better defined clusters with respect to geographical origins (Fig.2B and Fig.S13). The fineSTRUCTURE dendrogram separates first the individuals with a high proportion of steppe or HG ancestry from those with high ANF ancestry. Within the dendrogram’s first group, the next bifurcation separates the HG and steppe ancestry, and then, among the latter ones, the Yamnaya individuals are separated from the BBC, CWC, and Early Bronze age-associated Europeans with high steppe ancestry. Among this latter group, the final separation is shallow but notably distinguishes Central European from a subpart of more western individuals that are found in either of the final two subgroups. The reconstituted YY genome belongs to this western subgroup. Among the group with high ANF ancestry, the first bifurcation separates the individuals with steppe ancestry from those devoid of this ancestry. Among those with no steppe ancestry, the next bifurcation separates individuals according to their WHG ancestry: a first group comprises all initial ANF and later splits into two groups, a European group with little WHG ancestry and a second group with higher WHG ancestry. This latter then divides into three subgroups, one with individuals buried in Poland, one with individuals from France, mostly northern France, and Italy, and one with individuals from France, mostly southern France, and Iberia. Strikingly, the Bréviandes individuals devoid of steppe ancestry (namely the females A, B, C, D, E) are found in this latter subgroup consisting mostly of southern France/Iberian origin. The individuals with both significant WHG and

steppe ancestries, and dating to periods corresponding to the Late Neolithic and the Middle Bronze Age, separate first into two clusters discriminated by their extent of steppe ancestry: a first group with higher steppe ancestry later subdivided into a group originating from Croatia and Hungary, and a group mostly originating from France that includes individual FK. The second group carrying lower steppe ancestry can be subdivided into two groups, one including mostly Iberian individuals as well as individual HI and a group of more diverse geographic origins.

Since the hierarchical clustering approach of the fineSTRUCTURE analysis produced clusters that could be connected to their geographic origin once the major effect of the admixture of the three source populations has been taken into account, we explored whether this geographic signal can also be detected in the PCA. Indeed, PC3-PC4 reveal a correlation between haplotypes and geography mostly for samples dating to later periods, i.e., from the Late Neolithic to the Middle Bronze Age (Fig. 2C showing individuals colored according to geographic origin and Fig. S14-15). West to east and south to north clines are clearly visible, irrespective of the cultural association of individuals, indicating that a fine-scale genetic structure was established at continental level during this time period, possibly due to the accumulation of genetic drift in different parts of Europe. Strikingly, if the proportion of the three source ancestries account for most of the signal detected in PC1 and 2, these proportions appear to contribute only marginally to PC3 and 4 as there seems to be no correlation between location on the 2D plane and any of these proportions (Fig. S16). Late Neolithic individuals from southern France and Iberia group together and are clearly distinct from most samples from northern France with the notable exception of the five female Bréviandes individuals with no steppe ancestry, A, B, C, D and E (see Fig. S14-15). The BBC individual SMGB54 is also associated with the southern France cluster. This indicates that these individuals all have a stronger genetic link to southern than to local populations. In contrast, the reconstituted YY individual has a clear northern signature that is localized in PC3-4 at the outer boundary of the cluster constituted of individuals from northern France/Netherlands/British Isles/Czechia, referred to herein as the northern European cluster (Fig. S14-15). YY's son FK occupies an intermediate position due to the southern component brought about by his mother E, which puts him in the middle of the northern European cluster (Fig. 2C). Finally, his grandson HI is again located halfway between his father FK of northern European affinity and his mother A of southern affinity, which puts him in between the two clusters (Fig. S14-15).

The Bréviandes burial recapitulates the steppe ancestry admixture dynamics in Europe

To explore whether we could extend our findings to a larger number of individuals than those for whom complete genome data allowed inference of haplotypes, we used SNP-based approaches to incorporate the individuals that had been genotyped using sequence capture with the 1240k SNP array. We used qpAdm analysis (56) grouping samples based on their location and associated cultural complex or context dates (Fig.3A and Table S6-7). As previously reported (6), we observe a dilution of steppe-related ancestry proportions through a northeast to southwest cline. It is striking that the Bréviandes burial recapitulates the general continental admixture dynamics, in particular the dilution of steppe ancestry with ANF-WHG ancestry. To better pinpoint the date when this ancestry dilution took place, we plotted the percentage of steppe-ancestry as a function of calibrated radiocarbon age of the individuals and color-coded them according to their geographical location (Fig. 3B). We further highlighted the earliest individuals from France, Switzerland and the Netherlands to evaluate the first arrival of steppe-related ancestry into France. We observe a progressive decrease in steppe ancestry in Western Europe spreading over a 1000-year-period between ~3000 and ~2000 BCE starting from a high level of steppe ancestry to levels found in early CWC-associated individuals from Czechia in the 3000-2500 BCE period, followed by a slow decrease for the late CWC- and BBC-associated individuals from Central Europe, as previously reported (14). In the following 1000 years (2000-1000 BCE), the percentage of steppe ancestry appears to reach a broad equilibrium with regional differences in equilibrium values, the values being at least twice as high in the Netherlands and the British Isles than in Iberia (6). In France, we observe a rapid decrease in steppe ancestry within a few hundred years between 2500-2300 BCE approaching the steady-state value found later in southwestern Europe. This rapid decrease suggests that the dynamics of the admixture between populations with steppe ancestry and local Neolithic farmers differed in this area from the general trends observed in Central and northwestern Europe. Strikingly, the Bréviandes collective grave amounts to a real-time snapshot of this process capturing it in action at the moment it occurred. The reconstituted YY corresponds to the earliest documented arrival of steppe ancestry in France, between 2600 and 2500 BCE, followed by the nearly contemporaneous individual CBV95 from nearby Ciry-Salsogne in the Paris Basin. This ancestry is found later in BBC individuals I1389/I1390 (first degree relatives from Sierentz, Les Villas d'Aurele, Upper Rhine, Alsace and dated respectively to 2468-2278 and 2480-2210 cal BCE (6)) and GBVPK (Grotte Basse de la Vigne Perdue, southern France, dated to 2461-2299 cal BCE (15)), in eastern and southern France, respectively. However, earlier individuals from

these same regions do not carry this ancestry, whether they are associated with the BBC or not (BBC-associated individuals I1392 from Hegenheim, Alsace, eastern France, dated to 2833-2475 cal BCE (6), and PEI2 from the Dolmen des Peirières, Villedubert, Aude, southern France, dated to 2564-2302 cal BCE (4) and individuals not associated with BBC, GBVPL (15) and I16247 (41)).

To infer the admixture dates between European Late Neolithic and steppe-related groups, we used fastGLOBETROTTER relying on phased haplotypes (57), in contrast to DATES that uses only unphased data (58). Indeed, because of the use of unphased SNPs, DATES relies on the ancestry covariance pattern between the test genome and those of two source populations, and simulation reveals that it provides varying values in case of complex admixture scenarios involving unequal contributions of three ancestral groups (58). Accordingly, when evaluating the admixture between steppe ancestry and Late Neolithic farmers that are themselves admixed between ANFs and HGs, extremely different admixture dates are obtained using DATES for some test individuals depending on the individuals used to represent the Late Neolithic source population (Table S10). In contrast, the admixture dates estimated with fastGLOBETROTTER are much more resistant to the uncertainties of the true structure of the source population (Table S10). The fastGLOBETROTTER approach was applied on the admixed individuals, FK and SMGB54, as well as other admixed individuals dating to the period after 2500 cal BCE. We estimated the admixture dates of the incorporation of the first admixture event in terms of number of generations with a jack-knifing procedure to generate 95% CI. (see Methods section). We plotted the calibrated mean age of the samples and their inferred admixture dates assuming 28 years per generation. We detected two admixture pulses differing by ~500 years between the admixture dates of the two individuals FK and SMGB54 analyzed herein (Fig. 3C). These two pulses are also observed when analyzing 19 additional BBC and Bronze age individuals as visible on the density plot, which recapitulates all phased genomes analyzed (Fig. 3C, Table S8). Note that even though individual FK is an F1 descendant of the second admixture pulse the dating of this admixture is that of his father YY because no recombination between the steppe and Neolithic ancestry haploid genomes had yet occurred. Individuals associated with the BBC from southern France, the Netherlands and Czechia gave rise to initial admixture dates similar to that of FK approximately dating back to 2900 BCE (YY gave a similar date, Table S8). In contrast, more recent admixture dates, around 2500 BCE, were inferred for other individuals from France including SMGB54, but also from the British Isles and Iberia.

We further exploited the haplotypic coancestry matrix by calculating the distribution of haplotypic length and chunk donation by Yamnaya individuals in order to understand the effect on the steppe ancestry haplotypes of different mixing behaviors between admixing groups. Fig.3D represents the distribution of the genome-wide ratio of steppe haplotype lengths to chunk number where steppe-related ancestry proportions are represented by a color scale to highlight that there is no overall correlation between the chunk length/count ratio and the ancestry proportion. Individual HI whose genome experienced a rapid dilution of steppe haplotypes after two successive introgressions of Neolithic haplotypes, was observed as an outlier compared to the other individuals, and individual FK is also at the end of the overall distribution since both experienced a 4-fold and 2-fold reduction in chunk number, respectively. The small reduction in chunk length may be due to the small generation number insufficient for recombination to exert its effect on chunk size. This analysis, alongside the previously presented dating estimate of the admixture between individuals with steppe and Neolithic farmer ancestries suggests that most steppe ancestry individuals preferentially interbred with individuals of similar ancestries and only rarely mated with Neolithic farmers, thus preserving for many generations high levels of steppe ancestry. In contrast, when an individual with steppe ancestry integrated a Neolithic farmers' group, his/her steppe ancestry would have been quickly diluted in a few generations. Such an event can be detected clearly only in the first generations, as observed here in the collective burial of Bréviandes.

To detect any potential sex-bias in the spread of steppe ancestry we performed a ChromoPainter analysis on the imputed X chromosome dataset followed by a PCA analysis. Surprisingly, the X-chromosome diversity on PC2 of the source populations (HG, ANF, Yamnaya) is higher than for the entire genomes since the latter cluster more closely together (compare Fig. S17A and Fig. 2A). The PCA of X chromosome coancestry matrix was plotted with a heatmap of the autosomal steppe ancestry proportions (Fig. S17). This analysis shows that on PC2 the individuals with HG/ANF ancestry but lacking steppe-ancestry on the X-chromosome are the furthest away from the Yamnaya individuals, while individuals with the highest amount of steppe-ancestry on the X-chromosome, and therefore closest to the Yamnaya individuals, also have the highest amount of steppe-ancestry on the autosomes. Thus, for most individuals, there is no clear sex bias in the genetic flux between Late Neolithic and steppe ancestries. There are, however, exceptions to this trend. Indeed, some of the individuals with low-steppe ancestry on the X, which are thus on the left side of PC2 in Fig. S17, have medium and high steppe-ancestry on the autosomes, as shown by the yellow and green colors in the

heatmap. (Fig. S17B). These individuals are mostly from France, but some are also from the British Isles. This reveals a novel phase of female-driven incorporation of Late Neolithic ancestry into the steppe ancestry individuals when they spread into western Europe associated with the second admixture pulse.

Discussion

Prior to the characterization of the genetic diversity in Europe during the 3rd millennium BCE, archaeological evidence pointed to the coexistence in western and central Europe of regionally diverse cultures, mainly the Late Neolithic cultures, the CWC and the BBC (e.g., (17, 22)). Genomic data showed that the spread of the CWC and the BBC coincided with population movements. While the CWC was associated with the spread of steppe-ancestry, which was described as massive (3), the latter is characterized by a more complex pattern, which cannot easily be attributed to human movements and hitherto remained elusive (6).

Real-time observation of the arrival of steppe-ancestry in France through the genome analysis of an unusual Late Neolithic collective burial in the Paris Basin

Our high-resolution genomic analysis of the Late Neolithic collective burial BRE445 of Bréviandes in the southern of the Paris Basin, brought to light the mechanisms and direction of the admixture processes in western Europe between Late Neolithic farmers and descendants of nomads from the Pontic-Caspian steppe. This Late Neolithic burial dated to around 2500 cal BCE lacked associated grave-goods. It appeared to the archaeologists as a family burial, despite an obvious lack of adult men, since it consisted of three adult women and one adult man as well as one male and two female children and a female neonate. Combining genetic, anthropological and archaeological data, we can estimate that the burial was in use for up to 20 years (see Supplementary Material), which suggests that the deaths were not caused by an epidemic. Furthermore, traces of violence were not found on the skeletons, making it unlikely that it was mass burial due to a catastrophic act of violence. The analysis of the genomes identified three kinship groups: (i) a group comprising three generations consisting of a grandmother E, her son FK, her grandson HI and his mother A, (ii) a group consisting of a mother B and her neonate

daughter D and (iii) an unrelated girl C. All female individuals of the three groups were carriers of Neolithic genomes characterized by a western hunter-gatherer proportion and a Neolithic proportion of Anatolian/Aegean origin. The genome coverage was high enough to impute and phase the haplotypes. Thus, we could use tools that maximize the higher precision inherent to the analysis of haplotypes and could detect signatures that distinguish Late Neolithic individuals according to their geographic origin. In our analysis the genomes of these Neolithic women show a closer affinity to Neolithic and Chalcolithic individuals from southern France and the Iberian Peninsula than to contemporary individuals from northern Europe suggesting that the women themselves, or their ancestors, came from southwestern Europe (Fig. 2C).

The $^{87}\text{Sr}/^{86}\text{Sr}$ isotopic signature of the BRE445 adult individuals enabled us to put forward a hypothesis about the history of the individuals that constituted the group buried at Bréviandes. Grandmother E might have grown up in a place with a slightly different $^{87}\text{Sr}/^{86}\text{Sr}$ isotopic signature than the one prevalent in Bréviandes. We cannot identify the exact area where she grew up, it could be anywhere in the blue area of Fig. S7, i.e., preferentially in northern or southeastern France) but when we superpose the $^{87}\text{Sr}/^{86}\text{Sr}$ isotopic map of France (Fig. S7) with the genomic PC3-PC4 plot correlated to the geographical map (Fig. 2C), we can propose that E came from an area south of the Paris Basin. This is true for individual B, too, although the childhood $^{87}\text{Sr}/^{86}\text{Sr}$ signature of B is slightly different from the one of E. According to the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio values, only E and B would have been non-locals while all the other individuals analyzed would have spent their lives in the area of Bréviandes. FK, the son of a man with steppe ancestry and E was born in Bréviandes and so was HI, the son FK had with the local Neolithic woman A. We could make use of the haplotype phases of FK and his mother E to reconstruct one haploid copy of the genome of the missing father of FK whom we have called YY. The pseudo-diploid genome of YY has a signature shared with individuals of steppe ancestry from northern Europe (Bohemia, Poland, Netherlands, northern France and the British Isles). FK as an offspring of individuals with these two distinct ancestries, therefore plots halfway between these two distinct groups. The admixed individual FK clusters with BBC individuals from France (SMGB54, analyzed here, and GBVPK, analyzed previously (15)), and with later Early and Middle Bronze age individuals from France. This indicates that the two-fold dilution of the typical northern/central European steppe ancestry with Late Neolithic south-western European ancestry observed in FK is akin to the admixture pattern found in BBC and Bronze age genomes in France. The further two-fold dilution following admixture with

Late Neolithic south-western European ancestry, as observed in HI, the son of FK and A, leads to an admixture combination typical of Bronze age individuals from the Iberian Peninsula.

On the basis of this pattern of genetic similarities between various Late Neolithic individuals, the kinship relationships between the BRE445 individuals as well as the $^{87}\text{Sr}/^{86}\text{Sr}$ isotopic signatures we can propose the following hypothetical narrative: we suggest that E moved to Bréviandes before or after having met YY, likely a member of a foreign group of different ethnicity related to the descendants of steppe nomads, where she gave birth to FK, the son of YY, who grew up in this Late Neolithic context. FK had a son, HI, with the Neolithic woman A. Father and son died at the same time, or HI shortly after his father. FK, A and B all died as middle-aged adults. The only older individual belonging to the first generation was E. Individual B who might also have come from another location, maybe as a child, settled at Bréviandes and had a child with a Neolithic man not present in the burial. Taking all these data together, the picture emerges of a group whose members were buried together because they had either kinship or social links. In contrast, in a nearby but partially destroyed, contemporary single burial of a ~30-year-old woman (BRE444) we did not detect first or second degree of kinship with any BRE445 individual (data not shown). This means that the BRE445 individuals were buried by someone aware of these relationships. The male adult individual, FK, excludes the possibility of an all-female burial, but the fathers of the other individuals were not included in this grave. The physical appearance of the individuals did not set them apart from other Neolithic individuals (e.g., (33, 34, 59)). Indeed, they all had genetic signatures suggesting brown eyes, dark skin, except for B and D who likely had a lighter skin color, and dark hair except for FK and E whose hair could have been a chestnut color.

The genomes of the Late Neolithic burial at Bréviandes reveal the European-wide sex-biased admixture process involving steppe ancestry

While most details of these individuals' life history of these individuals cannot be reconstructed, this peculiar burial reveals some aspects of processes involved in the arrival of steppe ancestry on the territory of present-day France. Grandfather YY was the carrier of the steppe-ancestry proportion that we found in the genomes of his son FK and grandson HI (Fig. 2A). While the steppe-ancestry proportion in the reconstructed pseudodiploid genome of YY resembles the proportion in contemporary CWC individuals in Poland (13, 34) and Bohemia

(14), it is reduced (~35-37%) in his son FK to a proportion found in Late Neolithic and Bronze Age individuals from Germany, Switzerland and southern France (6, 15, 37, 44). In his grandson it is even further reduced to levels also found in Bronze Age Spain and Italy (6, 60, 61). Thus, within only two generations, incorporation of Neolithic ancestry into the admixed steppe-Neolithic gene pool led to the sharp decrease of the steppe-ancestry proportion to a level observed in individuals post-dating 2500 cal BCE. The speed of this process and its low frequency during what we identify as the initial phase of the steppe-ancestry expansion in western Europe might explain why the process had escaped observation so far: it needed to be captured in “real-time”, as was possible in BRE445. This burial is a lucky case of a snapshot of a mating event involving a steppe-ancestry individual who could be considered as a mobile early steppe-ancestry pioneer. He probably came to France from the Lower Rhine as suggested by his northern European genetic signature.

This result unveils the fundamentals of the early-phase admixture process between local Neolithic and migrating steppe-ancestry groups with which steppe-ancestry introgressed into the European Neolithic gene pool, and, more so, about the groups’ most likely mating behavior. Indeed, our family snapshot shows that admixture between steppe-ancestry carriers and Neolithic individuals led to stark dilution of the steppe-ancestry portion of the genome within two generations. For individuals to keep for several generations a substantial steppe-ancestry proportion in their genomes, such as the 70% or more of steppe ancestry found in the majority of Late Neolithic and Bronze Age individuals of eastern and Central Europe (3, 6, 13, 14, 37, 44, 60), they must have mated mostly between members of their group with similar ancestry and only sporadically with Neolithic farmers in agreement with the inferences based on modelling of the first series of captured genome data (62). We estimated that the admixture between individuals with steppe and Neolithic ancestry was not linear but rather followed a pattern characterized by a bimodal distribution thanks to the use of an admixture dating tool adapted to haplotypes, fastGLOBETROTTER (57), a procedure more robust than the SNP-based approach used by DATES (58) when applied to the diversity of samples analyzed here (Fig. 3C and Table S10). The first admixture pulse occurred around 2900 BCE (Fig. 3B and C), i.e., around the time when the first individuals with steppe ancestry could be identified in eastern and central Europe, (e.g., (3, 13))(14, 63)). This first admixture pulse between migrating individuals with steppe-ancestry and Late Neolithic farmers appeared sex-biased as individuals with higher steppe ancestry on the autosomes than on the X chromosome were detected in the early CWC (44). Then followed the period of relative stasis discussed above, when CWC-

associated individuals mated preferentially with individuals of similar dominant steppe ancestry, allowing a large proportion of individuals to remain around 70-80% steppe ancestry. This slightly admixed population pursued its migration to the north and west of Europe following this internal mating pattern and only occasionally mated with local Late Neolithic farmers, as detected for example in southern Poland (34).

In western Europe, around 2500 BCE, we detected a new pulse of admixture between Late Neolithic and steppe ancestry in the genome of SMGB54 from Saint-Martin-la-Garenne “Les Bretelles” associated with the BBC (2339-2139 cal BCE) harboring ~68% steppe-ancestry (Fig. 3A and B and Table S7), but also in later Bronze age individuals from France, the Iberian Peninsula and the British Isles (Fig. 3C). Using a haplotype-based approach we also detected differences in the extent of steppe ancestry on the chromosome X and the autosomes affecting individuals from France, but also from the British Isles (Fig. S7). This pattern reveals that there was a sex-bias in the admixture process associated with this second pulse, with an excess of males over females of steppe ancestry, the former relying more on local females of Neolithic ancestries to produce offspring. This is in agreement with the mitochondrial lineages found in the analyzed individuals that lack steppe haplotypes W and U2 (Tables S5-S6). Individual FK from Breviandes could be considered as the earliest representative of this second pulse of admixture detected as it occurred. Since FK is an F1 hybrid between a Neolithic female, E, and a man of steppe ancestry, YY, the haplotype block length-based dating of admixture by fastGLOBETROTTER detected only the previous admixture pulse of YY, because no meiosis-dependent recombination between the maternal and paternal genomes had yet occurred in FK.

The high steppe-ancestry portion, long steppe-ancestry haplotype chunks and an admixture date estimated at 2500 cal BCE of the genomes of SMGB54 points to a recent admixture event, about twelve generations before his birth, which could have occurred in northern France where individuals with northern steppe ancestry blended with individuals with southern Neolithic genomes without steppe-ancestry. Thus, both SMGB54 and the male individuals from the Breviandes grave are evidence for the second pulse of admixture between steppe ancestry men and Neolithic women. Since this second pulse is detected also in later Bronze age individuals from Western Europe, which, like SMGB54, did not show a marked further erosion of steppe ancestry, we can infer that following the second pulse of admixture with Neolithic women, most individuals resume the mating preference with steppe-ancestry carriers. This brief but rather general pattern of admixture with Neolithic women in Western Europe around 2500 BCE suggests a modification of the pattern or rhythm of the westward expansion of steppe-ancestry

carriers when reaching the Rhine valley. A slow and regular expansion would allow the maintenance of contacts between the pioneers and the neighboring original groups, thus permitting sufficient opportunities of mating with partners of similar steppe-ancestry. A similar pattern of exchanges of females of similar ancestry has been detected in the Lech Valley in Germany during the BBC and Bronze Age (44). This pattern may thus have been a general trend associated with steppe-ancestry groups already before the BBC, which could explain the relative stasis after the first pulse of admixture. Pulses of admixture suggest, however, rapid expansion of male-dominated pioneer groups moving in a single generation over distances too long to maintain enough “mating links” with their community of origin. This process would have led to more active recruitment of local women of Neolithic ancestry. Therefore, the two pulses would testify two more rapid expansion phases, the first one around 2900 BCE in Eastern Europe and the second around 2500 BCE in Western Europe. Strikingly, the second admixture pulse occurred at the time when the BBC spread throughout Europe associated with steppe-ancestry. It is tempting to speculate that these two events were associated, in particular in light of the proposal that the BBC originated from a Neolithic ancestry community in the Iberian Peninsula (18).

These different lines of evidence could be taken as an indication for the phase when male pioneers with steppe ancestry arrived in France from the north and admixed with local Neolithic women, just as we observed it at Bréviandes where the steppe-ancestry proportion of grandfather YY was quickly diluted in his grandson BRE445HI. This dilution of the steppe-ancestry in the Neolithic gene pool is in agreement with archaeological data that provide no evidence for a replacement of the local Neolithic cultures by the early phase BBC (22). Subsequently, during BBC phase 2 in France, a steady-state was reached and the steppe-ancestry was not further diluted in the Neolithic genomes, again suggesting an endogamic mode of reproduction of the steppe-ancestry carriers, comparable to the mode observed during the CWC in Central Europe.

We observed in real-time this pioneer mating pattern in the Bréviandes family grave that reveals solely male-driven introgression of the steppe-ancestry into a Neolithic population. Grandfather YY represents such a pioneer and is the first witness of the arrival of steppe-ancestry in the territory of present-day France. The traces left in the later steppe-ancestry descendants reveal that similar events happened elsewhere. During westward progression, the steppe-ancestry individuals in northern France reached a steady-state level of 50 to 60% Neolithic ancestry, as seen in SMGB54 and FK. Once reaching the British Isles, however, the

immigrants kept their initial steppe-ancestry proportion for about 500 years, arguing for the almost complete absence of admixture with local Neolithics ((6) and Fig 3B). Further incorporation of Neolithic ancestry into the steppe-ancestry background occurred during the southward migration toward the Iberian Peninsula leading to the dilution of the steppe-ancestry typical of these individuals who are characterized by only about 25% steppe-ancestry on average ((6) and Fig. 3B)

Blending of populations and associated cultures

The genomes of the two male individuals buried in single graves in the Paris Basin and post-dating 2,500 cal BCE, i.e., CBV95 and SMGB54, are respectively characterized by ~68% and ~44% steppe-ancestry. The cultural attributes of the former (notably a corded beaker and a Grand-Pressigny flint dagger (30)) associate him with the AOC/AOO that probably developed out of the CWC in the Lower Rhine area and spread southwards through the Rhine Valley (17), in agreement with his Neolithic genomic portion that relates him to the northern cluster. Conversely, although at present a speculation, the stone wrist-guard of the latter might be the attribute of an archer, either a hunter and/or a warrior or both. Moreover, SMGB54 was probably related to the BBC. This cultural attribute together with the clustering of his genome with those from the south could make him a witness of admixture between northern AOC pioneers moving southwards and mixing with Neolithic or Chalcolithic farmers from southern France or the Iberian Peninsula. Alternatively, the northern pioneers with steppe-ancestry moved southwards and admixed with northwards expanding southern BB users. Thus, SMGB54 could reflect movement along the Atlantic façade followed by inland incursions along rivers like the Seine. Other individuals from southern France, e.g., I1388, I12575, I3874, I3875 (6), PEI2 (4) and GBVPK (15), but also from the Upper Rhine Valley (Hegenheim, I1392 (6)) may be witnesses of the Rhône-Saône valley expansion. It is likely that the Late Neolithic populations of the north and the south of present-day France experienced the confrontation with different cultures through similar demographic mechanisms. The blending between the northern corded ware and the southern Bell Beaker users could have led to a hybrid culture, known as BBC, that rapidly expanded all over Europe, as proposed previously (16, 64).

Thus, the genome data from this study in the context of those published previously from other areas unveil how the BBC spread in western Europe. It is important to note that not all BBC-associated individuals in France were carriers of steppe-ancestry (e.g., I1392 (6), GBVPL

(15) and PEI2 (4). This hints to a continuous process of assimilation of the BBC socio-cultural model without population replacement by people with steppe ancestry. Rather, the adoption of this new culture seemed to be accompanied by local and additional intra-regional admixture, which followed a heterogenous pattern depending on the groups involved and led to both the disappearance of the Neolithic culture and the transformation of the Neolithic genome. Thus, the process of the introduction in western Europe of the steppe-ancestry and of the BBC resembles the interaction between Mesolithic hunter-gatherers and Neolithic farmers a few millennia earlier that led to the disappearance of the Mesolithic culture and the integration of hunter-gatherer genomes into the Neolithic gene pool (e.g., (3) (4, 13)).

Conclusion

To conclude, our study of an unusual burial dated to the Late Neolithic enables direct observation, almost in real-time, of the bimodal admixture process in Europe of dispersing people with steppe ancestry who disperse in Europe and local Late Neolithic groups or individuals. The generalization of the results obtained from our data suggests that this genomic transformation process took place during a period of profound cultural change. We propose that the CWC would have been developed after the initial encounter of migrating nomads from the steppes with local Late Neolithic populations in eastern and central Europe, as has been previously suggested (e.g., (16, 17)) and references therein). Subsequently, no further admixture with Late Neolithic individuals would have occurred. In the west, the propagation mode changed to leap-frog dispersal through “pioneers”, i.e., mobile individuals, during the early phase mostly men, who admixed with local Neolithic women, causing the dilution of their genomes with the Neolithic gene pool. Since we found no correlation between the spread of steppe-ancestry and the spread of the BBC during this early phase, the cultural spread seems to have been stimulated by people with steppe-ancestry only during later phases. After 2500 cal BCE, the descendants of the admixed groups moved southwards, mating with Late Neolithic individuals of France and the Iberian Peninsula, thus recruiting southern Late Neolithic ancestry into the admixed steppe-Neolithic ancestry. This process explains the establishment of a north-south steppe ancestry gradient, as it is still present today. The persistence of this pattern and the maintenance of the steppe-ancestry portion in the European genome up to the present day argues for a higher reproductive rate of individuals with steppe-ancestry in the later phase,

ultimately out-competing the Neolithic farmers. This process also reflects a shift to the profoundly different societal system observed in the Bronze Age (44) and eventually led to the replacement of the Neolithic genomes. Thus, this transformation process concerned both the cultures and the biology of the populations involved and led to the establishment of the western European genome that has prevailed up to now.

Material & Methods

Radiocarbon Dating and Strontium Isotope Analysis

The samples were AMS ^{14}C -dated at Tandem Laboratory, Uppsala University, Uppsala, Sweden. The $^{87}\text{Sr}/^{86}\text{Sr}$ -isotope analyses were performed by Paul Fullagar under the direction of T. Douglas Price, The Archaeological Chemistry, at the University of Chapel Hill, NC 27599, USA.

DNA extraction and Library Construction

Ancient DNA work was carried out in a high containment laboratory as described (65-67). The dense pyramidal parts of the petrous bones were isolated using a flame-sterilized diamond-disc of a Dremel. Bone pieces were then ground to fine powder in a 6775 Freezer/Mill®Spex SamplePrep in liquid nitrogen as described previously (4). The further protocol deviated from the previous one. Briefly, about 100 mg of bone powder were washed three times for 15 min. each with 1 ml 0.5 M sodium phosphate buffer at pH 8.0 Each sample was rinsed with 1 ml TT buffer made of 10 mM Tris HCl pH 8.0 and 0.1% Tween 20 One ml extraction buffer (0.5 M EDTA pH 8.0, 0.05% Tween 20, 250 µg/ml Proteinase K, and 0.14 M β-mercaptoethanol) was added to each sample and incubated on a rotating wheel for ~30 hours at 37°C. After this first incubation, the samples were centrifuged at 6000 × g for 1 minute. The supernatant was removed and kept for purification. Another 1 ml extraction buffer was added to the pellet and after resuspension, incubated for another 30 hours on a rotating wheel at 37°C. After this second incubation step, all samples were centrifuged at 6000×g for 1 minute. The supernatants from both incubation periods were pooled together in 2 ml tubes, mixed well and then centrifuged at maximum speed for 2 minutes. One ml of the supernatant was used for the following purification step using 5M40 binding buffer (5M Guanidine HCl, 40% isopropanol,

120 mM Sodium Acetate (NaAc) pH 5.2, 0.05% Tween 20; (68). 1 ml of each sample extract was purified with ~10 ml 5M40 binding buffer through Qiagen columns under vacuum, rinsed twice with 1 ml Buffer PE (Qiagen) and eluted with twice 30 µl of TET buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.05% Tween 20) following (69)

Libraries were constructed using 6 µl of extract and the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, Ipswich, MA, USA) after a pretreatment with USER enzyme mix (NEB, Ipswich, MA, USA) as described by the manufacturer. An unindexed Y-shaped adapter was ligated, the ligated libraries were first amplified for 8 cycles with the Illumina Rd1 and Rd2 sequencing primers, then 20% of the library was further amplified for 12 cycles with Illumina unique dual indexed adapter primers. The libraries were purified using 1.3 volume of SPRI beads (Macherey-Nagel™ NucleoMag™ NGS Clean-up and Size Select) for two successive rounds of purification.

Processing of Sequencing Data

All data processing steps were performed using shell instructions and in-house written scripts. The raw output of Illumina sequencers was in the BCL format. Because of the high number of in-house barcodes used, we performed the initial demultiplexing and fastq conversion using bcl2fastq (Illumina) on the laboratory server. Sequence quality was assessed using FastQC (Babraham bioinformatics - citation). Paired-end reads were merged and trimmed for adapters using leeHom (70). Reads larger than 28 bp were selected using awk and mapped onto the Human Reference Genome hs37d5 using BWA (Burrows-Wheeler Aligner) version 0.7.17 (71) with a seed length of 18. Aligned sequences in BAM format were processed with Samtools v.1.9. (72) to create index files and flagstat reports, and to select and sort mapped reads. Picard tools (<http://broadinstitute.github.io/picard/>) was used to remove duplicates. Multiple libraries were merged by Samtools v.1.9 before, followed by filtering the mapping quality (-q 20). Damage patterns were analyzed using mapDamage v.2.2.1 followed by end-clipping using BamUtil (<https://genome.sph.umich.edu/wiki/BamUtil:trimBam>). According to the damage pattern, three to four bases at the ends of each read were clipped for the UDG-treated libraries and 10 bases for non-UDG treated libraries.

Uniparental Haplogroups, Sex Identification and Contamination Estimation

The clipped bam files were used as inputs to run Haplogrep2, Haplocheck and HaploGrouper (73, 74) for the classification of mitochondrial haplogroups and the estimation of contamination level, respectively, which showed no detected contamination in the mtDNA data. To determine the sex of each individual, chromosome counts were performed using index statistics (idxstats) of Samtools on bam files by using exclusively the reads with mapping quality higher than 20 to prevent confusing effects of repetitive regions. Then the ratio of the proportion of read counts normalized to chromosome length for both the X and Y, normalized for each to the average of the autosomal proportion was used. We used the ANGSD software to estimate contamination on the basis of X chromosome for male individuals (75). HaploGrouper approach (76) was performed on the bam files of male individuals to classify Y-haplogroups (Table S1).

Genotype Calling and Dataset Preparation

We followed two different genotyping approaches, which are genotype imputation and pseudo-haploid calling, in order to prevent the possible false positive calls that could come from direct diploid genotyping due to low coverage of the majority of ancient samples. To generate an imputed and phased dataset, GLIMPSE (v.1.1.1) (77) was used for imputing the ancient genomes, namely 8 individuals from the current study and 111 publicly available ancient individuals. We followed the procedure described in the online Glimpse tutorial (https://odelaneau.github.io/GLIMPSE/tutorial_hg19.html), using Genotype Likelihoods (GL) computed with Bcftools (mpileup -I -E, call -Aim parameters) using the 1000 Genomes Project phase 3 reference panel excluding the indels. We filtered out from the imputed genomes the positions that have MAF lower than 5% in the 1000G reference panel. Then, we applied a further filtration of Genotype Posterior Probability lower than 0.99 and genotype missingness rate of 0 for downstream analyses. We use the autosomal sites in the 1240k SNP capture (38) to select 1240k SNPs that were used for calling, with Samtools mpileup and the pileupCaller pipeline (<https://github.com/stschiff/sequenceTools>), the bam files resulting from either shotgun sequencing or in-solution target capture.

Estimation of Kinship, Pigmentation Profile and Runs of Homozygosity (ROH)

We inferred the relatedness between individuals coming from the complex burial described in the current study using a combination of two complementary approaches, NgsRelate and READ softwares (45, 46). We used a mixed group of 56 ancient Western European individuals covering the periods of the Middle Neolithic, Late Neolithic, and Bronze Age to provide a distribution of ancestral proportions in the background allele frequency similar to that found in the complex burial. The first NgsRelate (v.2) analysis was performed using as input the genotype likelihood and allele frequency of these 56 individuals calculated from the bam files using ANGSD and the autosomes. To infer the pedigree, we performed a second NgsRelate analysis using ANGSD and the X chromosome data in the same manner as with the autosomes. A third NgsRelate analysis was performed using as input a vcf containing the imputed genotypes of these 56 individuals. The estimated kinship (θ) and Cotterman (k_0, k_1, k_2) coefficients from these three analyses were compared to the ideal values obtained using simulations by Yaka et al., 2021 (Table S2). READ was used to analyze pseudo-haploidized genotypes of individuals from the complex burial to further corroborate the NgsRelate-inferred relationships. We also calculated the normalized 1-P0 (mismatch rate) values from the output of Read as described (78) (Table S2). To establish the pedigree shown in Fig. 1, we used the kinship relationships determined by these analyses, the uniparental haplogroups (mitogenomes and Y chromosomes), the genetic sex and the estimation of the anthropological age at death.

In order to predict the pigmentation profile of the newly reported individuals, we used HIrisplex-S (79) and the imputed genotypes. We checked the HIrisplex-S output for consistency by manual verification of the genotypes and allele depths obtained without imputation. We applied HapROH (47) in order to identify ROH blocks longer than 4 cM on the Late Neolithic individuals from France.

Population Genetic Analyses

We ran ChromoPainter/fineSTRUCTURE (53) on the imputed and phased data using different sets of individuals so that no relatives are analyzed simultaneously. We used four datasets including all unrelated individuals with the related individuals grouped as follows: 1) FK-A-D; 2) E-A-B; 3) HI-B; 4) YY-A-D. After estimation of switch and mutation rates using 10 expectation-maximization (EM) iterations, each haplotype was painted using the others (-a 0 0). We applied fineSTRUCTURE's Markov chain Monte Carlo (MCMC) model on the

resulting coancestry matrices with 3,000,000 burn-in and 2,000,000 sampling iterations. After the extraction of the maximum observed posterior probability state and 100,000 hill-climbing iterations, the trees were inferred by using the maximum concordance state as it is used in the PoBI study (80). To obtain input files for SOURCEFIND v2 (55) and fastGLOBETROTTER (57) we ran ChromoPainter v2 with 10 iterations of EM to infer parameters on chromosomes 1, 8, 15 and 22. Then, the copy vector and painting samples files were generated for the 22 autosomes by using the inferred parameters and the ancestral groups, ANF, WHG, and Yamnaya (as described in the text), as donor populations to paint recipient individuals. We ran fastGLOBETROTTER to infer the admixture dates for each individual by ANF, WHG, and Yamnaya ancestral populations as donor populations while using European Late Neolithic (without any Yamnaya ancestry) and Yamnaya populations as surrogate populations. This was followed by a jackknife resampling step as implemented in the software in order to calculate the confidence intervals. These CIs were calculated using the delete- m_j jackknife variance estimator formula of Busing et al. (81).

qpAdm (56) was run using the Admixtools package (82) on the 1240k dataset for individuals coming from both shotgun sequencing and in-solution target capture data. We used Mbuti, Ust'-Ishim, Kostenki14, Sunghir3, Karelia, Motala12, and Iranian Neolithic genomes as the reference groups. Target individuals that could not be modeled as the product of a three-way admixture event were modeled as the product of a two-way admixture event.

Acknowledgments

We thank T. Douglas Price and Paul D. Fullagar for performing the strontium isotope analysis and Elisabeth Vauquelin for providing the iconographic representation of burial 445 of Bréviandes. The authors also thank all the staff from the Imagine genomic core facility for advice and technical assistance and Cécile Masson from the Bioinformatic Core Facility for the data transfer. We are grateful to Michael Ilett for critical reading of the manuscript and corrections of the English language. We thank Olivier Lemerrier for stimulating discussions.

Funding

We acknowledge support from the French national research center CNRS. The paleogenomic facility of the Institut Jacques Monod obtained support from the University Paris Diderot within the program “Actions de recherches structurantes”. The sequencing facility of the Institut Jacques Monod, Paris, is supported by grants from the University Paris Diderot, the Fondation pour la Recherche Médicale (DGE20111123014), and the Région Ile-de-France (11015901). We acknowledge support from the French national research center CNRS. EUR G.E.N.E. (ANR-17-EURE-0013 ; IdEx #ANR-18-IDEX-0001 l'Université de Paris ; Programme d'Investissements d'Avenir) for support of OP.

Author contributions

EMG and TG conceptualized and supervised the project; OP performed the laboratory work and the formal data analyses; OP, EMG and TG analyzed and interpreted the data; CL, CP, LB, CDJ and AC provided the archaeological samples; CL and CDJ provided the archaeological contexts; LB, AC, CP, IR and EV provided anthropological data; CB supervised the sequencing platform; EMG and TG wrote the paper with help from OP and input from CL.

Competing Interests. The authors declare that they have no competing financial interests that might have influenced the work described in this manuscript.

Data availability

Sequence data generated for this study are available from EBI European Nucleotide Archive PRJEBXXX.

Figures

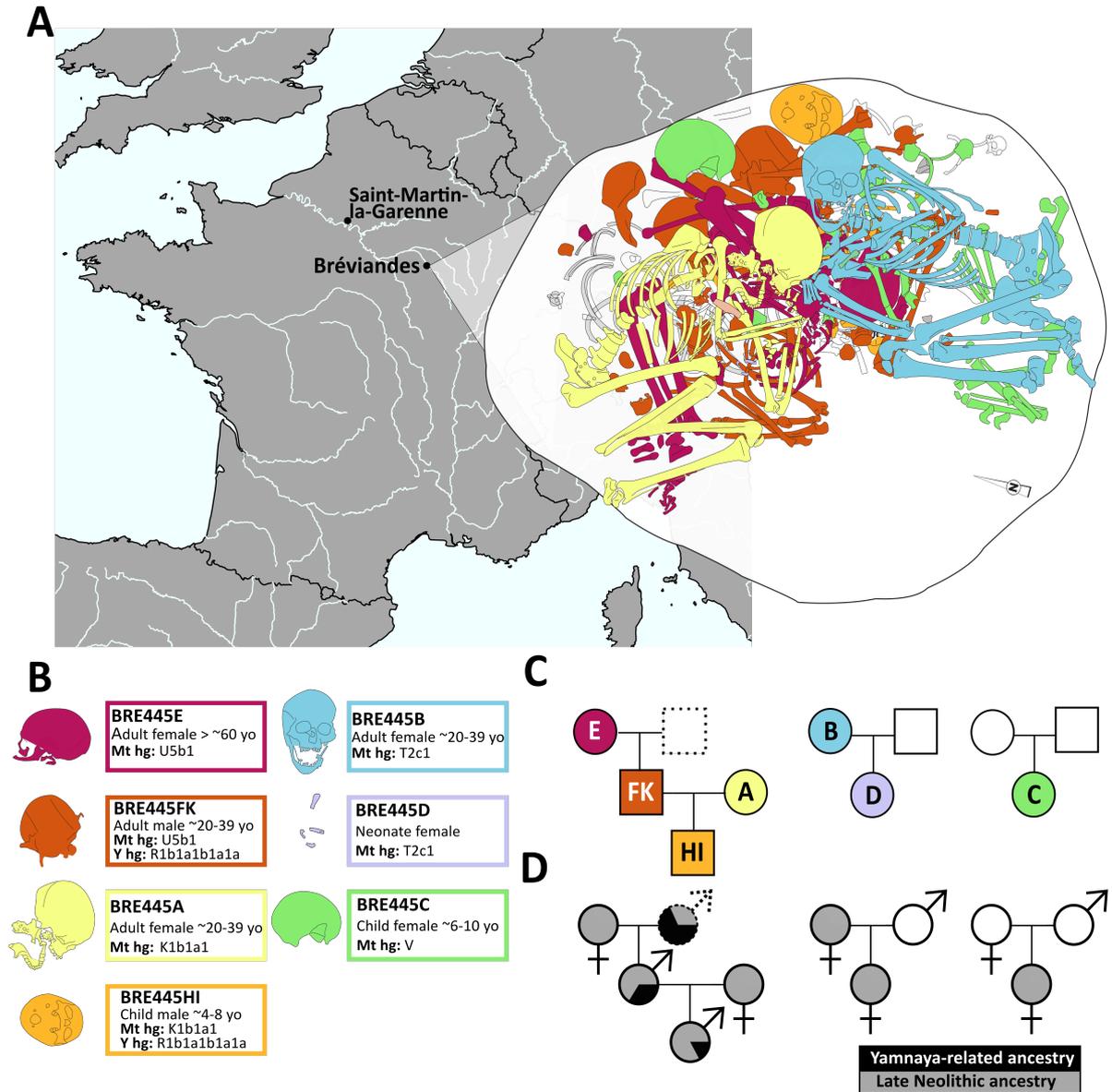


Fig. 1. A collective burial in the Paris Basin (A) Map showing geographical locations of samples analyzed in this study. Schematic representation of the collective burial at Bréviandes. (B) Estimated age of death and uniparental haplogroup information of the Bréviandes individuals. (C) Pedigree plots showing kin relationships between the Bréviandes individuals. (D) Pedigree plots showing the genomic component of Yamnaya-related steppe and Late Neolithic ancestries between the Bréviandes individuals. Empty boxes and circles indicate unsampled individuals. Dotted lines indicate the individual whose genome is reconstructed. Iconographic reconstruction of the burial: Elisabeth Vauquelin.

including other related Bréviandes individuals. See Figure S13 for the original dendrograms. **(C)** The PC3-PC4 plot of the same PCA (top) showing only European Late Neolithic and Bronze Age samples colored based on their geographical locations and the locations of these samples on the geographical map (bottom). The plot has been rotated clockwise to have a similar orientation as the geographical map. The variance explained by PC3 and PC4 are, respectively, 2,9% and 2,5%.

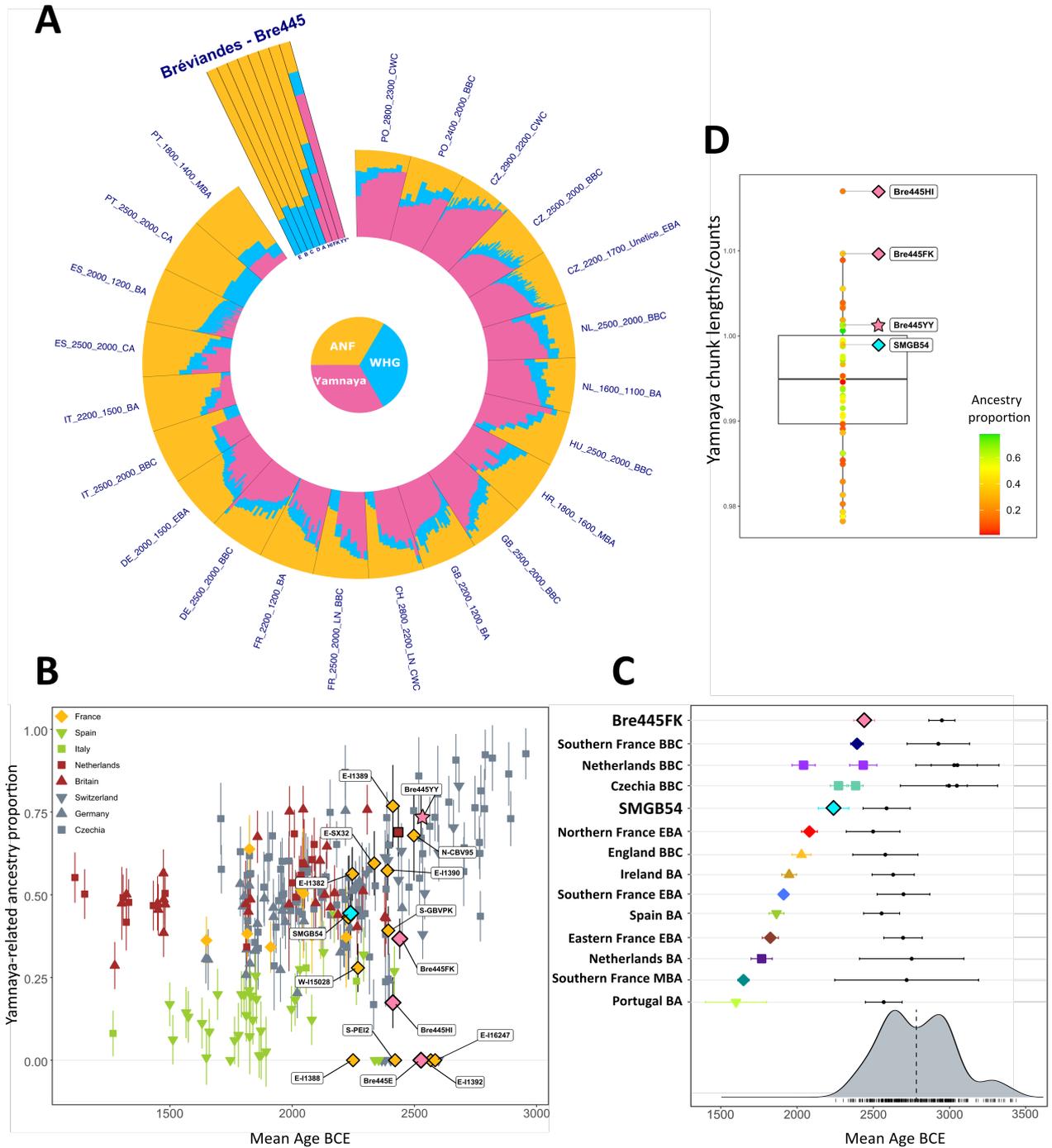


Figure 3: Spatiotemporal analysis of Yamnaya-related steppe ancestry (A) Circular graph display of the three-way qpAdm models for Late Neolithic and Bronze Age individuals grouped by location, archaeological context and dates. **(B)** Relative Yamnaya-related ancestry proportions plotted against mean radiocarbon dating values. Individuals of interest are highlighted with black outlines. The names of individuals from France are shown with boxes and indication of geographical regions in France. (E: East, W: West, S: South, N: North). For Bre445YY, the radiocarbon date of Bre445E is used. For Bre445HI, the radiocarbon date of Bre445FK minus one generation (28 years) is used. **(C)** Admixture dates and their 95% confidence interval of generation times (in black) inferred by fastGLOBETROTTER software and mean radiocarbon dating values of individuals (top). Density curve of inferred dates for each autosome of individuals calculated by the jackknife resampling step (bottom). **(D)** Box plot (n=51) of normalized Yamnaya-related haplotype chunk counts and chunk lengths ratios. Ancestry proportions calculated by SOURCEFIND of each sample are highlighted as a heatmap.

References

1. Q. Fu *et al.*, An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216-219 (2015).
2. Q. Fu *et al.*, The genetic history of Ice Age Europe. *Nature* **534**, 200-205 (2016).
3. W. Haak *et al.*, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207-211 (2015).
4. S. Brunel *et al.*, Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 12791-12798 (2020).
5. M. Rivollat *et al.*, Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers. *Science advances* **6**, eaaz5344 (2020).
6. I. Olalde *et al.*, The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* **555**, 190-196 (2018).
7. I. Lazaridis *et al.*, Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419-424 (2016).
8. A. Omrak *et al.*, Genomic Evidence Establishes Anatolia as the Source of the European Neolithic Gene Pool. *Current biology : CB* **26**, 270-275 (2016).
9. C. Gamba *et al.*, Genome flux and stasis in a five millennium transect of European prehistory. *Nature communications* **5**, 5257 (2014).
10. Z. Hofmanova *et al.*, Early farmers from across Europe directly descended from Neolithic Aegeans. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 6886-6891 (2016).
11. M. Lipson *et al.*, Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368-372 (2017).
12. I. Mathieson *et al.*, The genomic history of southeastern Europe. *Nature* **555**, 197-203 (2018).
13. D. M. Fernandes *et al.*, A genomic Neolithic time transect of hunter-farmer admixture in central Poland. *Scientific reports* **8**, 14879 (2018).
14. L. Papac *et al.*, Dynamic changes in genomic and social structures in third millennium BCE central Europe. *Science advances* **7**, (2021).
15. A. Seguin-Orlando *et al.*, Heterogeneous Hunter-Gatherer and Steppe-Related Ancestries in Late Neolithic and Bell Beaker Genomes from Present-Day France. *Current biology : CB* **31**, 1072-1083 e1010 (2021).
16. K. Kristiansen, in *The Oxford Handbook of Neolithic Europe*, C. Fowler, J. Harding, D. Hofmann, Eds. (Oxford University Press, 2014), chap. 56, pp. 1093-1118
17. V. Heyd, in *Yamnaya Interactions. The Yamnaya Impact of Prehistoric Europe 2*, V. Heyd, G. Kulcsár, B. Preda-Bălănică, Eds. (Archaeolingua, Helsinki, 2021), vol. 44, pp. 383-414.
18. J. L. Cardoso, Absolute chronology of the Beaker phenomenon north of the Tagus estuary: demographic and social implications. *Trab. Prehist.* **71**, 56-75 (2014).
19. J. Guilaine, La question campaniforme: sur quelques débats d'hier et d'aujourd'hui. *Estudos Arqueológicos de Oeiras* **25**, 9-46 (2019).
20. J. L. Cardoso, Os mais antigos vasos marítimos e sua difusão a partir do estuário do Tejo (Portugal). *Estudos Arqueológicos de Oeiras* **25**, 47-74 (2019).
21. L. Salanova, *Chronologie et facteurs d'évolution des sépultures individuelles campaniformes dans le Nord de la France*. L. Salanova, Y. Tchérémissinoff, Eds., Les sépultures individuelles campaniformes en France (Supplément à Gallia Préhistoire, CNRS, Paris, 2011), vol. 41.
22. O. Lemerrier, Think and Act. Local Data and Global Perspectives in Bell Beaker Archaeology. *Journal of Neolithic Archaeology* **8**, 77-96 (2018).
23. O. Lemerrier, in *XXVIIIe Congrès Préhistorique de France: La fin du Néolithique et la genèse du Bronze ancien dans l'Europe du nord-ouest*, N. Buchez, O. Lemerrier, I. Praud, M. Talon, Eds. (Société Préhistorique Française, Amiens, 2019), vol. session 5, pp. 239-250.

24. J. N. Lantin, J. D. Van der Waals, in *Glockenbechersymposium 1974*. (Bussum, Oberried, 1976), pp. 1-80.
25. M. Furholt, *Die absolutchronologische Datierung der Schnurkeramik in Mitteleuropa und Südkandinavien*. Universitätsforschungen zur Prähistorischen Archäologie (Dr. Rudolf Habelt GMBH Bonn, 2003).
26. E. Guerra Doce, C. Liesau von Lettow-Vorbeck, in *XVII UISPP world congress*, E. Guerra Doce, C. Liesau von Lettow-Vorbeck, Eds. (Archaeopress Archaeology, Burgos, 2016), vol. 6, pp. 141.
27. O. Lemerrier, in *XVIIIth UISPP World Congress*, T. Lachenal, R. Roure, O. Lemerrier, Eds. (Archaeopress, Paris, 2019), vol. Demography and Migration. Population Trajectories from the Neolithic to the Iron Age, pp. 116-140.
28. H. Fokkens, in *Background to Beakers. Inquiries into regional cultural backgrounds of the Bell Beaker complex*, H. Fokkens, F. Nicolis, Eds. (Sidestone Press, Leiden, 2012), pp. 9-35.
29. S. Chauvin *et al.*, Une sépulture campaniforme en plaine du Briennois (Blignicourt, Aube). *Internéo* **12**, 149-158 (2018).
30. L. Hachem *et al.*, in *Les sépultures individuelles campaniformes en France*, L. Salanova, Y. Tchéremissinoff, 978-2-271-07124-8. , Eds. (Gallia Préhistoire, Paris, 2011), vol. XLIIe Supplément, pp. 21-35.
31. E. Sangmeister, in *Oreluer Cikkijqe Atkabtuyqe*, P. R. Giot, Ed. (Laboratoire d'Anthropologie Préhistorique Rennes, Brest, 1961), vol. Les Civilisations Atlantiques du Néolithique à l'âge du Fer, pp. 25-56.
32. T. D. Price, G. Grupe, P. Schröter, Migration in the Bell Beaker period of central Europe. *Antiquity* **72** 405-411 (1998).
33. S. Brace *et al.*, Ancient genomes indicate population replacement in Early Neolithic Britain. *Nature ecology & evolution* **3**, 765-771 (2019).
34. A. Linderholm *et al.*, Corded Ware cultural complexity uncovered using genomic and isotopic analysis from south-eastern Poland. *Scientific reports* **10**, 6885 (2020).
35. A. Juras *et al.*, Mitochondrial genomes reveal an east to west cline of steppe ancestry in Corded Ware populations. *Scientific reports* **8**, 11603 (2018).
36. L. Saag *et al.*, Extensive Farming in Estonia Started through a Sex-Biased Migration from the Steppe. *Current biology : CB* **27**, 2185-2193 e2186 (2017).
37. A. Furtwängler *et al.*, Author Correction: Ancient genomes reveal social and genetic structure of Late Neolithic Switzerland. *Nature communications* **11**, 4759 (2020).
38. I. Mathieson *et al.*, Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499-503 (2015).
39. I. Olalde *et al.*, The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230-1234 (2019).
40. V. Villalba-Mouco *et al.*, Genomic transformation and social organization during the Copper Age-Bronze Age transition in southern Iberia. *Science advances* **7**, eabi7038 (2021).
41. N. Patterson *et al.*, Large-scale migration into Britain during the Middle to Late Bronze Age. *Nature* **601**, 588-594 (2022).
42. A. Goldberg, T. Gunther, N. A. Rosenberg, M. Jakobsson, Ancient X chromosomes reveal contrasting sex bias in Neolithic and Bronze Age Eurasian migrations. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 2657-2662 (2017).
43. I. Lazaridis, D. Reich, Failure to replicate a genetic signal for sex bias in the steppe migration into central Europe. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E3873-E3874 (2017).
44. A. Mittnik *et al.*, Kinship-based social inequality in Bronze Age Europe. *Science* **366**, 731-734 (2019).
45. K. Hanghoj, I. Moltke, P. A. Andersen, A. Manica, T. S. Korneliussen, Fast and accurate relatedness estimation from high-throughput sequencing data in the presence of inbreeding. *GigaScience* **8**, (2019).
46. J. M. Monroy Kuhn, M. Jakobsson, T. Gunther, Estimating genetic kin relationships in prehistoric populations. *PLoS One* **13**, e0195491 (2018).

47. H. Ringbauer, J. Novembre, M. Steinrücken, Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nature communications* **12**, 5425 (2021).
48. L. Chaitanya *et al.*, The HIrisPlex-S system for eye, hair and skin colour prediction from DNA: Introduction and forensic developmental validation. *Forensic science international. Genetics* **35**, 123-135 (2018).
49. B. L. Beard, C. M. Johnson, Strontium isotope composition of skeletal material can determine birth place and geographic mobility of humans and animals. *J. Forensic Sci.* **45**, 1049-1061 (2000).
50. T. D. Price, J. H. H. Burton, R. A. Bentley, The characterization of biologically available strontium isotope ratios for the study of prehistoric migration. *Archaeometry* **44**, 117-135 (2002).
51. T. D. Price, C. Knipper, G. Grupe, V. Smrčka, Strontium isotopes and prehistoric human migration: the bell beaker period in Central Europe. *Eur. J. Archaeol.* **7**, 9-40 (2004).
52. M. Willmes *et al.*, Mapping of bioavailable strontium isotope ratios in France for archaeological provenance studies. *Applied Geochemistry* **90**, 75-86 (2018).
53. D. J. Lawson, G. Hellenthal, S. Myers, D. Falush, Inference of population structure using dense haplotype data. *PLoS genetics* **8**, e1002453 (2012).
54. E. R. Jones *et al.*, Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature communications* **6**, 8912 (2015).
55. J. C. Chàcon-Duque *et al.*, Latin Americans show wide-spread Converso ancestry and imprint of local Native ancestry on physical appearance. *Nature communications* **9**, 5388 (2018).
56. E. Harney, N. Patterson, D. Reich, J. Wakeley, Assessing the performance of qpAdm: a statistical tool for studying population admixture. *Genetics* **217**, (2021).
57. P. Wangkumhang, M. Greenfield, G. Hellenthal, An efficient method to identify, date and describe admixture events using haplotype information. *bioRxiv* 2021.2008.2012.455263 (2021).
58. M. Chintalapati, N. Patterson, P. Moorjani, The spatiotemporal patterns of major human admixture events during the European Holocene. *Elife* **11**, e77625 (2022).
59. N. Marchi *et al.*, The genomic origins of the world's first farmers. *Cell* **185**, 1842-1859 e1818 (2022).
60. D. M. Fernandes *et al.*, The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. *Nature ecology & evolution* **4**, 334-345 (2020).
61. T. Saupe *et al.*, Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula. *Current biology : CB* **31**, 2576-2591 e2512 (2021).
62. J. Rio, C. S. Quilodran, M. Currat, Spatially explicit paleogenomic simulations support cohabitation with limited admixture between Bronze Age Central European populations. *Communications biology* **4**, 1163 (2021).
63. I. Lazaridis *et al.*, The genetic history of the Southern Arc: A bridge between West Asia and Europe. *Science* **377**, eabm4247 (2022).
64. S. Needham, Transforming Beaker culture in north-west Europe: processes of fusion and fission. *Proceedings of the Prehistoric Society* **71**, 171-217 (2005).
65. E. A. Bennett *et al.*, Morphology of the Denisovan phalanx closer to modern humans than to Neanderthals. *Science advances* **5**, eaaw3950 (2019).
66. E. A. Bennett *et al.*, The genetic identity of the earliest human-made hybrid animals, the kungas of Syro-Mesopotamia. *Science advances* **8**, eabm0218 (2022).
67. S. Guimaraes *et al.*, Ancient DNA shows domestic horses were introduced in the southern Caucasus and Anatolia during the Bronze Age. *Science advances* **6**, (2020).
68. J. Dabney *et al.*, Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 15758-15763 (2013).
69. N. Rohland, I. Glocke, A. Aximu-Petri, M. Meyer, Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nature protocols* **13**, 2447-2461 (2018).
70. G. Renaud, U. Stenzel, J. Kelso, leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic acids research* **42**, e141 (2014).

71. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
72. H. Li *et al.*, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
73. H. Weissensteiner *et al.*, Contamination detection in sequencing studies using the mitochondrial phylogeny. *Genome research*, (2021).
74. H. Weissensteiner *et al.*, HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic acids research* **44**, W58-63 (2016).
75. T. S. Korneliussen, A. Albrechtsen, R. Nielsen, ANGSD: Analysis of Next Generation Sequencing Data. *BMC bioinformatics* **15**, 356 (2014).
76. A. Jagadeesan *et al.*, HaploGrouper: a generalized approach to haplogroup classification. *Bioinformatics* **37**, 570–572 (2021).
77. S. Rubinacci, D. M. Ribeiro, R. J. Hofmeister, O. Delaneau, Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nature genetics* **53**, 120-126 (2021).
78. R. Yaka *et al.*, Variable kinship patterns in Neolithic Anatolia revealed by ancient genomes. *Current biology : CB* **31**, 2455-2468 e2418 (2021).
79. L. Chaitanya *et al.*, The HIrisPlex-S system for eye, hair and skin colour prediction from DNA: introduction and forensic developmental validation. *Forensic Sci. Int. Genet* **35**, 123–135 (2018).
80. S. Leslie *et al.*, The fine-scale genetic structure of the British population. *Nature* **519**, 309-314 (2015).
81. F. M. T. A. Busing, E. Meijer, R. V. D. Leeden, Delete-m Jackknife for Unequal m. . *Statistics and Computing* **9**, 3-8 (1999).
82. N. Patterson *et al.*, Ancient admixture in human history. *Genetics* **192**, 1065-1093 (2012).

Supplementary materials

Content

1. Descriptions of archaeological sites
 - 1.1. Bréviandes les Pointes & les Grèvottes [ZAC St Martin 1]
 - 1.2. Burial at Saint-Martin-la-Garenne
2. Stable Isotope analysis ($^{87/86}\text{Sr}$ ratio)

Additional Supplementary Figures

Figure S8. Calibrated (IntCal20) radiocarbon dating profile of individuals of interest for this study.

Figure S9. Runs of Homozygosity (ROH) values for blocks longer than 4 cM of Late Neolithic individuals in France

Figure S10. Three dimensional PC1-PC2-PC3 plots from different angles for the same dataset shown in Fig. 2A.

Figure S11. PC1-PC2 plots based on ChromoPainter autosomal coancestry matrix for datasets including different members of related Bréviandes individuals.

Figure S12. The schematic representation of building the pseudo-diploid genome for unsampled father of BRE445FK individual.

Figure S13. fineSTRUCTURE dendrograms based on ChromoPainter autosomal coancestry matrix for datasets including different members of related Bréviandes individuals.

Figure S14. PC3-PC4 plots based on autosomal ChromoPainter coancestry matrix for datasets including different members of related Bréviandes individuals.

Figure S15. Same PC3-PC4 plots shown in the Figure S14 highlighting geographical locations of Late Neolithic and Bronze Age individuals.

Figure S16. PC3-PC4 plots based on ChromoPainter X-chromosome coancestry matrix.

Figure S17. PC1-PC2 plots based on ChromoPainter X-chromosome coancestry matrix.

Supplementary Tables (excel file)

Table S1. Summary of archaeological and genetic information of the newly reported individuals

Table S2. Kinship estimation for the Bréviandes burial

Table S3-1. Phenotypic inference as deduced with the HirisPlex-S SNPs

Table S4. SOURCEFIND ancestry estimation for ancient individuals

Table S5. Summary of the information available for the previously published individuals used in the shotgun dataset

Table S6. Summary of the information available for the previously published individuals used in the 1240k dataset

Table S7. qpAdm admixture modelling for each individual

Table S8. fastGLOBETROTTER inference of Yamnaya and Late Neolithic surrogate groups for ancient individuals

Table S9. Summary of the Strontium isotopic ratios of the Bréviandes burial

Table S10. Comparison of the estimation of the number of generations since the admixture between Neolithic- and steppe-ancestry individuals using either fastGLOBETROTTER or DATES and varying source populations

1. Descriptions of archaeological sites

1.1. Bréviandes les Pointes & les Grèvottes [ZAC St Martin 1] (Aube, France)

Christophe Laurelut, Lola Bonnabel

The site, just in the south of the city of Troyes (Champagne, France) is located on an old terrace of the river Seine outside its alluvial plain. A first excavation took place in 2006 and mainly concerned a pioneer settlement of the middle/younger LBK (~5200 BCE). The site was then repeatedly used as a burial ground during the Middle Neolithic (Cerny culture, c. 4700-4500 BCE), the end of the Neolithic and the Early to Middle Bronze Age (c. 1750-1300 BCE). It was finally occupied again during the middle phase of the Late Bronze Age until the early Iron Age (Hallstatt A2-C, c. 1100-620 BCE).

The burials dated to the end of the Neolithic are concentrated in three clusters distributed on the edges of the LBK village. Their absolute chronology runs from the end of the 4th millennium (~3100-2900 BCE) to the Final Neolithic with an apparent chronological gap between features that can be attributed to its first part (~2900-2600 BCE) and the later features (~2600-2300 BCE). Thus, these clusters presumably do not belong to a single occupation; but they all contribute to the creation of a funerary landscape that seems to focus on the former LBK settlement area, over 2000 years after its abandonment.

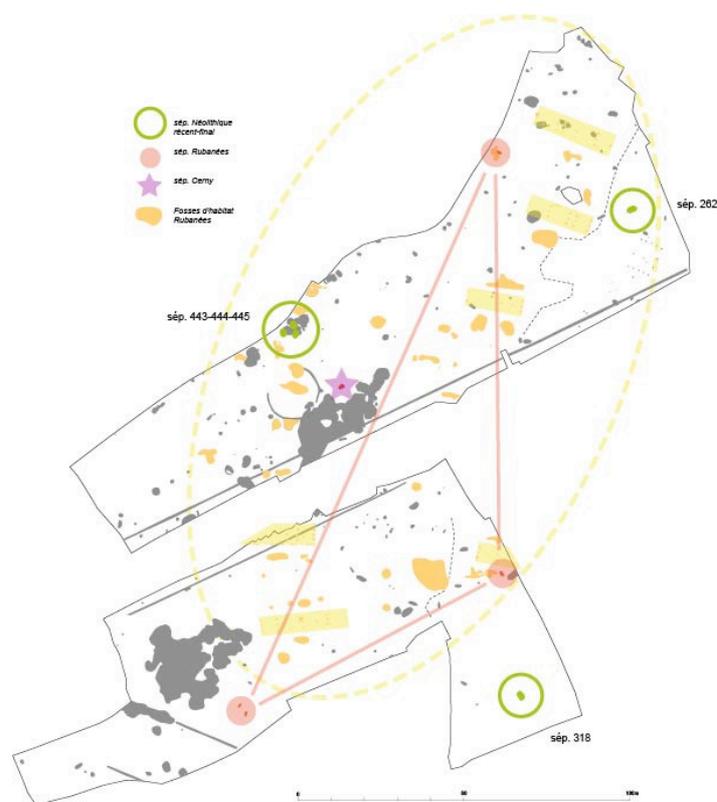


Figure S1: Plan of the site Bréviandes les Pointes & les Grèvottes [ZAC St Martin 1] with the burials of the different occupation phases highlighted.

Recut by a later (Iron Age) extraction pit, the western cluster remained unnoticed until the end of excavation. It includes three burials, two individual (burials BRE443 and BRE444) and one collective (BRE445). The morphology of these features was diversified, with each one corresponding to a specific model. Burials BRE444 and BRE445 were analyzed genetically in the present study and will be presented in more detail. They are the latest Neolithic burials dating to 2600-2300 BCE.

Burial BRE444



Burial BRE444 was partially destroyed by an Iron Age pit. The burial is of a young woman (<30 years old) who had been placed on her right side with her (missing) head oriented toward the north. The presence of empty spaces in the grave could be deduced from the morphology of the skeleton as well as the presence of a cover retaining the sediment. The burial is contemporary with the grave BRE445.

Figure S2 : Burial BRE444
(photo © Inrap).

Burial BRE445

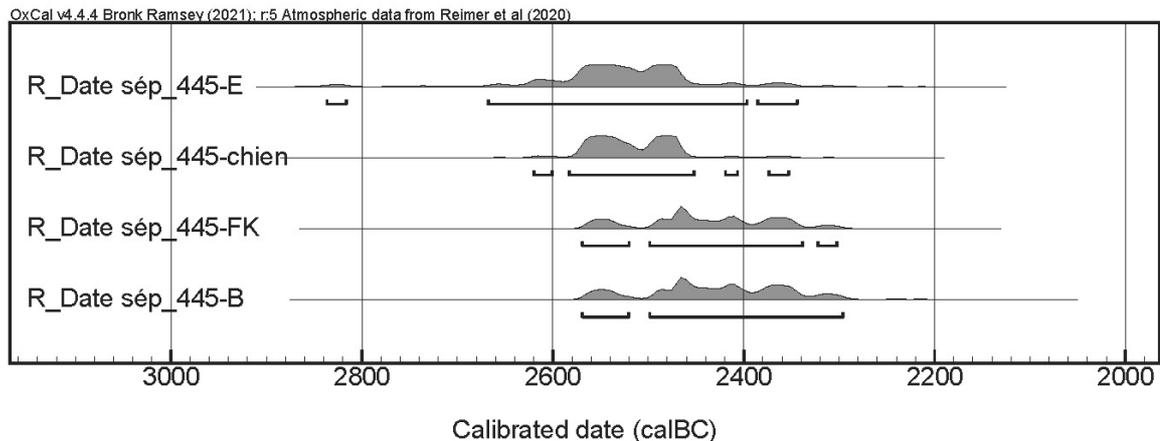
The Iron Age extraction pit cutting the burial destroyed the upper part of the grave and its access. Around 1 m below the excavation surface, a ~30 cm thick layer was all that remained from the subcircular 1.55m*1.25 m funerary chamber. The only associated grave goods were a circular bone bead found on the neck of individual A, and the paw of a young dog, dated to the same period as the human remains.

Seven individuals were buried in this grave: an adult woman A, lying on the right side in a bent position with her head toward the north; a young adult woman B, lying on the right side in a crouched position with her head toward the south; a 8 ± 2 year-old young child crouched on its right side C; a perinatal D; an elderly woman E, lying in crouched position on her right side, head toward the south; an adult man FK lying on his back slightly turned to the right side, head toward the north; a 6 ± 2 year-old child HI, crouched on its left side, head toward the west.



Figure S3: Collective burial BRE445. (photo © Inrap).

Radiocarbon dating of the individuals of burial BRE445: Table S1



Radiocarbon profiles (OxCalv4.4.4. Bronk Ramsey (2021); for the individuals of burial BRE445 (A to HI), and the dog from the same burial (BRE445_chien) calibrated using *inca* 20, (I).

The depositional order was partially reconstructed from the stratigraphic data:

1) Placed on the bottom of the pit, the young man FK, the child HI and the child C were the first to be deposited. There is no clear stratigraphic relationship between them that would indicate their order of deposition, but the spatial layout of the part of the skeletons not displaced by later burials shows that the man FK was deposited in the center of the grave, the child HI along his back and the child C in a more peripheral position to the south. This could suggest that the man was the first inhumated, but he may also have been at the center of a simultaneous double or triple burial.

2) The elderly woman E was also deposited in the center of the grave, above FK and HI whose bones were partially displaced /pushed away on this occasion to make room for the new burial.

3) Lastly, the young women A and B were deposited simultaneously or at a short interval, B above the woman E and A above the child C. The bones of the previously buried individuals were again partially moved to make room for the new incomers.

The remains of the perinatal D lay in an ambiguous position at the level of E's belly, suggesting her possible pregnancy, but this hypothesis did not fit with her estimated age. The association of D with B, whose arms seem to cross on D's remains, is more probable.

Individuals A and B were deposited on a 10 cm thick layer of sediment. A similar layer of sediment was also found in the upper part of the burial on top of their skeletons. The orientation of perinatal D could not be determined because the skeleton was incomplete.

Anthropological study of the individuals:

Cécile Paresys, Isabelle Richard

The anthropological study of the buried individuals suggests a family burial without exclusion of a particular sex or age class.

Based on the surface degeneracies of the sacroiliac pelvis, individual A was attributed to a 20 to 39-year-old woman. No infectious or traumatic lesion on the bones was diagnosed. The individual has only slight cribra orbitalia and osteoarthritis. The dentition is incomplete with ante mortem loss of molars, the teeth show caries and tartar.

The age of individual B was estimated at 20 – 30 years based on the absence of synostosis of the medial epiphyses of the clavicles and at 20-39 years based on the surface degeneracies of the sacroiliac pelvis. The individual has only slight cribra orbitalia. The dentition is complete with little caries and tartar and some signs of hypoplasia. The sex was determined at 100% to be female.

Individual C was a 8 + 2-year-old child as deduced from its dental mineralization stage, and 5 to 6.5 years old based on the metrical estimation of the length of the limbs (humerus : 192 mm ; radius : 141 mm ; ulna : 156 mm ; femur : 264 mm ; tibia : 216 mm). The individual shows a slight cribra orbitalia and some tartar on the teeth.

Only a few bone fragments are preserved from the perinatal individual D (fragments of the skull, the clavicle, the ribs, vertebrae, as well as humerus, ulna, femur and coxal fragments). The metrical estimation of its femur (71 mm) suggests, based on the method of Adalian, that it was between the 34,4 et 41,08th week of its intra-uterine life. The hypothesis was put forward that D was the child of E and that it was buried with its mother while still being in utero. This hypothesis was invalidated by the genetic results.

Individual E is represented by its skull, the mandible, a maxilla fragment, five teeth, and several post-cranial bone fragments. According to the sacroiliac pelvis surface degeneracies, the individual was more than 60 years old. The sex could not be clearly determined but was suspected to be female. No stress indicators could be identified, but signs of osteoarthritis in the upper part of the skeleton. No infectious or traumatic bone lesion was found. There are no indications for ante mortem tooth loss, but the teeth show caries and tartar.

The skeleton of individual HI is almost complete and well preserved. It is a 6 + 2-year-old child according to the mineralization stage of the teeth and 4.5 to 5.5 years old according to the metrical estimation of the limb bones (humerus : 168 mm ; radius : 132 mm ; femur : 232 mm ; tibia : 187 mm). The only stress indicator is a cribra orbitalia of stage 2. No infectious or traumatic bone lesion was found. The permanent molars do not show any traces of wear, but there are slight traces of wear on the temporary teeth. There are no traces of tartar or caries on the teeth.

The skeleton of individual FK is almost complete and well preserved. The absence of synostosis of the medial epiphyses of the clavicles suggests that the individual was 20 to 30 years old, and 20-39 years based on the surface degeneracies of the sacroiliac pelvis. The probabilistic sex diagnosis yielded a male at 97% confidence. The only stress indicator is a slight cranial hyperostosis. The 27 teeth show some tartar and traces of wear, but not ante mortem loss. The lower jaw has stage 2 periodontitis.

Reconstruction of the constitution of the collective burial BRE445 based on stratigraphic and genetic data

The stratigraphic succession observed in the collective burial establishes that the first individuals to be buried either simultaneously or shortly after each other were FK and HI, followed by E. C is at the bottom of the grave, but without any direct relation to FK, HI and E. The sedimentation between E and C and the following individuals B, D and A, maybe due to the first decay of the feature after some time of use, could suggest a longer interval between burial E/C and B/D/A. Individual A was probably the last individual inhumated, shortly after or simultaneously with B/D. As an alternative to a linear, successive single burials model, the data allow for three simultaneous inhumation phases comprising first FK, C and HI, then E, and finally B, D and A. Based on stratigraphic data alone it is not possible to decide between these two scenarios, or any other intermediate scenario.

The genealogical relationships deduced from the genetic results and the age differences between the related individuals E, A, FK and HI were used to reconstruct the burial's duration use. The first to be buried were generations 2 and 3 (FK and HI), followed by generation 1 (E). The last individuals to be buried belonged to generation 2 (A). Thus, while generation 2 was involved in the initial and the final phase of the burial, generation 1 was involved in the intermediate phase, and generation 3 in the initial phase.

The position of the remains of B strengthens the hypothesis that B died during delivery and that she was buried with her baby D in her arms.

There are two cases in this burial where the mother survives her child. This is the case for E and A whose sons, FK and HI, respectively, died before them. The total number of deceased infants (3) for four adults is high and, although statistically not significant, corresponds to an overrepresentation of this age class.

The age difference between the elderly woman (+60 years old at death) and her son, the man FK (20-30 years old at death) enables a probabilistic evaluation of the time between the two burials. If an early age of 15 years for E's motherhood is assumed, and an early death at 20 years for her son FK, the upper time span between the two burials could be 25 years [$60 - (15 + 20) = 25$]. If a relatively late age of 25 years for E's motherhood is assumed and a later death at 30 years for FK, the time between the two burials would be of only 5 years [$60 - (15 + 20) = 5$].

The same estimation can be made between female A (20-40 years old at death) and her son HI (6 ± 2 years old at death). There may be around 20 years between the two burials if A's motherhood is estimated at 15 years and her age at death at 40 years [$40 - (15 + 6) = 19 \pm 2$]. If A

gave birth to HI at an age of 20 years, the time difference would only be 5 years [30-(20+6)=5].

Although these estimations are based on different values (age differences at death of an elderly and an adult woman on one hand, and of an adult woman and a child on the other), the differences are comparable. While in the first case this difference concerns the first and second inhumation level, however, in the second case it concerns the whole burial (level 1 and 3). Therefore, the duration of the use of the burial is situated in a relatively reduced time span of one generation at most, but probably less.

1.2. Saint-Martin-la-Garenne

Cynthia Domenech-Jaulneau, Aloïs Corona

The individual primary burial at Saint-Martin-la-Garenne, Yvelines, “Les Bretelles” (SMGB54) was discovered at the bank of the Seine (2). The burial reflects the funerary ritual of the Bell Beaker Culture and dates to the extreme end of the Neolithic (~2200 BCE). AMS radiocarbon-dating of the SMGB burial yielded 3800 ± 30 BP, calibrated via IntCal2020 (1) to 2410-2129 cal BCE (97.8%).



The grave, dug into the subsoil, was probably accompanied by a structure on the surface since four concrete limestone blocks were placed on the northern side of the pit. It was probably located on the edge of a funerary area. Domestic occupations from the 4th millennium were discovered nearby.

Figure S4 : burial at Saint-Martin-la-Garenne

The anthropological analysis showed that the individual was a robust, healthy 30-49-year-old man without any visible bone pathology. He was deposited on his left side in a crouched position and with the head oriented to the north-east. The corpse was inhumated into the subsoil. The skeleton is not complete and although the bones are brittle but their morphological preservation is relatively good.



The colluvial deposits on top of the burial contained mixed finds dating from the Neolithic to modern times. The burial itself contained only one grave good, a shale wrist-guard characteristic of the Bell Beaker culture. It was found next to the elbow of the right arm. It is a polished, rectangular object with a perforation at each end. Nowadays, this kind of object is made of leather and used by archers to protect the internal side of the lower arm against the friction of the bow strings after the shot. Alternatively, it could have been used as sharpener for copper or bronze blades.

Figure S5 : The shale wrist-guard discovered from the burial at Saint-Martin-la-Garenne.

1.3. Ciry-Salsogne « La Bouche à Vesle » (Aisne, France)

Although no new genomic data from the burial at Ciry-Salsogne « La Bouche à Vesle » have been produced in the present study, the sequence data we had previously obtained (3) were reanalyzed in our present study and used in the comparative analyses showing that it is an important specimen. Therefore, it is described here.

CBV is an individual burial located in the Aisne-valley in northern France at Ciry-Salsogne « La Bouche à Vesle » (4). The skeleton is rather poorly preserved. The corpse was placed on its back and had not been wrapped in a shroud. It was radiocarbon dated to 3970 +/- 30BP (GrA-32767), 2575-2359 cal BCE using IntCal20 (1, 3).

The grave goods comprise a Bell Beaker vase, a flint flake, a deer antler and a flint dagger (4). The technique used to decorate the vessel is known from the Netherlands. This decoration style relates to the Dutch All-Over-Corded (AOC) group, characterized by individual burials and corded beakers that are different from classical Bell Beaker vessels (4). The dagger, made of flint from Grand-Pressigny (Centre-Val de Loire, southwestern part of the Paris Basin), is another feature that this burial shares in terms of raw material and style with burials in the north of Germany and the Netherlands, in association with vases of the style AOC and others (4). As in the case of similar burials in northern Germany and the Netherlands, no arrowheads were present. The deer antler is perforated in the center recalling arc-shaped Bell Beaker pendants (4).

To conclude, both structural elements of the CBV grave and the burial practice are comparable to graves in the Moselle region and in Luxembourg (4).

2. Stable Isotope analysis ($^{87/86}\text{Sr}$ ratio)

Eva-Maria Geigl

Life-time mobility of individuals across geologically distinct areas can be evaluated by determining the bioavailable $^{87}\text{Sr}/^{86}\text{Sr}$ ratio (5, 6). This determination is based on the strontium that had entered the skeletons through soils and the food chain after its release from eroding bedrock (7). Depending on the rubidium content of a rock and its age, the amount of ^{87}Sr – expressed as the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio – varies among geological units between about 0.7000 and 0.7500 and occasionally above. Since strontium in the diet replaces the calcium in biological apatite but is not metabolized, the measured $^{87/86}\text{Sr}$ ratio in skeletal tissues reflects the uptake during their formation (5). Therefore, changes in the food source can be reconstructed through the measurement of the $^{87/86}\text{Sr}$ ratio in the skeletal tissue (8). Owing to the existence of other, non-geological sources of Sr isotopes in the biosphere, one has to compare the isotopic signature obtained for an individual with the biologically-available signature in the area around the burial (6, 8). This necessitates the comparison with a baseline map of the strontium isotopic variation across a region. Such a terrestrial baseline map for archaeological provenance studies has been established for continental France based on the $^{87/86}\text{Sr}$ ratios of plants and soil leachates (9).

While the $^{87/86}\text{Sr}$ ratios in bone reflect the strontium ingested over the last one or two decades before death, the $^{87/86}\text{Sr}$ ratios obtained from dental enamel represent the childhood signal. Bone is remodeled throughout an individual's lifetime. During this lifelong process mature bone tissue is removed from the skeleton and new bone tissue is formed (10). Thus, the $^{87/86}\text{Sr}$ ratios reflect the bone metabolism during the last one or two decades before death and provide a summary measure of the local $^{87/86}\text{Sr}$ ratio (6, 8, 11). Conversely, the tooth enamel forms during tooth-specific childhood periods, is not remodeled afterwards and resists rather well post-mortem taphonomic alteration (6, 8, 11). As a consequence, the $^{87/86}\text{Sr}$ ratio in tooth enamel is a signal from the childhood of an individual and deviations of this ratio from the local baseline values can indicate that an individual was a first-generation migrant, more so if the $^{87/86}\text{Sr}$ ratio in long bone diaphysis corresponds to a local signal since the latter shifts toward local values over time (6, 8, 11).

The $^{87/86}\text{Sr}$ ratios of small animals testify for the local signature and were shown to be strongly correlated with those of indigenous prehistoric humans (6, 11). As a note of caution, one has to consider, however, that bone has a higher likelihood of containing diagenetic Sr isotopes than enamel.

To analyze life-time mobility patterns of the individuals of burial BRE445, their $^{87/86}\text{Sr}$ ratio was determined by T. Douglas Price, Archaeological Chemistry, Middleton WI, USA. Both teeth and bones were analyzed to detect differences between childhood and adulthood and thus differences in provenance. In particular, premolars were analyzed, which erupt between 9 and 13 years (12), and shaft fragments of femurs yielding a signal summarizing the last one or two decades of the life of the individual. As a local control, bones of a Neolithic cow and a Neolithic hare were subject to the same analysis. The hare is supposed to carry a local signature as it would not have a lifetime mobility.

Sample	species	tissue	skeletal element	$^{87}\text{Sr}/^{86}\text{Sr}$ ratio
BRE445A	Human	enamel	PM2	0,7081
BRE445A	Human	bone	shaft frag femur	0,7081
BRE445B	Human	enamel	PM1	0,708673
BRE445B	Human	bone	shaft frag	0,708165
BRE445E	Human	enamel	PM1	0,7088
BRE445E	Human	bone	shaft frag femur	0,7082
BRE445FK	Human	enamel	PM2	0,7081
BRE445FK	Human	bone	shaft frag femur	0,7081
BRE149t	hare	bone	shaft frag	0,7082
BRE45-2	cow	bone	shaft frag	0,7085

Table S9: $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of the analyzed samples

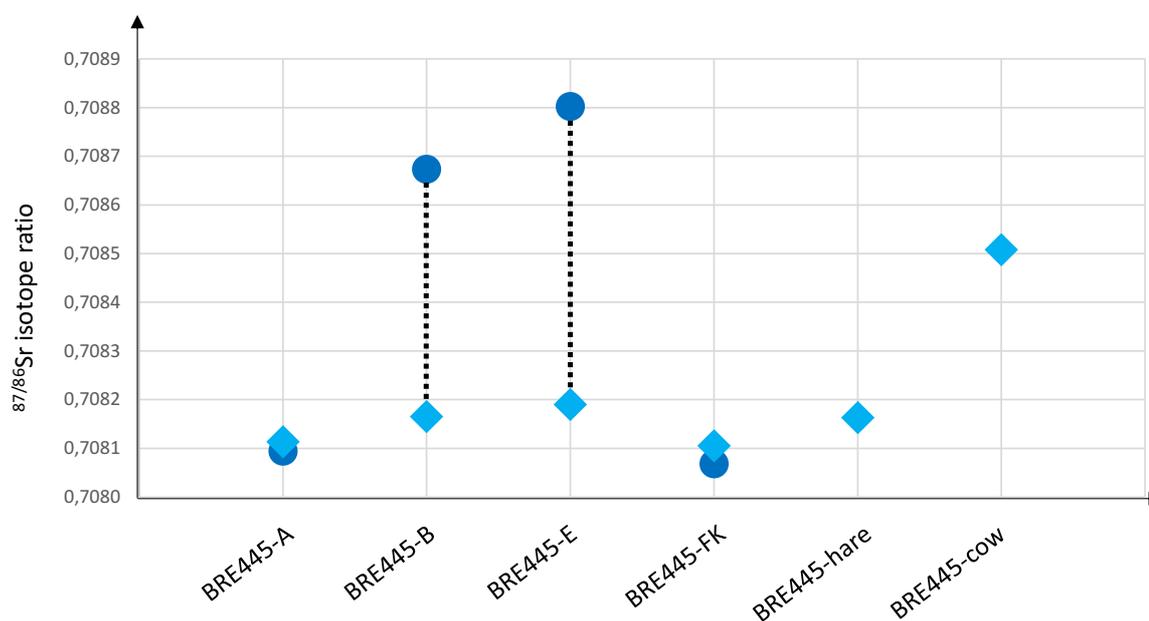


Figure S6: $^{87}/^{86}\text{Sr}$ ratio of the analyzed samples, the adult Bréviandes individuals and faunal samples; dark blue triangles: values obtained from teeth; light blue diamonds: values obtained from bones.

The $^{87/86}\text{Sr}$ values of both teeth and bones of individuals A and FK are in the same range. This argues for these individuals having spent their life in the surroundings of the burial, in agreement with the value of the hare representing a local, low-mobility control. The value for the domestic cow is different (higher) suggesting that the cow could have come from somewhere else. Since domestic animals such as cattle and pigs were mobile in these societies and circulated in social networks and were subject to transhumance, this is a plausible result. The value for the premolar of individual E, however, is higher by 0.0006 compared to the bone value. Since the teeth enamel forms during childhood and is not remodeled later this suggests that E spent her childhood in a location with a different geology than the one around Bréviandes les Pointes & les Grévottes (ZAC St Martin 1) meaning that she might originally not be from this area. She possibly spent her adult life in Bréviandes, at least after giving birth to FK who carries a local Sr isotopic signature, in agreement with the bone and teeth values of individual A and the hare. The difference between the tooth and the bone values of individual B is with 0.005 less pronounced than in E, but might hint to a non-local provenance, too.

We compared the obtained $^{87/86}\text{Sr}$ values to the baseline map of continental France (9). The site of Bréviandes is located in the Paris Basin that is characterized by a rather homogenous geology (Figure S7) (9). The site itself is located on an ancient terrace of the river Seine and the substrate is a relatively thin silt cover, probably of Weichselian origin, which is eroded and remodeled at the site. The isotope values of this location are in the range of 0.7076-0.7090, in agreement with the values obtained from the analysis of the human and faunal individuals of the site.

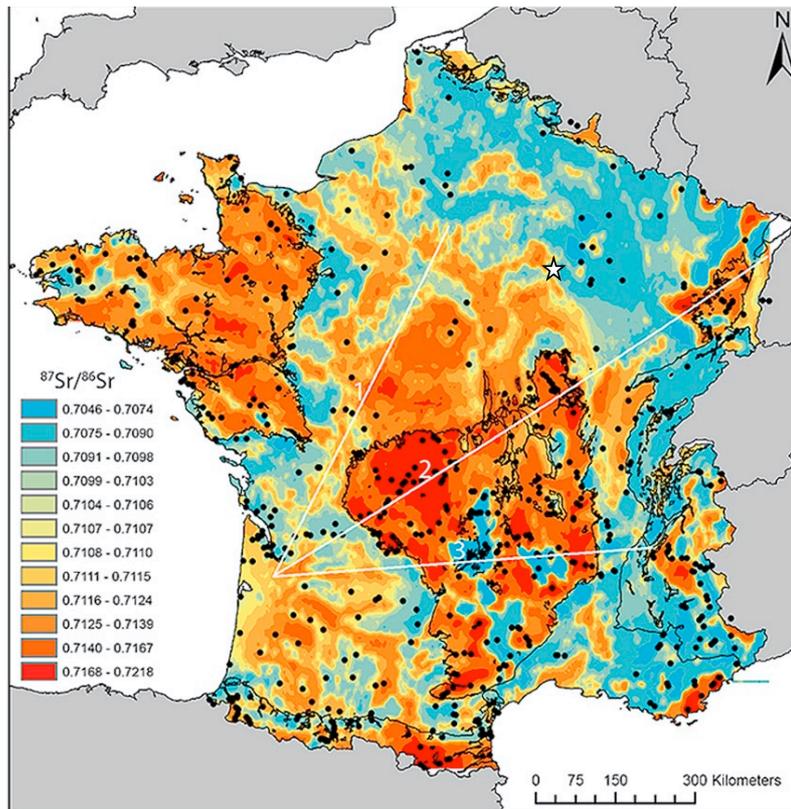


Figure S7: Surface geological map of France (BRGM: Geoscience for a sustainable earth; <https://www.brgm.fr/en/news/news/open-science-free-access-brgm-digital-geological-maps>) with sample sites from the IRHUM dataset marked as black dots (9). The site of Bréviandes (ZAC St Martin 1) is indicated as white star.

Conclusion

The $^{87}/^{86}\text{Sr}$ values obtained in the present study suggest that the life-time mobility of individuals E and B of BRE445 could have comprised an area that goes beyond the Paris Basin. These two individuals, especially individual E, may have spent their childhood either in the north or northeast of France, or in the east/southeast (Rhône valley) of France before arriving in the Paris Basin and Bréviandes.

Additional Supplementary Figures

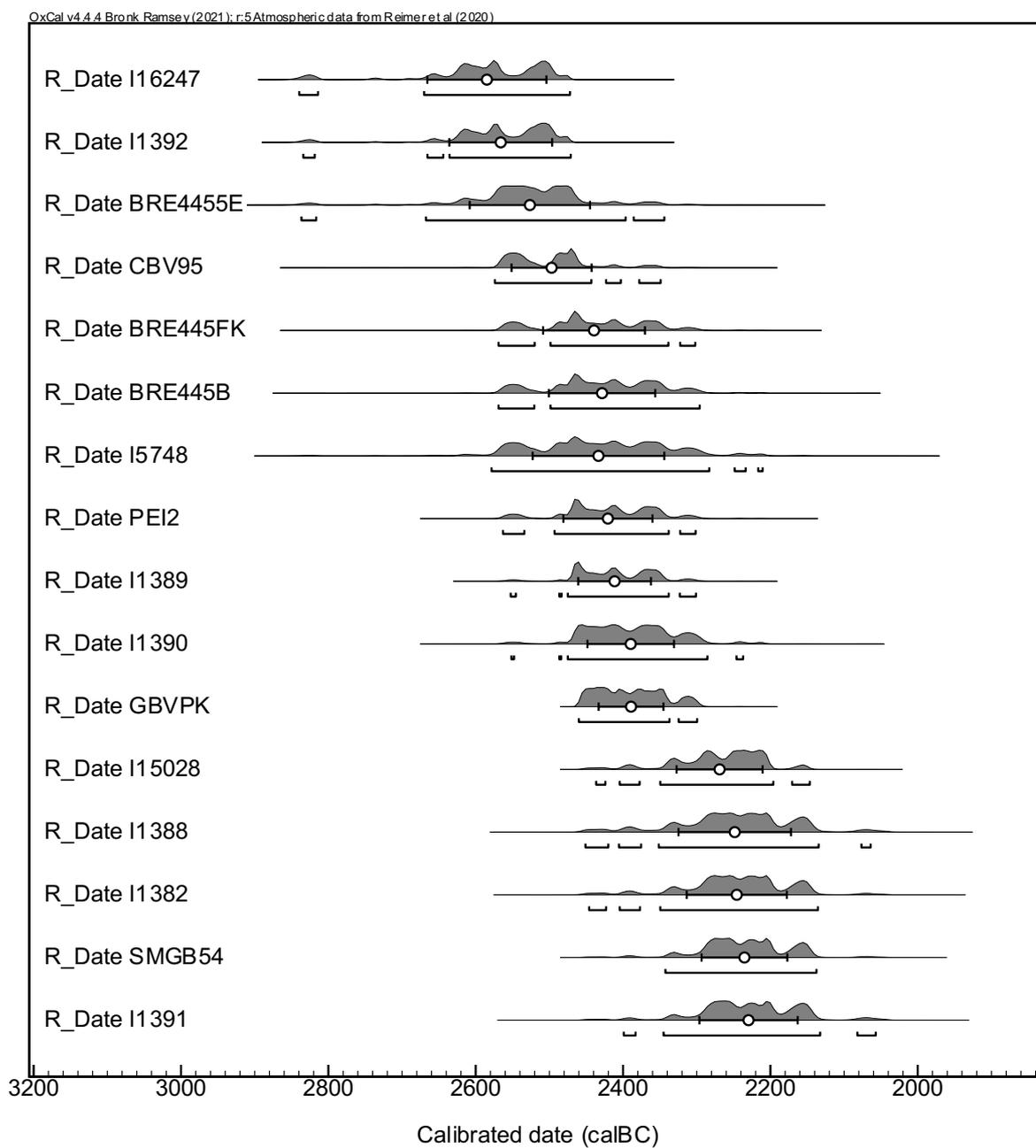


Figure S8. Calibrated (IntCal20; *I*) radiocarbon dating profile of individuals of interest for this study.

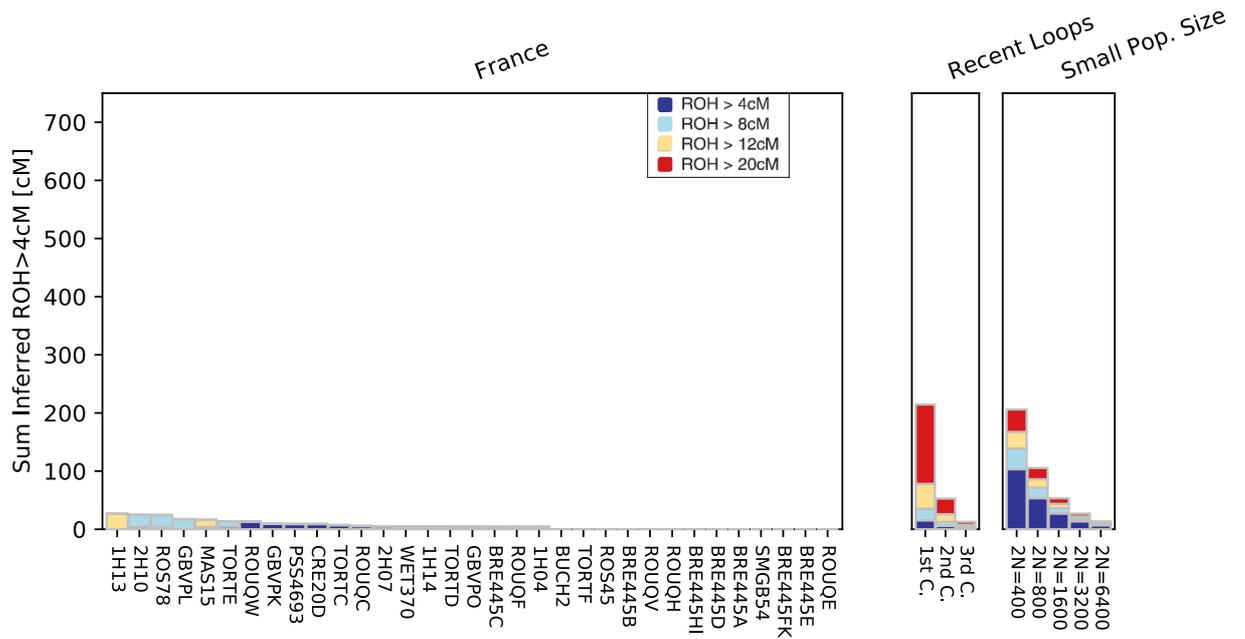


Figure S9. Runs of Homozygosity (ROH) values for blocks longer than 4 cM of Late Neolithic individuals in France

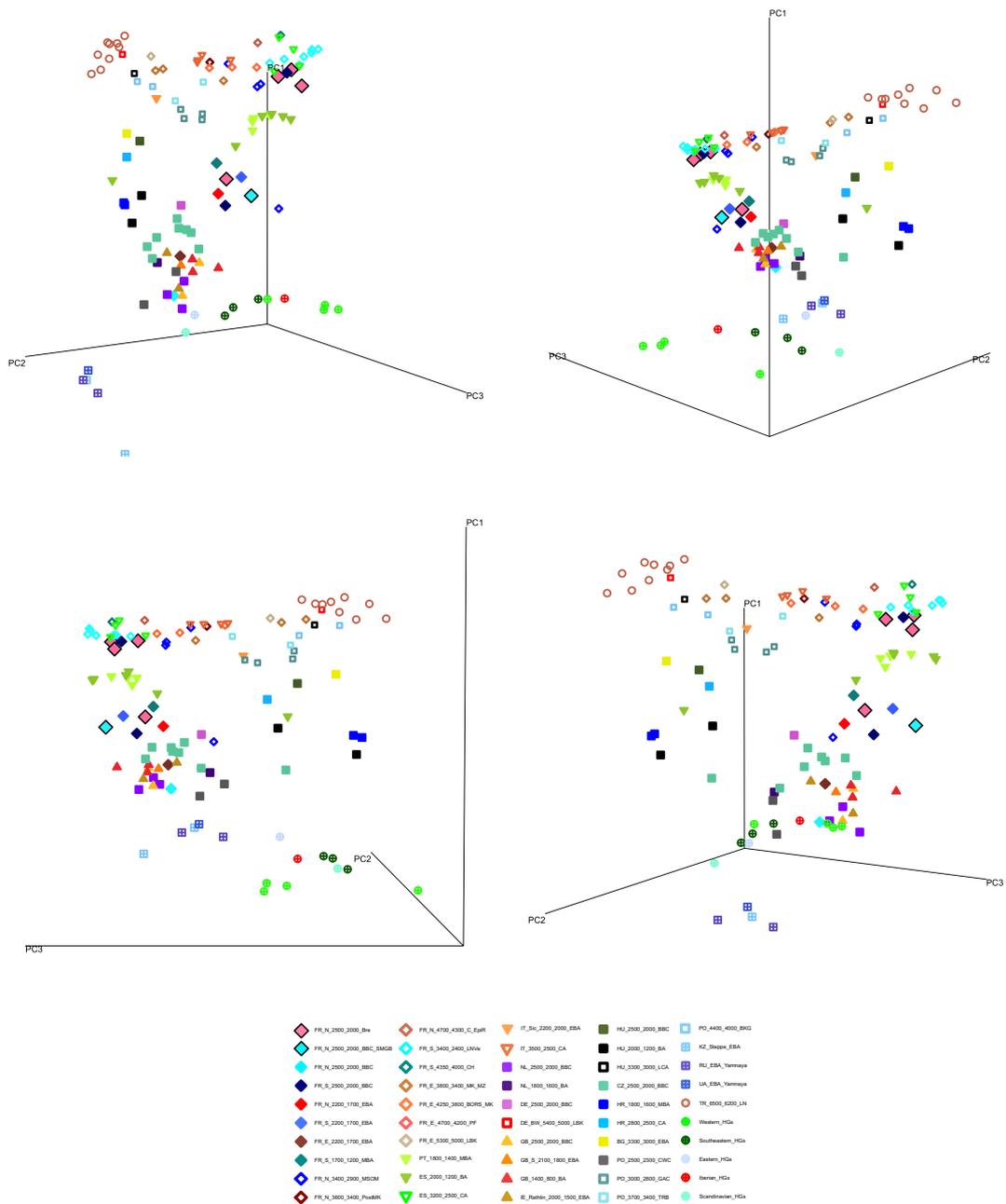


Figure S10. Three dimensional PC1-PC2-PC3 plots from different angles for the same dataset shown in Fig. 2A.

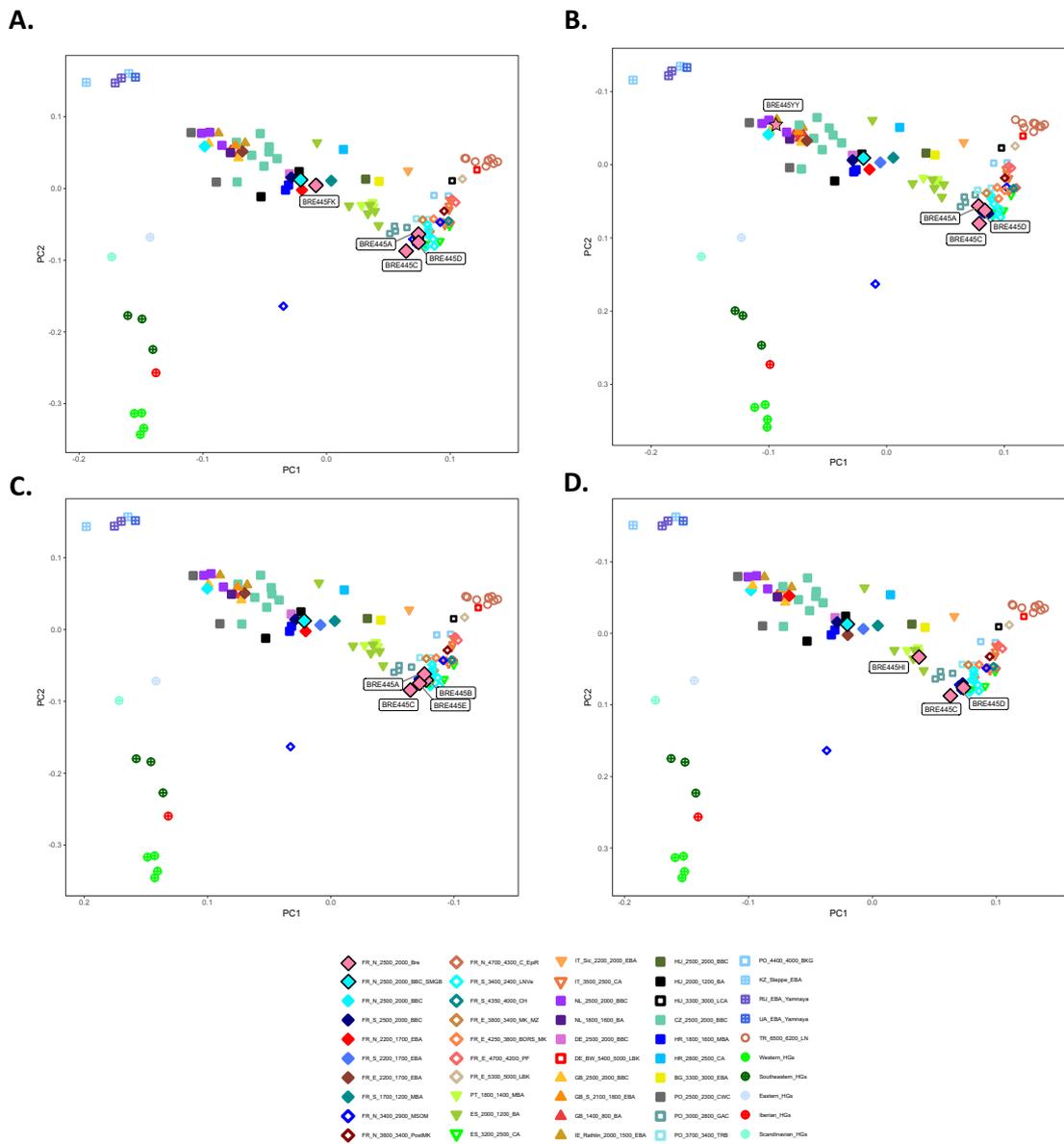


Figure S11. PC1-PC2 plots based on ChromoPainter autosomal coancestry matrix for datasets including different members of related Bréviandes individuals. Plot A is used in the Figure 2A.

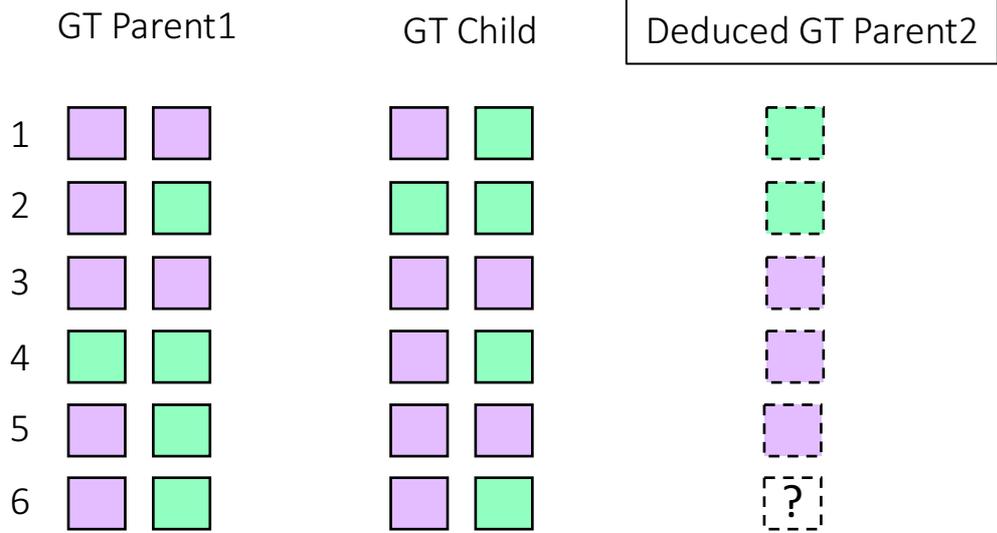
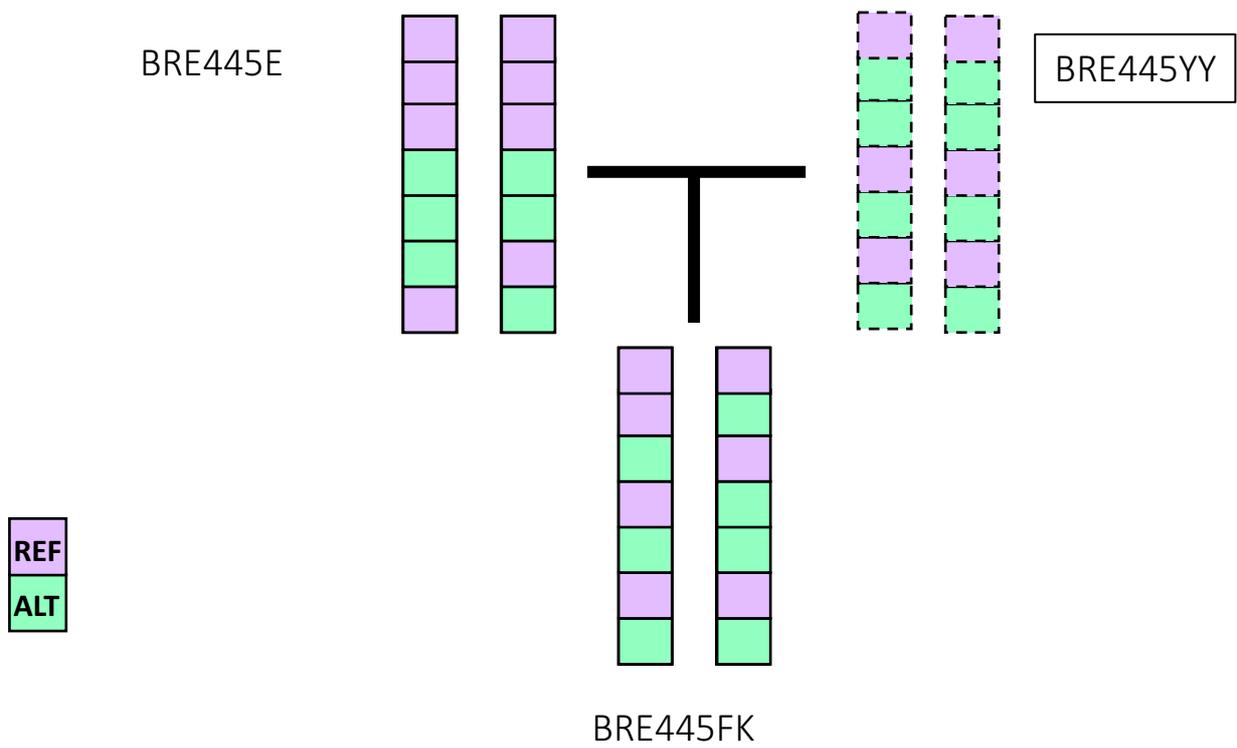
A**B**

Figure S12. The schematic representation of building the pseudo-diploid genome for unsampled father of BRE445FK individual.

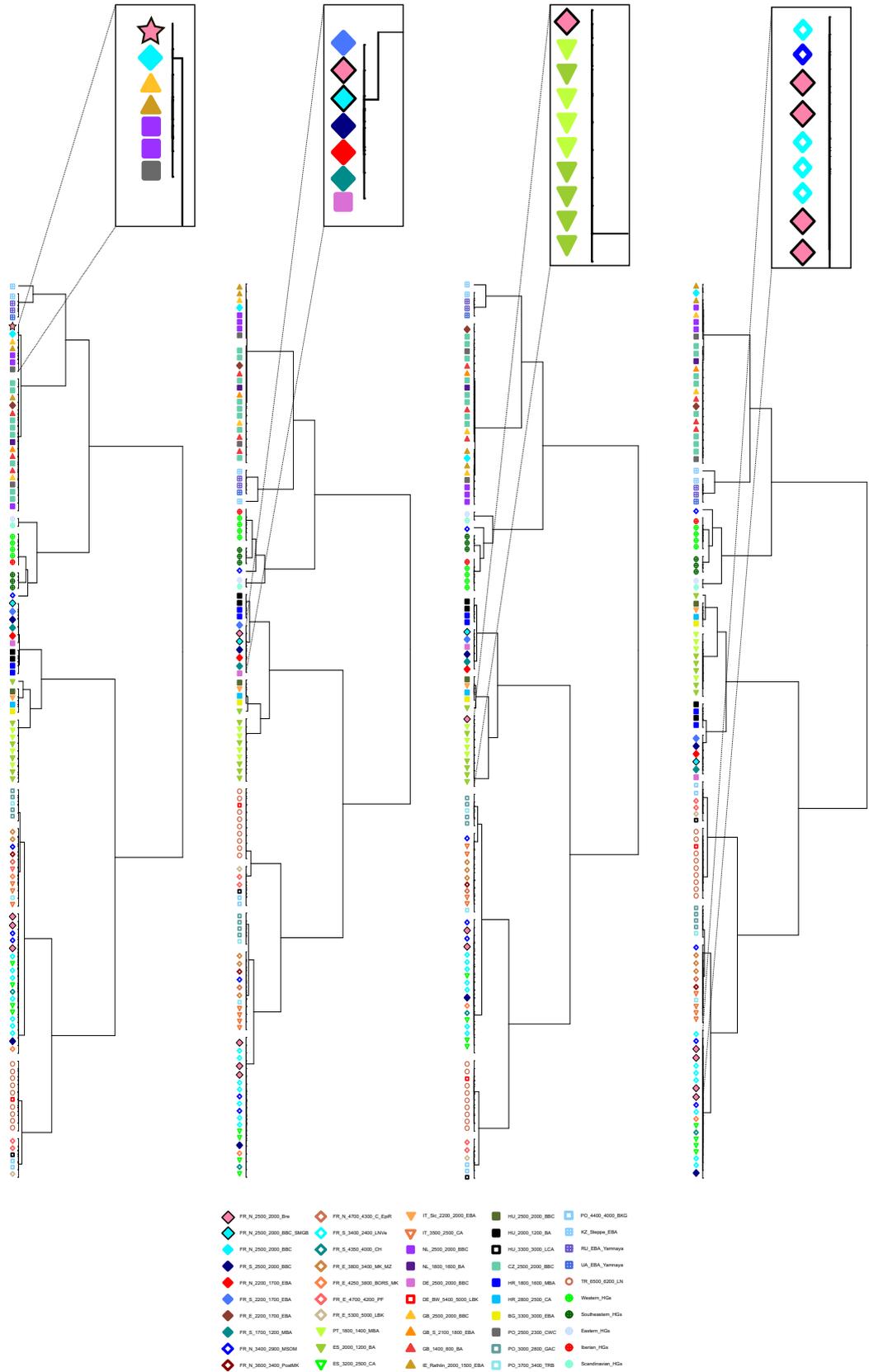


Figure S13. fineSTRUCTURE dendrograms based on ChromoPainter autosomal coancestry matrix for datasets including different members of related Bréviandes individuals. The four clusters highlighted (top) are used in the Figure 2B.

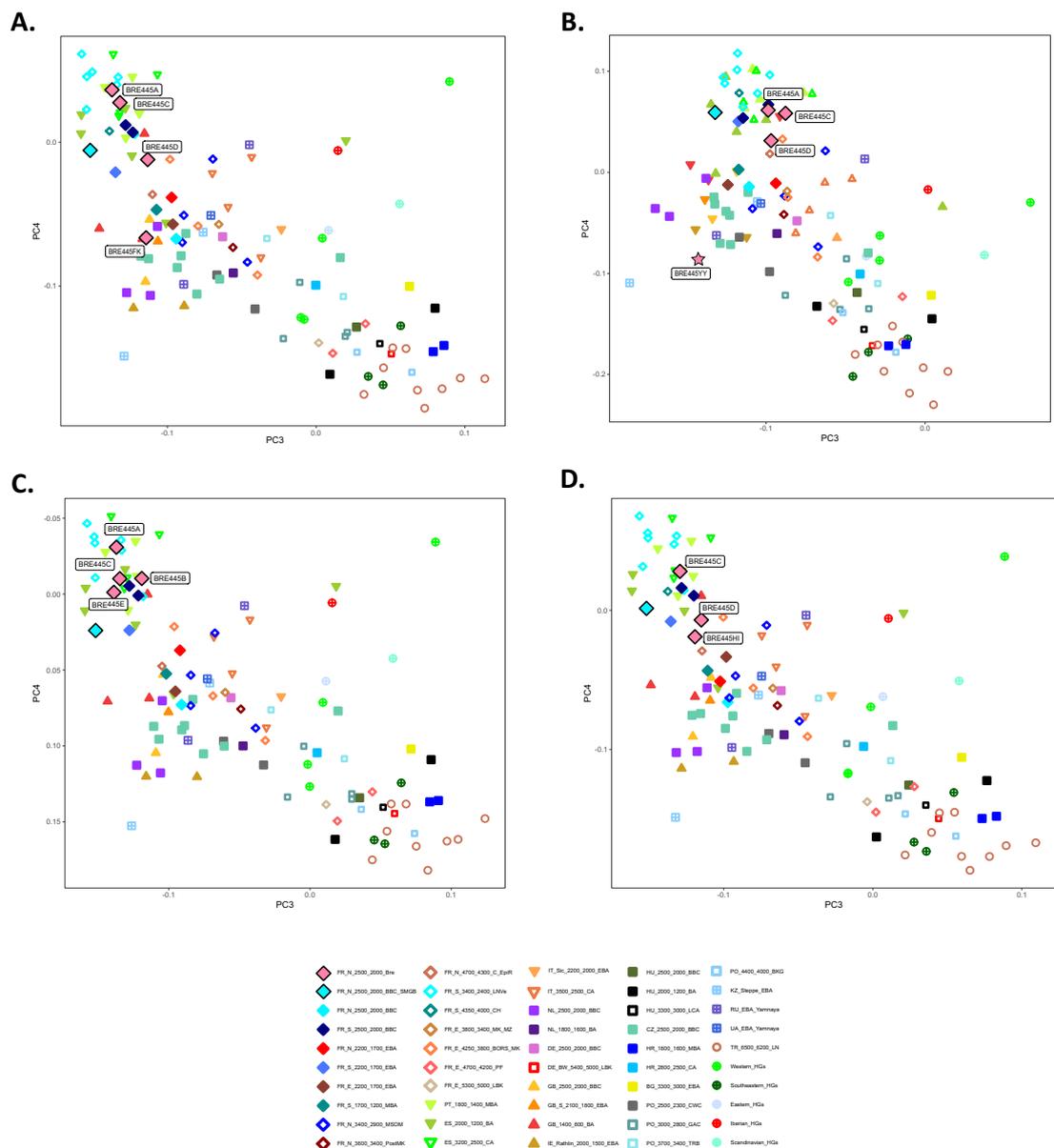


Figure S14. PC3-PC4 plots based on autosomal ChromoPainter coancestry matrix for datasets including different members of related Bréviandes individuals.

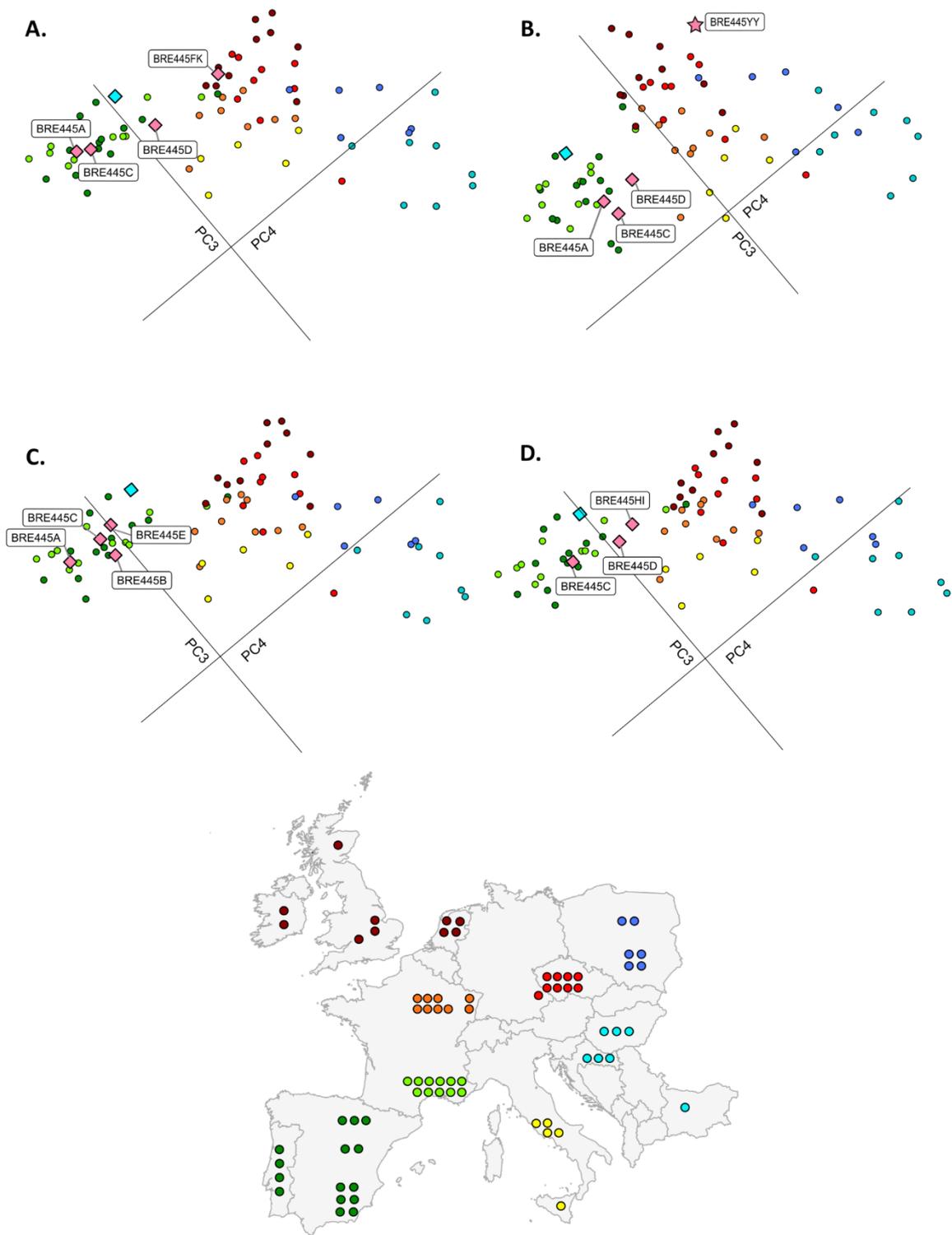


Figure S15. Same PC3-PC4 plots shown in the Figure S14 highlighting geographical locations of Late Neolithic and Bronze Age individuals. Plot A is used in the Figure 2C.

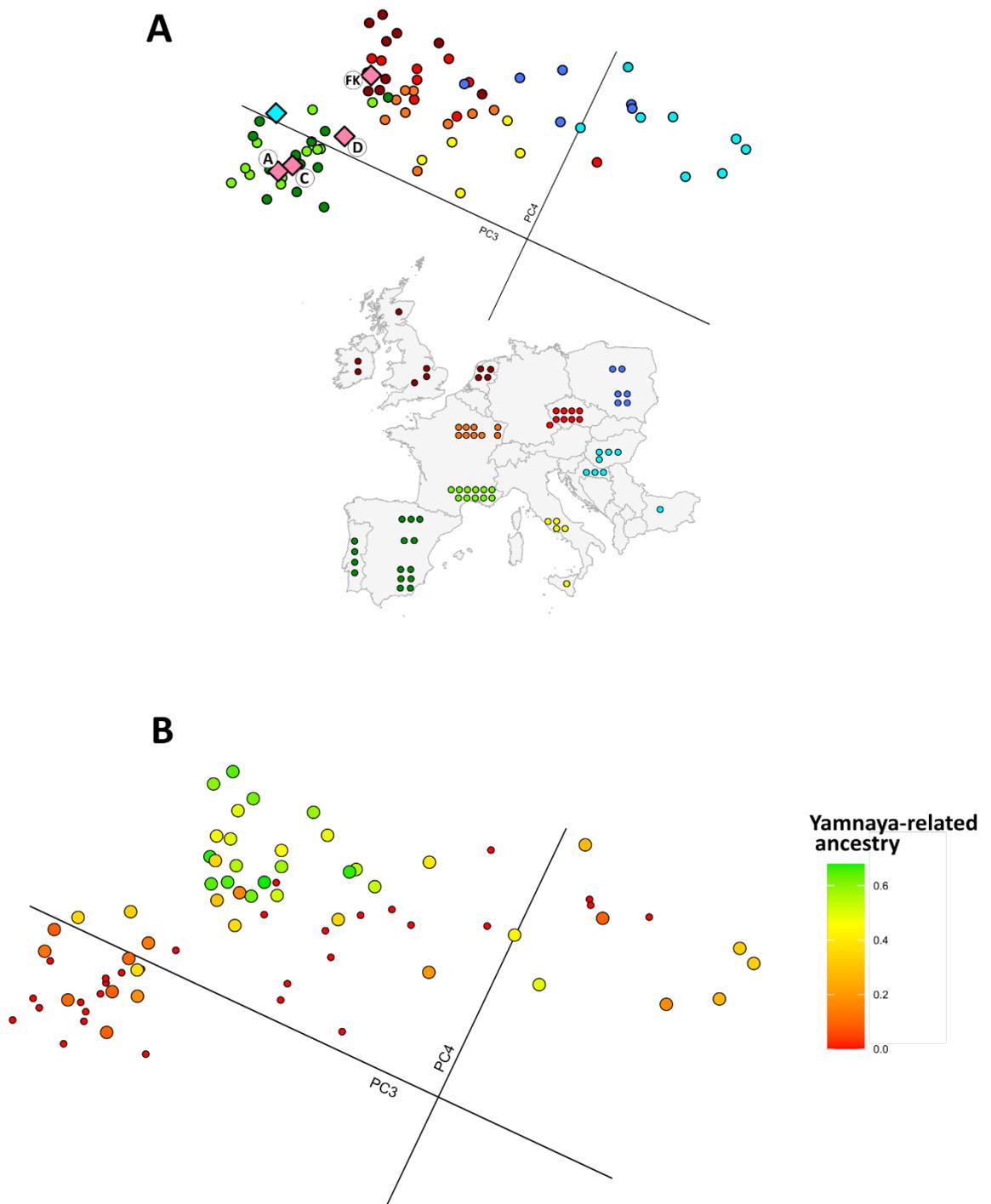


Figure S16. PC3-PC4 plots based on ChromoPainter X-chromosome coancestry matrix. **A.** The plot identifying the genomes used. **B.** The plot showing the autosomal steppe ancestry proportions (color gradient). The smaller red circles correspond to Neolithic individuals without any steppe ancestry.

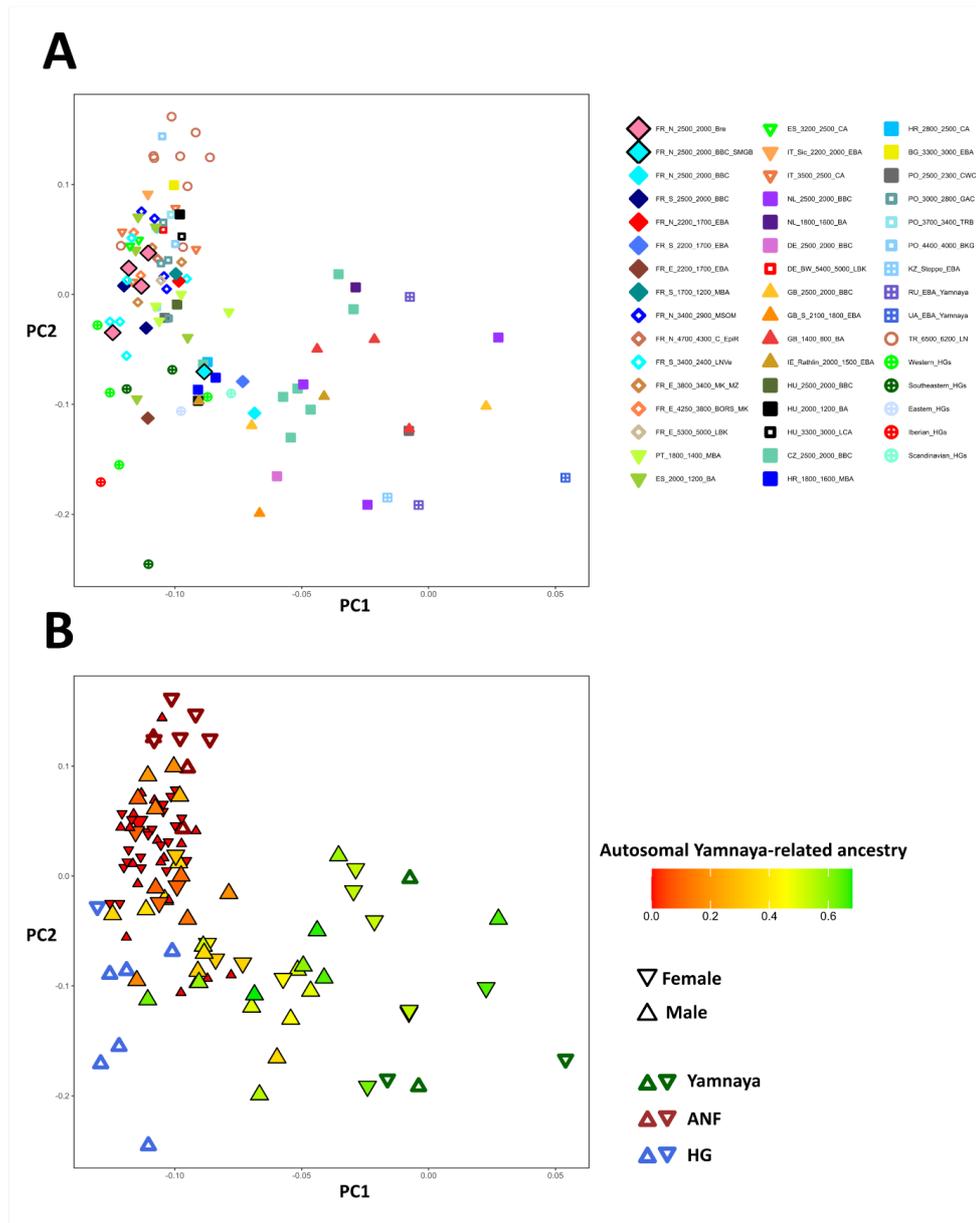


Figure S17. PC1-PC2 plots based on ChromoPainter X-chromosome coancestry matrix. A. The plot identifying the genomes used. **B.** The plot showing the autosomal steppe ancestry proportions (color gradient) and the sex of the individuals. European Neolithic individuals without steppe ancestry are shown as smaller triangles. Ancestral groups that are used to infer the ancestry proportion (HG, ANF, Yamnaya) are shown as empty triangles.

Although the X-chromosome PCA shows that the genetic heterogeneity of the source population on the X-chromosome is higher than on the autosomes (compared with Fig. 2A), there is a global correlation as revealed by a red-to-green gradient from right (steppe-ancestry side) to left (HG/ANF ancestry side). Some individuals with relatively high steppe-ancestry on the autosomes (green to yellow) have a higher Neolithic ancestry on the X-chromosome as shown by their position on the left of the PCA plot.

Table S1. Summary of archaeological and genetic information on individuals from the newly reported data

Sample ID	Location	Country	Latitude	Longitude	Radiocarbon Dates	Sex	MtDNA hg	Y hg	Sampled material	Estimated age of death	Coverage (X)	Contamination Xchr	Contamination MtDNA
Bre445A	les Pointes et les Grévottes, Bréviandes, Aube	France	48.256N	4.095E	-	F	K1b1a1	-	petrous	20-39	0.75	-	Not Detected
Bre445B	les Pointes et les Grévottes, Bréviandes, Aube	France	48.256N	4.095E	2580-2275 calBCE (98.9%) 3940 ± 40 BP (Gn-31284)	F	T2c1e	-	petrous	20-39	0.77	-	Not Detected
Bre445C	les Pointes et les Grévottes, Bréviandes, Aube	France	48.256N	4.095E	-	F	V/(Hv0a)	-	petrous	6-10	0.95	-	Not Detected
Bre445D	les Pointes et les Grévottes, Bréviandes, Aube	France	48.256N	4.095E	-	F	T2c1e	-	petrous	Neonate	0.96	-	Not Detected
Bre445E	les Pointes et les Grévottes, Bréviandes, Aube	France	48.256N	4.095E	2706 - 2287 calBCE (96.9%) 4000 ± 50 BP (Gn-31285)	F	U5b	-	petrous	>60	1.1	-	Not Detected
Bre445FK	les Pointes et les Grévottes, Bréviandes, Aube	France	48.256N	4.095E	2567 - 2339 cal BCE (95.4%) 3945 ± 35 BP (Ua-69965)	M	U5b	R1b1a1b1a1a	petrous	20-39	4.6	0.006305	Not Detected
Bre445Hl	les Pointes et les Grévottes, Bréviandes, Aube	France	48.256N	4.095E	-	M	K1b1a1	R1b1a1b1a1a	petrous	4-8	1.2	0.010296	Not Detected
SMG65A	Saint-Martin-la-Garnme, Yvelines, Île-de-France	France	49.040N	1.089E	2339-2139 cal BCE (4189 calBP beta)	M	T2b3-151	R1b1a1b1a1a	petrous	30-49	2.3	0.003613	Not Detected

Table S2. Kinship estimation for newly reported data

Ind1	Ind2	nSites	NgsRelate - ANGSD GI autosome				NgsRelate - Imputed GI			
			θ	k0	k1	k2	nSites	θ	k0	k1
BRE445B	BRE445D	720361	0.24	0.11	0.83	0.06	523883	0.26	0.00	
BRE445E	BRE445FK	1541359	0.23	0.10	0.90	0.01	523883	0.25	0.00	
BRE445FK	BRE445Hl	1723629	0.23	0.09	0.90	0.01	523883	0.26	0.00	
BRE445A	BRE445Hl	804074	0.22	0.17	0.77	0.06	523883	0.25	0.00	
BRE445E	BRE445Hl	1013374	0.12	0.61	0.32	0.07	523883	0.14	0.43	

Table S2. Kinship estimation for newly reported data

Kin-relationship	Autosomal θ	chr.X θ	k_0	k_1	k_2
Mother-son	0.25	0.5	0	1	0
Father-daughter	0.25	0.5	0	1	0
Father-son	0.25	0	0	1	0
Sisters	0.25	0.375	0.25	0.5	0.25
Brothers	0.25	0.25	0.5	0.5	0.25
Brother-sister	0.25	0.25	0.25	0.5	0.25

Table S2 (Y aka et al., Variable kinship patterns in Neolithic Anatolia revealed by ancient genomes, Current Biology (2021), <https://doi.org/10.1016/j.cub.2021.04.011>)

Table S4. SOURCEFIND ancestry estimation for ancient individuals

Sample ID	Group label	HG	ANF	Yamnaya
Bre445A	FR_N_2500_2000_Bre	0,244	0,756	0
Bre445C	FR_N_2500_2000_Bre	0,259	0,741	0
Bre445D	FR_N_2500_2000_Bre	0,217	0,783	0
Bre445E	FR_N_2500_2000_Bre	0,308	0,692	0
Bre445FK	FR_N_2500_2000_Bre	0,195	0,453	0,352
Bre445HI	FR_N_2500_2000_Bre	0,223	0,645	0,131
Bre445YY	FR_N_2500_2000_Bre	0,079	0,261	0,660
SMGB54	FR_N_2500_2000_BBC	0,158	0,484	0,358
OBE3626-1	FR_E_2200_1700_EBA	0,117	0,276	0,607
ROS45	FR_E_4700_4200_PF	0,121	0,879	0
ROS78	FR_E_4700_4200_PF	0,104	0,896	0
WET370	FR_E_3800_3400_MK_MZ	0,135	0,865	0
BERG02-2	FR_E_4250_3800_BORS_MK	0,155	0,845	0
BERG157-2	FR_E_4250_3800_BORS_MK	0,154	0,846	0
BERG157-7	FR_E_4250_3800_BORS_MK	0,191	0,809	0
Sch72-15	FR_E_5300_5000_LBK	0,021	0,979	0
PSS4170	FR_N_2200_1700_EBA	0,231	0,512	0,257
CBV95	FR_N_2500_2000_BBC	0,128	0,193	0,679
1H13	FR_N_3400_2900_MSOM	0,251	0,749	0
2H07	FR_N_3400_2900_MSOM	0,192	0,808	0
2H11	FR_N_3400_2900_MSOM	0,184	0,816	0
2H10	FR_N_3400_2900_MSOM	0,562	0,438	0
PSS4693	FR_N_3600_3400_PostMK	0,135	0,865	0
BUCH2	FR_N_4700_4300_C_EpiR	0,148	0,852	0
GBVPO	FR_S_1700_1200_MBA	0,135	0,557	0,308
TORTF	FR_S_2200_1700_EBA	0,140	0,524	0,335
GBVPK	FR_S_2500_2000_BBC	0,222	0,404	0,374
GBVPL	FR_S_3400_2400_LNVe	0,194	0,806	0
ROUQE	FR_S_3400_2400_LNVe	0,189	0,811	0
ROUQF	FR_S_3400_2400_LNVe	0,211	0,789	0
ROUQH	FR_S_3400_2400_LNVe	0,196	0,804	0
ROUQV	FR_S_3400_2400_LNVe	0,189	0,811	0
ROUQW	FR_S_3400_2400_LNVe	0,237	0,763	0
TORTE	FR_S_3400_2400_LNVe	0,219	0,781	0
CRE20D	FR_S_4350_4000_CH	0,139	0,861	0
esp005	ES_2000_1200_BA	0,206	0,636	0,158
ALM017	ES_2000_1200_BA	0,179	0,700	0,122
ALM041	ES_2000_1200_BA	0,173	0,740	0,087
ALM050	ES_2000_1200_BA	0,254	0,607	0,139
CBR004	ES_2000_1200_BA	0,198	0,715	0,087
ATP16	ES_3200_2500_CA	0,121	0,879	0
ATP2	ES_3200_2500_CA	0,188	0,812	0
I5838	ES_3200_2500_CA	0,193	0,807	0
I1838	ES_3200_2500_CA	0,215	0,785	0
TV32032	PT_1800_1400_MBA	0,187	0,724	0,089
TV3831	PT_1800_1400_MBA	0,182	0,636	0,183
VO10207	PT_1800_1400_MBA	0,225	0,672	0,103
MG104	PT_1800_1400_MBA	0,217	0,694	0,090
RM127	IE_Rathlin_2000_1500_EBA	0,076	0,272	0,652
RSK1	IE_Rathlin_2000_1500_EBA	0,075	0,335	0,589
I4071	NL_1800_1600_BA	0,146	0,319	0,535
I5748	NL_2500_2000_BBC	0,099	0,246	0,655
I4073	NL_2500_2000_BBC	0,123	0,252	0,626
I4075	NL_2500_2000_BBC	0,152	0,245	0,603
I2520	BG_3300_3000_EBA	0,139	0,691	0,171
I7286	CZ_2500_2000_BBC	0,227	0,284	0,489
I7278	CZ_2500_2000_BBC	0,089	0,415	0,496
I7212	CZ_2500_2000_BBC	0,144	0,421	0,434
I7211	CZ_2500_2000_BBC	0,104	0,314	0,582
I4889	CZ_2500_2000_BBC	0,064	0,417	0,519
I4888	CZ_2500_2000_BBC	0,083	0,461	0,456
I7281	CZ_2500_2000_BBC	0,081	0,378	0,541
I7289	CZ_2500_2000_BBC	0,127	0,387	0,486
I5835	DE_2500_2000_BBC	0,194	0,453	0,353
I7574	GB_1400_800_BA	0,120	0,207	0,673
I3130	GB_1400_800_BA	0,148	0,320	0,531
I2859	GB_1400_800_BA	0,129	0,335	0,535
I3255	GB_2500_2000_BBC	0,096	0,254	0,650
I2445	GB_2500_2000_BBC	0,185	0,317	0,497
I3132	GB_S_2100_1800_EBA	0,065	0,383	0,552
JAG78	HR_1800_1600_MBA	0,196	0,473	0,331
JAG85	HR_1800_1600_MBA	0,178	0,494	0,328
Pop39	HR_2800_2500_CA	0,055	0,522	0,423
BR2	HU_2000_1200_BA	0,137	0,580	0,283
R15E479	HU_2000_1200_BA	0,237	0,492	0,271
I7045	HU_2500_2000_BBC	0,121	0,786	0,093
R5	IT_3500_2500_CA	0,124	0,876	0
R1014	IT_3500_2500_CA	0,140	0,860	0
R24	IT_3500_2500_CA	0,145	0,855	0
R4	IT_3500_2500_CA	0,067	0,933	0
I3123	IT_Sic_2200_2000_EBA	0,076	0,724	0,200
N45	PO_2500_2300_CWC	0,119	0,209	0,672
N47	PO_2500_2300_CWC	0,225	0,375	0,400
Koszyce_3_10	PO_3000_2800_GAC	0,236	0,764	0
Koszyce_3_12	PO_3000_2800_GAC	0,203	0,797	0
Koszyce_3_1	PO_3000_2800_GAC	0,264	0,736	0
Koszyce_3_8	PO_3000_2800_GAC	0,270	0,730	0
N20	PO_3700_3400_TRB	0,211	0,789	0
N18	PO_3700_3400_TRB	0,214	0,786	0
N25	PO_4400_4000_BKG	0,127	0,873	0
N27	PO_4400_4000_BKG	0,127	0,873	0

Table S5. Summary of previously published individuals used in the shotgun dataset

Sample ID	Group label (in this study)	Cal 20	95s Cal 20 (95s)	Radiocarbon Date (BP)	Radiocarbon Date (s)	Group label (published)	Source	Location	Country	Sex	MDNA hg	Y hg
CBV95	FR_N_2500_2000_BBC	2575	2350	3970	30	France_Bell_Beaker	BruneiPNAS2020	La Bouche-à-Vesle, Ciry, Salsogne	France	M	J2a1a	R1b1a1a2
PE12	FR_N_2500_2000_BBC	2564	2302	3935	30	France_Bell_Beaker	BruneiPNAS2020	Dolmen des Peirnières, Villardubert	France	F	H1	
OBE3626-1	FR_E_2200_1700_EBA	1930	1701	3505	35	France_Bronze_Age	BruneiPNAS2020	Obernal	France	M	R1b	R1b1a1b1
PSS4170	FR_N_2200_1700_EBA	2197	1975	3690	30	France_Bronze_Age	BruneiPNAS2020	Pont-sur-Seine	France	M	H25	R1b1a1b1
PSS4693	FR_N_3600_3400_PostMK	3605	3449	4650	45	France_Neolithic	BruneiPNAS2020	Pont-sur-Seine	France	F	US5b3b	
CRE200	FR_S_4350_4000_CH	4350	4000			France_Neolithic	BruneiPNAS2020	Le Crès, Béziers	France	F	H-1629a	
BER602-2	FR_E_4350_3800_BORS_MK	4041	3802	5135	35	France_Neolithic	BruneiPNAS2020	Saulgier, Berghheim	France	F	W5b	
BER6157-2	FR_E_4250_3800_BORS_MK	4300	3900			France_Neolithic	BruneiPNAS2020	Saulgier, Berghheim	France	F	J1c1b	
BER6157-7	FR_E_4250_3800_BORS_MK	4318	4020	5335	35	France_Neolithic	BruneiPNAS2020	Saulgier, Berghheim	France	M	US5b1	I2a1a2
BUCH2	FR_N_4700_4300_C_Epil	4700	4300			France_Neolithic	BruneiPNAS2020	Buchères	France	M	US5b1-16189	H2a1
ROS45	FR_E_4700_4200_Pf	4789	4607	5833	29	France_Neolithic	BruneiPNAS2020	Rosheim "Mittelfeld" "Rosenmeer"	France	M	H5u	I2a1a2
ROS78	FR_E_4700_4200_Pf	4700	4500			France_Neolithic	BruneiPNAS2020	Rosheim "Mittelfeld" "Rosenmeer"	France	M	US81b1	I2a1a2
WE1370	FR_E_3800_3400_MK_MZ	3800	3400			France_Neolithic	BruneiPNAS2020	Wettolsheim	France	M	H1c5f	H2
Sch72-15	FR_E_5300_5000_LBK	5466	5215	6340	40	France_Early_Neolithic	BruneiPNAS2020	Schwindratsheim	France	M	T2f	C1a2
1H13	FR_N_3400_2900_MSOM	3092	2921	4395	20	France_Mont_Aime	Seguin-OrlandoPNAS2021	Mont-Aimé hypogée I	France	M	K1a4a1	I2a1b1b1
2H07	FR_N_3400_2900_MSOM	3366	3104	4540	20	France_Mont_Aime	Seguin-OrlandoPNAS2021	Mont-Aimé hypogée II	France	F	K1a4a1h	
2H11	FR_N_3400_2900_MSOM	3351	3102	4510	20	France_Mont_Aime	Seguin-OrlandoPNAS2021	Mont-Aimé hypogée II	France	M	J1c1	I2a1a1b1
2H10	FR_N_3400_2900_MSOM	3341	3098	4495	20	France_Mont_Aime	Seguin-OrlandoPNAS2021	Mont-Aimé hypogée II	France	M	H3	I2a1b1b1
GBVPK	FR_S_2500_2000_BBC	2461	2299	3890	15	France_Basse_Vigne_Perdue	Seguin-OrlandoPNAS2021	Grotte Basse de la Vigne Perdue, les Monges	France	M	J2b1a	R1b1a1a1a2a1
GBVPL	FR_S_1700_1200_MBA	1737	1547	3360	15	France_Basse_Vigne_Perdue	Seguin-OrlandoPNAS2021	Grotte Basse de la Vigne Perdue, les Monges	France	F	US5b	
ROUQE	FR_S_3400_2400_LNve	3321	2937	4435	15	France_Rouquet	Seguin-OrlandoPNAS2021	Grotte du Rouquet, Moujan	France	F	H1-152	
ROUQF	FR_S_3400_2400_LNve	3304	3096	4490	20	France_Rouquet	Seguin-OrlandoPNAS2021	Grotte du Rouquet, Moujan	France	M	V	I2a1a2
ROUQH	FR_S_3400_2400_LNve	3346	3101	4505	20	France_Rouquet	Seguin-OrlandoPNAS2021	Grotte du Rouquet, Moujan	France	M	U2e1c1	I2a1b1b1
ROUQV	FR_S_3400_2400_LNve	3354	3102	4515	15	France_Rouquet	Seguin-OrlandoPNAS2021	Grotte du Rouquet, Moujan	France	F	H3	
ROUQW	FR_S_3400_2400_LNve	3356	3101	4515	25	France_Rouquet	Seguin-OrlandoPNAS2021	Grotte du Rouquet, Moujan	France	M	X2c1	I2a1b1b1
TORTF	FR_S_2200_1700_EBA	2028	1825	3565	15	France_Tortues	Seguin-OrlandoPNAS2021	Grotte des Tortues, les Monges	France	F	H5a1a	
TORTF	FR_S_3400_2400_LNve	2839	2500	4080	15	France_Tortues	Seguin-OrlandoPNAS2021	Grotte des Tortues, les Monges	France	F	K1a2b	
I2520	BG_3300_3000_EBA	3336	3029	4470	25	Bulgaria_EBA	MathiesonNature2018	Veliiko Tarnovo, Dzulyunitsa	Bulgaria	M	H	H
I7286	CZ_2500_2000_BBC	2403	2199	3830	20	Czech_BellBeaker	OlaldeNature2018	Radosevice	Czech Republic	M	I4a	R1b1a1a1a2b1
I7278	CZ_2500_2000_BBC	2462	2298	3890	20	Czech_BellBeaker	OlaldeNature2018	Radosevice	Czech Republic	M	I4a	R1b1a1a1a2b1
I7212	CZ_2500_2000_BBC	2454	2201	3840	25	Czech_BellBeaker	OlaldeNature2018	Radosevice	Czech Republic	M	K1b1a1-199	R1b1a1a1a2b1
I7211	CZ_2500_2000_BBC	2296	2143	3800	20	Czech_BellBeaker	OlaldeNature2018	Radosevice	Czech Republic	F	US5b	
I7045	CZ_2500_2000_BBC	2440	2201	3835	20	Hungary_EBA_BellBeaker	OlaldeNature2018	Szigetszentmiklós-Udúlósr	Hungary	F	T2c1d-152	
H889	CZ_2500_2000_BBC	2284	2056	3765	20	Czech_BellBeaker	OlaldeNature2018	Prague 8, Kobylisy, Ke Strice Street	Czech Republic	M	H9	R1b1a1a1a2b1
H898	CZ_2500_2000_BBC	2199	2027	3700	20	Czech_BellBeaker	OlaldeNature2018	Prague 8, Kobylisy, Ke Strice Street	Czech Republic	M	K1c2c	R1b1a1a1a2b1
I7281	CZ_2500_2000_BBC	2340	2146	3810	20	Czech_BellBeaker	OlaldeNature2018	Prague 5 - Mala Ohrada	Czech Republic	F	H5a1g	R1b1a1a1a2b1
I7289	CZ_2500_2000_BBC	2456	2209	3860	20	Czech_BellBeaker	OlaldeNature2018	Radosevice	Czech Republic	M	US81b1	R1b1a1a1a2b1
IS835	DE_2500_2000_BBC	2500	2000			Germany_BellBeaker	OlaldeNature2018	Ilbich LKR	Germany	M	J1c1	R1b1a1a1a2b1
BK_Suttgart	DK_BW_5400_5000_LBK	5071	5071	6246	30	LBN_EBA	LazaridisNature2014	Viesenhauer Hof, Stuttgart-Muehlhausen	Germany	F	Ta21b	
esd005	ES_2000_1200_BA	2000	1200			Spain_BA	ValdioseraPNAS2018	Cueva de los Lagos, La Riga	Spain	M	K1a	R1b1-DF27
ALM017	ES_2000_1200_BA	1882	1698	3474	25	Almoloja_Argar_Early	Villalba-MoucoScienceAdvances2021	La Almoloja, Pilego, Murcia	Spain	M	H2b1	R1b1a1a1a2a1
ALM014	ES_2000_1200_BA	1944	1774	3540	19	Almoloja_Argar_Early	Villalba-MoucoScienceAdvances2021	La Almoloja, Pilego, Murcia	Spain	M	T1a1	R1b1a1a1a2a1
ALM050	ES_2000_1200_BA	2000	1750			Almoloja_Argar_Early	Villalba-MoucoScienceAdvances2021	La Almoloja, Pilego, Murcia	Spain	M	H5a1a	R1b1a1a1a2a1
CR004	ES_2000_1200_BA	1538	1441	3244	20	St_Cabozelindo_BA	Villalba-MoucoScienceAdvances2021	Cabezo Redondo, Villena, Alicante	Spain	F	H5a1a	R1b1a1a1a2a1
ZAP002	ES_2000_1200_BA	1750	1550			ZAP002	Villalba-MoucoScienceAdvances2021	Lorca, Murcia	Spain	M	H5a1	R1b1a1a1a2a1
ATP16	ES_3200_2500_CA	3265	2913	4400	30	C_Iberia_CA	ValdioseraPNAS2018	El Portalon Cave, Sierra de Atapuerca	Spain	F	X2c	
ATP2	ES_3200_2500_CA	2900	2675	4210	30	C_Iberia_CA	ValdioseraPNAS2018	El Portalon Cave, Sierra de Atapuerca	Spain	M	US5b3	H2
IS838	ES_3200_2500_CA	2900	2300			Spain_C	LipsonNature2017	Burgos, Atapuerca, El Mirador Cave	Spain	M	K1a-195	I2a1b1a1*
IS838	ES_3200_2500_CA	3356	2936	4470	50	Spain_C	LipsonNature2017	Basque Country, Alava, Las Yurdias 1b	Spain	F	J1c3	
I7574	GB_1400_800_BA	1408	1226	3055	30	England_MBA	OlaldeNature2018	England, Cambridgeshire, Trumpington, Clay Farm	Great Britain	F	H1a3	
I3130	GB_1400_800_BA	984	826	2758	29	Scotland_LBA	OlaldeNature2018	Scotland, Coevessa Caves	Great Britain	F	US5a1c	
I2859	GB_1400_800_BA	912	808	2714	29	Scotland_LBA	OlaldeNature2018	Scotland, Coevessa Caves	Great Britain	M	K1a2c	R1b1a1a1a2a
I2555	GB_2500_2000_BBC	2139	1947	3661	21	England_BellBeaker	OlaldeNature2018	England, Cambridgeshire, Trumpington Meadows	Great Britain	F	T2b	
I2445	GB_2500_2000_BBC	2200	1985	3709	27	England_BellBeaker	OlaldeNature2018	England, Oxfordshire, Yarnon	Great Britain	M	X2b6	R1b1a1a1a2c1
I3132	GB_S_2100_1800_EBA	2126	1886	3614	33	Scotland_C_EBA	OlaldeNature2018	Scotland, Coevessa Cave 2	Great Britain	M	Ta21b1a	R1b1a1a1a2c1a5c
JAG78	HR_1800_1600_MBA	1800	600			Croatia_Jag_MBA	FreilichScientificReports2021	Jagodnjak-Krecevic, Osijek	Croatia	M	T2b11	G2a2a1a2a2a1-231430
JAG85	HR_1800_1600_MBA	1876	1687	3445	20	Croatia_Jag_MBA	FreilichScientificReports2021	Jagodnjak-Krecevic, Osijek	Croatia	F	H9a	
Pop99	HR_2800_2500_CA	2859	2500	4095	28	Croatia_Pop_CA	FreilichScientificReports2021	Belj Manastir-Popova zemlja, Osijek	Croatia	F	H9f	
BR2	HU_2000_1200_BA	987	833	2769	24	Hungary_LBA	GambaNatureCommunications2014	Ludas-Varju-Dulo	Hungary	M	K1a1a	J2a1
R1E5479	HU_2000_1200_BA	2000	1500			Hungary_BA	AllentoftNature2015	Erd	Hungary	M	T2b	I2a21a1a2a
I5119	HU_3300_3000_LCA	3330	3013	4445	20	Hungary_LateC_EBA_Baden_Yamnaya	OlaldeNature2018	Mezőcsát-Horcsadócs	Hungary	F	J1c2	
RW127	IE_Rathlin_2000_1500_EBA	2000	1400	3601	30	Ireland_BA	CassidyPNAS2016	Rathlin Island, County Antrim	Ireland	M	US5a1b	R1b1a1a1a2c1a2c1g
RSK1	IE_Rathlin_2000_1500_EBA	2025	1700	3539	54	Ireland_BA	CassidyPNAS2016	Rathlin Island, County Antrim	Ireland	M	US5b2a2	R1b1a1a1a2c1a2c1
RS	IE_3500_2500_CA	2902	2877			Italy_CA	AntonioScience2019	Grotta Continenza	Italy	F	H1	H1
R1014	IT_3500_2500_CA	3500	2500			Italy_CA	AntonioScience2019	Monte San Biagio	Italy	M	H1a1a3a3	H-1901
R24	IT_3500_2500_CA	3626	3374			Italy_CA	AntonioScience2019	San Croce/Issa	Italy	F	H1a1a3	
R4	IT_3500_2500_CA	2960	2880			Italy_CA	AntonioScience2019	Grotta Continenza	Italy	M	K1a-195	G-F1193
I3123	IT_Sik_2200_2000_EBA	2287	2041	3760	30	Italy_Sicily_EBA	FernandesNatureEcologyEvolution2020	Sicily, Buffa	Italy	M	US81b1	R1b1a1a1a2a
I4071	NL_1800_1600_BA	1883	1634	3450	40	Netherlands_BA	OlaldeNature2018	Noord-Holland, Oostwoud, De Tuithorn	Netherlands	F	H6a1a	
I5748	NL_2500_2000_BBC	2579	2211	3945	55	Netherlands_BellBeaker	OlaldeNature2018	Noord-Holland, Oostwoud, De Tuithorn	Netherlands	M	X2d4	R1b1a1a1a2
I4073	NL_2500_2000_BBC	2197	1897	3660	50	Netherlands_BellBeaker	OlaldeNature2018	Noord-Holland, Oostwoud, De Tuithorn	Netherlands	M	US5b1	R1b1a1a1a2
I4075	NL_2500_2000_BBC	2127	1933	3635	20	Netherlands_BellBeaker	OlaldeNature2018	Noord-Holland, Oostwoud, De Tuithorn	Netherlands	F	H5a1	
N45	PO_2500_2300_CWC	2570	2340			Poland_CWC	FernandesScientificReports2018	Pikutkowo	Poland	F	H5a8	
N47	PO_2500_2300_CWC	2570	2340			Poland_CWC	FernandesScientificReports2018	Pikutkowo	Poland	M	US5b2a2	I2a2a
Koszyce_3_4	PO_3000_2800_GAC	2872	2578	4126	36	Poland_Koszyce_GAC	SchroederPNAS2019	Koszyce, site 3	Poland	M	H10a	I2a-1801
Koszyce_3_1	PO_3000_2800_GAC	2900	2696	4213	36	Poland_Koszyce_GAC	SchroederPNAS2019	Koszyce, site 3	Poland	F	K1a1b1a	
Koszyce_3_2	PO_3000_2800_GAC	2884	2671	4180	22	Poland_Koszyce_GAC	SchroederPNAS2019	Koszyce, site 3	Poland	F	T2b	
Koszyce_3_8	PO_3000_2800_GAC	3072	2888	4330	34	Poland_Koszyce_GAC	SchroederPNAS2019	Koszyce, site 3	Poland	F	J1c3f	
N20	PO_3700_3400_TRB	3636	3389			Poland_TRB	FernandesScientificRep2018	Pikutkowo	Poland	F	H9f	
N18	PO_3700_3400_TRB	3636	3389			Poland_TRB	FernandesScientificRep2018	Pikutkowo	Poland	F	N1a1a1	
N25	PO_4400_4000_BKG	4600	4000			Poland_BKG	FernandesScientificRep2018	Oslonki	Poland	F	K2a	
N27	PO_4400_4000_BKG	4600	4000			Poland_BKG	FernandesScientificRep2018	Oslonki	Poland	M	V14	
TV3032	PT_1800_1400_MBA	1800	1400			Portugal_MBA	MartiniopLoSGenetics2017	Torre Velha 3	Portugal	M	X2b-226	R1b1a2
TV3831	PT_1800_1400_MBA	1800	1400			Portugal_MBA	MartiniopLoSGenetics2017	Torre Velha 3	Portugal	M	H1-152	R1b1a2a1a

Table S6. Summary of previously published individuals used in the 1240k dataset

ID	Group label (in this study)	Location	Country	a	High CoX	Low CoX	Carbon	Date	Collection	Sex	MIDNA hg	Y hg	Group label (published)	Source
MX188	CH 2800 2200_CWC	Spreitenbach CWC	Switzerland	2465	2309					M	U2a1a2	I2a1a2	Spreitenbach CWC	FurtwanglerNatureCommunications2020
MX191	CH 2800 2200_CWC	Spreitenbach CWC	Switzerland	2570	2190					M	H9+16189	I2a1a2	Spreitenbach CWC	FurtwanglerNatureCommunications2020
MX196	CH 2800 2200_CWC	Spreitenbach CWC	Switzerland	2480	2290					F	J1c3	I2a1a2	Spreitenbach CWC	FurtwanglerNatureCommunications2020
MX197	CH 2800 2200_CWC	Spreitenbach CWC	Switzerland	2490	2294					F	H11a+152	I2a1a2	Spreitenbach CWC	FurtwanglerNatureCommunications2020
MX198	CH 2800 2200_CWC	Spreitenbach CWC	Switzerland	2900	2300					F	J1c9	I2a1a2	Spreitenbach CWC	FurtwanglerNatureCommunications2020
MX199	CH 2800 2200_CWC	Spreitenbach CWC	Switzerland	2499	2197					F	J2b1a	I2a1a2	Spreitenbach CWC	FurtwanglerNatureCommunications2020
MX195	CH 2800 2200_CWC	Spreitenbach CWC	Switzerland	2470	2050					M	H11a+152	I2a1a2	Spreitenbach CWC	FurtwanglerNatureCommunications2020
MX190	CH 2800 2200_CWC	Spreitenbach CWC	Switzerland	2571	2013					M	H11a+152	I2a1a2	Spreitenbach CWC	FurtwanglerNatureCommunications2020
MX190	CH 2800 2200_CWC	Spreitenbach CWC	Switzerland	2860	2460					M	U2a1a2	I2a1a2	Spreitenbach CWC	FurtwanglerNatureCommunications2020
Aesch25	CH 2800 2200_CWC	Aesch	Switzerland	2864	2501					M	X2b+226	R1b1a1a1a1a	Aesch	FurtwanglerNatureCommunications2020
MX298	CH 2800 2200_CWC	Wartau	Switzerland	2620	2448					M	X2d4	G2a2b	Wartau	FurtwanglerNatureCommunications2020
SX21	CH 2800 2200_FN	Wartau	Switzerland	2462	2309					M	H9	I2a1a2	Wartau	FurtwanglerNatureCommunications2020
SX26	CH 2800 2200_FN	Wartau	Switzerland	2461	2295					F	K2b1	I2a1a2	Wartau	FurtwanglerNatureCommunications2020
SX17	CH 2800 2200_FN	Wartau	Switzerland	2470	2336					F	U4a2f	G2a2a1a2a1a	Wartau	FurtwanglerNatureCommunications2020
SX10	CH 2800 2200_FN	Rapperswil Zurichstrasse	Switzerland	2695	2481					M	U4a1	G2a2a1a2a1a	Rapperswil Zurichstrasse	FurtwanglerNatureCommunications2020
KN002	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2201	1978	3705	34			F	T2a1b1a1	I2a2	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
KN006	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2131	1946	3648	19			F	J1c	I2a2	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
KN006	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2131	1947	3651	20			F	K2a11	I2a2	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
KN007	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2133	1947	3651	18			F	T1a1	I2a2	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
KN106	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2020	1780	3570	22			F	H9	I2a2	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
KO107	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1881	1691	3460	25			M	R1b1	I2a2	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
KO108	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1881	1691	3460	25			M	U2e2a1c	R1a - Z645	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
KO109	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2036	1883	3589	24			F	H9	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
KO110	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2027	1890	3599	24			F	H1b	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
KO111	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2136	1950	3660	25			F	W1c	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
KO113	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2029	1893	3605	24			F	H5b	R1a - Z645	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
KO114	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1887	1625	3416	25			F	H9	R1a - Z645	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
KO115	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1918	1746	3509	24			F	U5a11f	R1a - Z645	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
KO701	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2130	1947	3655	24			M	H4a	R1a - Z645	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
KZ702	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1919	1751	3516	20			F	H2b	R1b - P312	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
MIB01	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1884	1749	3494	19			F	U4b1b1	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
MIB03	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1942	1773	3538	19			F	K1a1b2a	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
MIB04	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1888	1723	3503	18			M	R1b1	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
MIB05	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1886	1749	3498	19			M	H2a1b	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
MIB024	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2020	1827	3575	22			F	H1b	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
MIB028	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2013	1779	3562	22			F	U5b2a2c	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
MIB029	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1887	1625	3416	25			F	H9	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
MIG10	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2135	1951	3659	23			F	T2e	R1b - P312	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
MIG11	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2136	1949	3658	24			F	R1b	R1b - P312	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
MIG12	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2136	1949	3658	24			F	H1c1b1a	G2a2b2a	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
MS001	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2026	1892	3601	19			M	U5b1a	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
MS002	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2020	1883	3582	19			M	T2c1d1	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
MS004	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1959	1773	3550	23			M	H17	I2a2	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
MS006	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2452	2152	3833	27			M	L1'2'4'5'6'	R1b - P312	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
PM003	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2182	1947	3670	22			F	H9	R1b - P312	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
PM004	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2141	1956	3675	30			M	H2a2b	I2a2	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
PM006	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2284	2037	3750	30			M	U5a1g1	I2a1	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
PM008	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2026	1873	3580	30			F	T2b	R1b*	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
PM010	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2033	1831	3595	30			F	H3a1a	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
PM012	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1877	1564	3410	40			F	J2b1a	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
ROU01	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1959	1773	3550	23			F	U5a1g2	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
ROU02	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1888	1723	3503	18			M	R1b1	R1b - P312	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
VI039	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1958	1748	3536	34			F	T2b	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
VI042	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1940	1696	3504	41			M	U5a1g2	I2a2	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
VI045	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2184	1778	3623	58			F	H1c1b1a	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
VI046	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1958	1748	3536	34			F	H9	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
VI047	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1958	1748	3536	34			F	T2c1d1a	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
VI050	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1950	1745	3526	36			F	H7a1	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
VI054	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2203	2032	3725	28			F	W111	R1b - P312	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
VI050	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1950	1745	3526	36			F	T1a1	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
VI051	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	1950	1745	3526	36			F	V	R1b - P312	Bohemia_Uneitic_preClassical	PapaScienceAdvances2021
VI065	CZ 2200 1700_Uneitic_EBA	Bohemia	Czechia	2140	1780	3618	52			F	H2b	R1b - P312	Bohemia_Uneitic_Classical	PapaScienceAdvances2021
I4885	DE 2500 2000_BBC	Bohemia	Czechia	2291	2142	3790	20			M	U5a1f	R1b - P312	Bohemia_BB_Late	OlaldeNature2018
I4886	DE 2500 2000_BBC	Bohemia	Czechia	2270	2039	3740	20			M	H8a	R1b - P312	Bohemia_BB_Late	OlaldeNature2018
I4887	DE 2500 2000_BBC	Bohemia	Czechia	2201	2037	3730	20			M	K1a2c	R1b - P312	excluded (kinship)	OlaldeNature2018
I4888	DE 2500 2000_BBC	Bohemia	Czechia	2195	2027	3700	20			M	K1a2c	R1b - P312	Bohemia_BB_Late	OlaldeNature2018
I4891	DE 2500 2000_BBC	Bohemia	Czechia	2284	2056	3765	20			M	U4a	R1b - P312	Bohemia_BB_Late	OlaldeNature2018
I4892	DE 2500 2000_BBC	Bohemia	Czechia	2284	2056	3765	20			M	J2a1a2	R1b - P312	Bohemia_BB_Late	OlaldeNature2018
I4895	DE 2500 2000_BBC	Bohemia	Czechia	2276	2042	3750	20			M	H10+16093	R1b - P312	Bohemia_BB_Late	OlaldeNature2018
I4896	DE 2500 2000_BBC	Bohemia	Czechia	2289	2141	3785	20			F	U5a2+16294	R1b - P312	Bohemia_BB_Late	OlaldeNature2018
I514	DE 2500 2000_BBC	Bohemia	Czechia	2287	2140	3780	20			M	H11a+152	R1b - P312	Bohemia_BB_Late	OlaldeNature2018
I7205	DE 2500 2000_BBC	Bohemia	Czechia	2286	2140	3780	20			M	H10+16093	R1b - P312	Bohemia_BB_Late	OlaldeNature2018
I7210	DE 2500 2000_BBC	Bohemia	Czechia	2201	2037	3730	20			M	H6a1b4	R1b - P312	Bohemia_BB_Late	OlaldeNature2018
I7211	DE 2500 2000_BBC	Bohemia	Czechia	2286	2143	3800	30			F	H9	R1b - P312	Bohemia_BB_Late	OlaldeNature2018
I7212	DE 2500 2000_BBC	Bohemia	Czechia	2454	2201	3840	25			M	K1b1a1+199	R1b - P312	excluded (kinship)	OlaldeNature2018
I7213	DE 2500 2000_BBC	Bohemia	Czechia	2440	2201	3835	20			F	H9	R1b - P312	Bohemia_BB_Late	OlaldeNature2018
I7214	DE 2500 2000_BBC	Bohemia	Czechia	2454	2201	3840	25			F	U8b1b1	R1b - P312	excluded (kinship)	OlaldeNature2018
I7249	DE 2500 2000_BBC	Bohemia	Czechia	2138	1973	3670	20							

15601	DE 2500 2000 BBC	Alburg-Lerchenhaid, Spedition Häring, Str. Straubing, Bavaria	Germany				F	H10e			Beaker Central Europe	OldaleNature2018	
15519	DE 2500 2000 BBC	Augsburg Sportgelände, Augsburg, Bavaria	Germany				M	Ia4	R1b1a1a2a1a2b1		Beaker Central Europe	OldaleNature2018	
15521	DE 2500 2000 BBC	Augsburg Sportgelände, Augsburg, Bavaria	Germany				F	J1b1a1			Beaker Central Europe	OldaleNature2018	
15833	DE 2500 2000 BBC	Irbach, County of Straubing-Bogen, Bavaria	Germany				M	J1c	R1b1a1a2a1a2b1		Beaker Central Europe	OldaleNature2018	
15834	DE 2500 2000 BBC	Irbach, County of Straubing-Bogen, Bavaria	Germany				F	W6S			Beaker Central Europe	OldaleNature2018	
16590	DE 2500 2000 BBC	Irbach, County of Straubing-Bogen, Bavaria	Germany				F	Xc21			Beaker Central Europe	OldaleNature2018	
16591	DE 2500 2000 BBC	Irbach, County of Straubing-Bogen, Bavaria	Germany				F	USb2c			Beaker Central Europe	OldaleNature2018	
16592	DE 2500 2000 BBC	Irbach, County of Straubing-Bogen, Bavaria	Germany				M	J1c	R1b1a1a2a1a2b1		Beaker Central Europe	OldaleNature2018	
17226	ES 2000 1200 BA	La Resquejada, San Román de Huerquia, Vallad中国家	Spain				F	K1a1b1			C Iberia BA_Cst	OldaleNature2018	
ALM001	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1960	1769	3544	27	F	X2b-22b		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM002	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1943	1769	3532	24	M	K1b1a1	R1b1a1b1a1a2	Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM003	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	X2b-22b			Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM006	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1919	1770	3520	14	M	H1	R1b1a1b1a1a2	Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM007	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	H1	R1b1a1b1a1a2		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM008	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F				Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM015	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1898	1746	3507	24	F	K1a4a1		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM016	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1945	1769	3534	24	M	H1+16189	R1b1a1b1a1a2	Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM017	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1882	1698	3474	25	M	T2b21	R1b1a1b1a1a2	Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM019	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	2127	1901	3631	24	F	K1b1a1c		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM020	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	K1a3a	R1b1a1b1a1a2		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM021	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	USb3			Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM025	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	H3	R1b1a1b1a1a2		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM026	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	K1a1b1			Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM028	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	K1a195	R1b1a1b1a1a2		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM029	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	H1b1			Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM032	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	T2b3+151	R1b1a1b1a1a2		Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM033	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	USb3a			Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM036	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	USb3	R1b1a1b1a1a2		Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM038	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1741	1541	3366	32	F	H1		Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM039	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1739	1535	3354	33	M	J1c2	R1b1a1b1a1a2	Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM042	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	2012	1775	3555	36	M	H1a1b1a		Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM041	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1944	1774	3540	19	M	H1c1a	R1b1a1b1a1a2	Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM042	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	H4a1a			Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM043	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1896	1750	3512	19	F	H1b1		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM047	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	H1a1b1	R1b1a1b1a1a2		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM050	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	H3	R1b1a1b1a1a2		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM051	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	V			Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM052	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	H1			Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM055	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	K1a4+1e			Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM057	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1874	1642	3436	19	M	H1	R1b1a1b1a1a2	Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM063	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	K1a195	R1b1a1b1a1a2		Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM064	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	USb11a	R1b1a1b1a1a2		Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM067	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	K1a4b			Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM068	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	H1a1	R1b1a1b1a1a2		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM069	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	USb11a			Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM075	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	USb1e			Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM076	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	H1+16189			Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM078	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	K1a	R1b1a1b1a1a2		Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM082	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	USb3	R1b1a1b1a1a2		Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM084	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	K2a			Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM086	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1932	1767	3524	21	F	USb11f		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
ALM087	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	K2			Almoloya Argar Late	Villalba-MoucoScienceAdvances2021	
ALM088	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1669	1504	3310	30	F	K1a195		Almoloya Argar Early	Villalba-MoucoScienceAdvances2021	
BA5002	ES 2000 1200 BA	La Bastida, Totana, Murcia	Spain	1620	1506	3298	27	M	H1c1a	R1b1a1b1a1a2	Bastida Argar	Villalba-MoucoScienceAdvances2021	
BA5022	ES 2000 1200 BA	La Bastida, Totana, Murcia	Spain	1739	1535	3354	33	M	K1a195	R1b1a1b1a1a2	Bastida Argar	Villalba-MoucoScienceAdvances2021	
BA5023	ES 2000 1200 BA	La Bastida, Totana, Murcia	Spain	2130	1941	3642	30	M	H1a1b1		Bastida Argar	Villalba-MoucoScienceAdvances2021	
BA5024	ES 2000 1200 BA	La Bastida, Totana, Murcia	Spain	2192	1980	3688	23	M	K1a195	R1b1a1b1a1a2	Bastida Argar	Villalba-MoucoScienceAdvances2021	
BA5025	ES 2000 1200 BA	La Bastida, Totana, Murcia	Spain	2134	1947	3653	21	M	H105a	R1b1a1b1a1a2	Bastida Argar	Villalba-MoucoScienceAdvances2021	
BA5027	ES 2000 1200 BA	La Bastida, Totana, Murcia	Spain				F	V3a			Bastida Argar	Villalba-MoucoScienceAdvances2021	
CB0004	ES 2000 1200 BA	Cabezo Redondo, Villena, Alicante	Spain	1538	1447	3244	20	F	USb1e		SE Iberia BA Argar	Villalba-MoucoScienceAdvances2021	
CM0001	ES 2000 1200 BA	Cerro del Morrón, Moratilla, Murcia	Spain	1881	1687	3455	28	F	T2c1d+152		SE Iberia BA Argar	Villalba-MoucoScienceAdvances2021	
EFA011	ES 2000 1200 BA	Forat de ses Arriçes, Ciutadella, Menorca	Spain				F	USb2a			Artigues LA	Villalba-MoucoScienceAdvances2021	
EH0001	ES 2000 1200 BA	Aspe, Alicante	Spain	1611	1441	3250	30	M	K1a195	R1b1a1b1a1a2	SE Iberia BA Valencian	Villalba-MoucoScienceAdvances2021	
EH0002	ES 2000 1200 BA	Aspe, Alicante	Spain	1669	1504	3310	30	F	USb1a		SE Iberia BA Valencian	Villalba-MoucoScienceAdvances2021	
PL2001	ES 2000 1200 BA	Villena, Alicante	Spain				F	K1a195			SE Iberia BA Valencian	Villalba-MoucoScienceAdvances2021	
PUC001	ES 2000 1200 BA	Villena, Alicante	Spain	1879	1693	3461	19	F	H3		SE Iberia BA Valencian	Villalba-MoucoScienceAdvances2021	
PUC002	ES 2000 1200 BA	Villena, Alicante	Spain	1741	1566	3371	19	M	H1a1	R1b1a1b1a1a2	SE Iberia BA Valencian	Villalba-MoucoScienceAdvances2021	
PUC003	ES 2000 1200 BA	Villena, Alicante	Spain	1882	1746	3486	20	F	USb2a+16192		SE Iberia BA Valencian	Villalba-MoucoScienceAdvances2021	
PUC004	ES 2000 1200 BA	Villena, Alicante	Spain	1748	1625	3407	19	F	V		SE Iberia BA Valencian	Villalba-MoucoScienceAdvances2021	
ALM031	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				F	J2b1a			Almoloya Argar 1stdegree relative	Villalba-MoucoScienceAdvances2021	
ALM024	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1954	1772	3544	24	M	K1a195	R1b1a1b1a1a2	Almoloya Argar 1stdegree relative	Villalba-MoucoScienceAdvances2021	
ALM058	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	USb11f	R1b1a1b1a1a2		Almoloya Argar 1stdegree relative	Villalba-MoucoScienceAdvances2021	
ALM060	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain	1878	1644	3444	23	F	USa1+16192		Almoloya Argar 1stdegree relative	Villalba-MoucoScienceAdvances2021	
ALM080	ES 2000 1200 BA	La Almoloya, Piñero, Murcia	Spain				M	USb1a	R1b1a1b1a1a2		Almoloya Argar 1stdegree relative	Villalba-MoucoScienceAdvances2021	
ES0000	ES 2000 1200 BA	Cueva de los Tucos, La Alfranca, Murcia	Spain				M	H1a1			ES 2000 1200 BA	OldaleNature2018	
EHJ001	ES 2500 2000 CA	El Hundido, Monasterio de Rodilla, Burgos, Castilla y León	Spain	2287	2041	3760	30	M	USa1b1	R1b1a1a2a1a2f(R1b1a1a2a1a2a5)	C Iberia CA_Stp	OldaleScience2019	
EHJ002	ES 2500 2000 CA	El Hundido, Monasterio de Rodilla, Burgos, Castilla y León	Spain	2564	2299	3933	32	M	K1a4a1	R1b1a1a2a1a2f(R1b1a1a2a1a2a5)	C Iberia CA_Stp	OldaleScience2019	
IA0459	ES 2500 2000 CA	Arroyal I, Alfoz de Quintanadueñas, Burgos, Castilla y León	Spain				F	USb2b			C Iberia CA_Stp	OldaleNature2018	
IA0460	ES 2500 2000 CA	Arroyal I, Alfoz de Quintanadueñas, Burgos, Castilla y León	Spain	2460	2206	3860	30	F	H1a1	Ia2a		C Iberia CA_Stp	OldaleNature2018
IA0461	ES 2500 2000 CA	Arroyal I, Alfoz de Quintanadueñas, Burgos, Castilla y León	Spain	2463	2201	3837	25	F	H1a1			C Iberia CA_Stp	OldaleNature2018
12328	ES 2500 2000 CA	Cueva de la Paloma, Soto de las Requejas, Asturias	Spain				M	H1a152			NW Iberia CA_Stp	OldaleScience2019	
IA4247	ES 2500 2000 CA	C Camino de las Yegaras, Community of Madrid	Spain	2462	2209	3870	30	M	T2a			C Iberia CA_Stp	OldaleNature2018
15665	ES 2500 2000 CA	Virzaçal, Tabla de Rodón, Burgos, Castilla y León	Spain	2283	1983	3730	40	M	K1a24a	R1b1a1a2a1a2	C Iberia CA_Stp	OldaleNature2018	
16539	ES 2500 2000 CA	Humanejo, Parla, Community of Madrid	Spain				M	T2b3+151	R1b1a1a2a1a2f(R1b1a1a2a1a2c)		C Iberia CA_Stp	OldaleNature2018	
1623	ES 2500 2000 CA	C Camino de las Yegaras, San Fernando de Henares, Community of Madrid	Spain				F	USb11a			C Iberia CA_Stp	OldaleNature2018	
BI0180	FR 2000 1200 BA	Muehlacker, Grand-Est (East), Bischwiller	France				F	USb11a			France Early Bronze Age A2a	BrunePNAS2020	
BI0385	FR 2000 1200 BA	Muehlacker, Grand-Est (East), Bischwiller	France				F	T2c1a2a	</				

126774	HR 1800 1600 MBA	Mulego/Mon Sego	Croatia					F	J1c2		Croatia MBA	PattersonNature2021
126893	HR 1800 1600 MBA	Mulego/Mon Sego	Croatia					F	U8b1b1		Croatia MBA	PattersonNature2021
JAG06	HR 1800 1600 MBA	Jagodnjak-Krčevine	Croatia					M	T2b11	G2a2a1a2a2a1	Croatia Jag MBA	FreilichScientificReports2021
JAG58	HR 1800 1600 MBA	Jagodnjak-Krčevine	Croatia					M	T2b11	G2a2a1a2a2a1	Croatia Jag MBA	FreilichScientificReports2021
JAG78	HR 1800 1600 MBA	Jagodnjak-Krčevine	Croatia					M	U5b1a1a	G2a2a1a2a2a1	Croatia Jag MBA	FreilichScientificReports2021
JAG82	HR 1800 1600 MBA	Jagodnjak-Krčevine	Croatia					M	U2e1a1	G2a2a1a2a2a1	Croatia Jag MBA	FreilichScientificReports2021
JAG85	HR 1800 1600 MBA	Jagodnjak-Krčevine	Croatia					F	K1a1b1		Croatia Jag MBA	FreilichScientificReports2021
JAG93	HR 1800 1600 MBA	Jagodnjak-Krčevine	Croatia					F	U5a1g		Croatia Jag MBA	FreilichScientificReports2021
12364	HU 2500 2000 BBC	Budapest-Békašmegyey	Hungary	2284	2059	3779	28	M	U5a2b	H2	Beaker Central Europe	MathiesonNature2015
12365	HU 2500 2000 BBC	Budapest-Békašmegyey	Hungary	2459	2206	3858	32	M	V3	R1b1a1a2a1a2b1	Beaker Central Europe	MathiesonNature2015
12741	HU 2500 2000 BBC	Szigetszentmiklós,Felső Úrge-hegyi dűlő	Hungary	2456	2150	3835	35	M	H1-16189	I2a1a1	Beaker Central Europe	MathiesonNature2015
12786	HU 2500 2000 BBC	Szigetszentmiklós,Felső Úrge-hegyi dűlő	Hungary	2457	2204	3850	35	M	I1	I2a2a	Beaker Central Europe	MathiesonNature2015
12787	HU 2500 2000 BBC	Szigetszentmiklós,Felső Úrge-hegyi dűlő	Hungary	2456	2200	3840	35	M	T2b	R1b1a1a2a2	Beaker Central Europe	MathiesonNature2015
13529	HU 2500 2000 BBC	Budakalász, Csajkéske (M0 Site 12)	Hungary					F	I3a		Beaker Central Europe	MathiesonNature2015
14178	HU 2500 2000 BBC	Szigetszentmiklós,Felső Úrge-hegyi dűlő	Hungary					M	J1c1b1a	R1b1a1a2	Beaker Central Europe	OlaideNature2018
17044	HU 2500 2000 BBC	Szigetszentmiklós-Událódor	Hungary					M	U5b1a1b	R1b1a1a2a1c1	Beaker Central Europe	LazaridisNature2016
17045	HU 2500 2000 BBC	Szigetszentmiklós-Událódor	Hungary					F	T2c1a152		Beaker Central Europe	LazaridisNature2016
BR0003	IT 2200 1500 BA	Grottna dei Cavoloni del Broion	Italy					M	U4a2f	R1b1'5-P312	Italy Broion BA	SaueCurrentBiology2020
BR0007	IT 2200 1500 BA	Grottna dei Cavoloni del Broion	Italy					F	K1a1b1		Italy Broion BA	SaueCurrentBiology2020
BR0010	IT 2200 1500 BA	Grottna dei Cavoloni del Broion	Italy					F	H610291T		Italy Broion EBA	SaueCurrentBiology2020
131442	IT 2200 1500 BA	Sicily, Buffa	Italy					M	K1a3a	R1b1a1a2a1a2 (xR1b1a1a2a1a2c)	Sicily EBA	FernandesNatureEcologyEvolution2020
131422	IT 2200 1500 BA	Sicily, Buffa	Italy					F	T2b		Sicily EBA	FernandesNatureEcologyEvolution2020
13123	IT 2200 1500 BA	Sicily, Buffa	Italy					M	U8b1b1	R1b1a1a2a1a2a1	Sicily EBA	FernandesNatureEcologyEvolution2020
13124	IT 2200 1500 BA	Sicily, Buffa	Italy					M	H1-16189	R1b1a1a2a1a2 (xR1b1a1a2a1a2c)	Sicily EBA	FernandesNatureEcologyEvolution2020
13125	IT 2200 1500 BA	Sicily, Buffa	Italy					M	K1a3a	G2a2b2a1a1c1	Italy Sicily MBA o2	FernandesNatureEcologyEvolution2020
13878	IT 2200 1500 BA	Sicily, Marcita	Italy	1382	1134	3015	20	F	K1a+195C1		Sicily LBA	FernandesNatureEcologyEvolution2020
18561	IT 2200 1500 BA	Sicily, Isello	Italy	2399	2153	3825	20	M	K1b1+16093C1	R1b1a1a2a1a2a1	Sicily EBA5561	FernandesNatureEcologyEvolution2020
13123	IT 2200 1500 BA	Sicily, Buffa	Italy					M	U8b1b1	R1b1a1a2a1a2	Italy Sicily EBA	FernandesNatureEcologyEvolution2020
115979	IT 2500 2000 BBC	Via Guidorossi, Parma, Emilia Romagna	Italy					F	T2b3		Beaker Northern Italy	OlaideNature2018
12477	IT 2500 2000 BBC	Via Guidorossi, Parma, Emilia Romagna	Italy					F	T2b3+151		Beaker Northern Italy	OlaideNature2018
12478	IT 2500 2000 BBC	Via Guidorossi, Parma, Emilia Romagna	Italy	2196	1939	3671	40	M	K1a2a	R1b1a1a2a1a2	Beaker Northern Italy	OlaideNature2018
113229	NL 1600 1100 BA	Noord-Holland, Hoogkarspel-Houterpolder-West	Netherlands					F	T2a1		Netherlands MBA LBA	PattersonNature2021
111971	NL 1600 1100 BA	Noord-Holland, Westwoud-Binnenwijdend	Netherlands	1210	1015	2915	30	F	H3		Netherlands LBA	PattersonNature2021
111972	NL 1600 1100 BA	Noord-Holland, Westwoud-Binnenwijdend	Netherlands	1501	1319	3155	30	M	U5b2b1	R1b1a1b1a1a1c	Netherlands MBA	PattersonNature2021
111973	NL 1600 1100 BA	Noord-Holland, Westwoud-Binnenwijdend	Netherlands	1415	1230	3065	30	F	H5a1		Netherlands MBA	PattersonNature2021
112081	NL 1600 1100 BA	Noord-Holland, Hoogkarspel-Houterpolder-West	Netherlands	1391	1221	3040	20	F	U5a1c		Netherlands MBA	PattersonNature2021
112082	NL 1600 1100 BA	Noord-Holland, Hoogkarspel-Houterpolder-West	Netherlands					F	K1d		Netherlands MBA	PattersonNature2021
112083	NL 1600 1100 BA	Noord-Holland, Westwoud-Merkenwaardweg	Netherlands	1260	1048	2945	30	F	J2b1a1		Netherlands MBA LBA	PattersonNature2021
117019	NL 1600 1100 BA	Zuid-Holland, Vlaardingen-Kralbeek	Netherlands	1421	1216	3060	40	M	T2b		Netherlands MBA	PattersonNature2021
120062	NL 1600 1100 BA	Noord-Holland, Bovenkarspel-Het Valkje	Netherlands	920	806	2710	35	F	H1-16189	R1b1a1b1a1a1c1a	Netherlands LBA	PattersonNature2021
126828	NL 1600 1100 BA	Noord-Holland, Wervershoof-Zwaagdijk	Netherlands					F	H2a2		Netherlands MBA	PattersonNature2021
126829	NL 1600 1100 BA	Noord-Holland, Wervershoof-Zwaagdijk	Netherlands					M	H50	R1b1a1b1a	Netherlands MBA	PattersonNature2021
126830	NL 1600 1100 BA	Noord-Holland, Wervershoof-Zwaagdijk	Netherlands	1620	1311	3210	60	M	H2a+16093	R1b1a1b1a1a2c1	Netherlands MBA	PattersonNature2021
126831	NL 1600 1100 BA	Noord-Holland, Wervershoof-Zwaagdijk	Netherlands					F	U5a1a1		Netherlands MBA	PattersonNature2021
126832	NL 1600 1100 BA	Noord-Holland, Wervershoof-Zwaagdijk	Netherlands					F	U4a3		Netherlands MBA	PattersonNature2021
113022	NL 2500 2000 BBC	Zuid-Holland, Molensangraaf	Netherlands	2136	1892	3635	40	M	K2b1a	R1b1a1b1a1a	Netherlands LNB EBA, BellBeaker	PattersonNature2021
113027	NL 2500 2000 BBC	Zuid-Holland, Molensangraaf	Netherlands	2127	1983	3700	25	F	T2b1		Netherlands LNB EBA, BellBeaker	PattersonNature2021
113028	NL 2500 2000 BBC	Zuid-Holland, Oltoland-Kromme Elieboog	Netherlands					M	H1a	R1b1a1b1a1a2	Netherlands LNB, BellBeaker	PattersonNature2021
14068	NL 2500 2000 BBC	De Tuilthorn, Oostwoud, Noord-Holland	The Netherlands	2134	1950	3655	20	M	U5a2a1	R1b1a1a2a1a2	Beaker The Netherlands	OlaideNature2018
14069	NL 2500 2000 BBC	De Tuilthorn, Oostwoud, Noord-Holland	The Netherlands	2192	1887	3640	50	M	U5a1a1	R1b1a1a2a1a2	Beaker The Netherlands	OlaideNature2018
14073	NL 2500 2000 BBC	De Tuilthorn, Oostwoud, Noord-Holland	The Netherlands	2197	1897	3660	50	M	U5a2b3	R1b1a1a2a1a2	Beaker The Netherlands	OlaideNature2018
14074	NL 2500 2000 BBC	De Tuilthorn, Oostwoud, Noord-Holland	The Netherlands	2281	1899	3690	60	M	H	R1b1a1a2a1a2	Beaker The Netherlands	OlaideNature2018
14075	NL 2500 2000 BBC	De Tuilthorn, Oostwoud, Noord-Holland	The Netherlands	2127	1933	3635	20	F	H5a1		Beaker The Netherlands	OlaideNature2018
14076	NL 2500 2000 BBC	De Tuilthorn, Oostwoud, Noord-Holland	The Netherlands	1883	1747	3490	20	F	K1d		Beaker The Netherlands	OlaideNature2018
15748	NL 2500 2000 BBC	De Tuilthorn, Oostwoud, Noord-Holland	The Netherlands	2579	2211	3945	55	M	X2b4	R1b1a1a2a1a2	Beaker The Netherlands	OlaideNature2018
16531	PO 2400 2000 BBC	Dzielnica	Poland	2288	2037	3755	35	M	U5b2b1a1	R1a1a	Poland BA	OlaideNature2018
16537	PO 2400 2000 BBC	Racibórz-Stara Wieś	Poland	2293	2039	3765	35	M	I1a1	R1b1a1a2a1a	Poland BA	OlaideNature2018
16576	PO 2400 2000 BBC	Janów	Poland	2340	2042	3780	35	F	K1a1a1b2b		Poland BA	OlaideNature2018
16580	PO 2400 2000 BBC	Janów Śląski	Poland					M	K1a1b2b	R1b1a1a2a1a2b1	Beaker Central Europe	MathiesonNature2015
16581	PO 2400 2000 BBC	Kornice	Poland	2454	2145	3825	35	M	H6a1a	R1b1a1a2a1a2b1	Beaker Central Europe	LazaridisNature2016
16582	PO 2400 2000 BBC	Kornice	Poland	2395	2051	3790	35	F	U5a2b		Beaker Central Europe	MathiesonNature2015
16583	PO 2400 2000 BBC	Zemik Wierkie	Poland	2291	2047	3770	30	F	U5b1a1b		Beaker Central Europe	LazaridisNature2016
pcw260	PO 2400 2000 BBC	Malopolska Upland	Poland	2455	2147	3830	35	F	T2b		Poland BBC	LinderholmScientificReports2020
pcw280	PO 2400 2000 BBC	Malopolska Upland	Poland					F	U5a1		Poland BBC	LinderholmScientificReports2020
N45	PO 2800 2300 CWC	Pikulowo	Poland					F	H5a8		Poland CWC	FernandesScientificReports2018
N47	PO 2800 2300 CWC	Pikulowo	Poland					M	U5b2a2	I2a2a	Poland CWC	FernandesScientificReports2018
pcw040	PO 2800 2300 CWC	Rzeszow Foothills	Poland	2571	2305	3950	35	M	U5b2b1a1	R1b1a1a2	Poland CWC	LinderholmScientificReports2020
pcw041	PO 2800 2300 CWC	Rzeszow Foothills	Poland	2571	2305	3950	35	M	J1c2c	R1b1a1a2	Poland CWC	LinderholmScientificReports2020
pcw061	PO 2800 2300 CWC	Rzeszow Foothills	Poland	2469	2211	3890	35	F	H7a		Poland CWC	LinderholmScientificReports2020
pcw070	PO 2800 2300 CWC	Rzeszow Foothills	Poland	2469	2211	3890	35	M	I2	R1b1a1a2	Poland CWC	LinderholmScientificReports2020
pcw350	PO 2800 2300 CWC	Sokar Ridge	Poland	2462	2206	3865	35	M	T1a1	R1b1a1a2	Poland CWC	LinderholmScientificReports2020
pcw361	PO 2800 2300 CWC	Sokar Ridge	Poland					M	K1a1b	R1b1a1a2a1a	Poland CWC	LinderholmScientificReports2020
pcw382	PO 2800 2300 CWC	Sokar Ridge	Poland	2465	2209	3875	35	M	U5a1a2a	R1b1a1a2a1a	Poland CWC	LinderholmScientificReports2020
TV3032	PT 1800 1400 MBA	Torre Velha 3	Portugal					M	X2b+226	R1b1a2	Portugal MBA	MartiniandPLoGenetics2017
TV3831	PT 1800 1400 MBA	Torre Velha 3	Portugal					M	H1+152	R1b1a2a1a2	Portugal MBA	MartiniandPLoGenetics2017
V010209	PT 1800 1400 MBA	Monte do Vale do Ouro 3	Portugal					F	U5a1		Portugal MBA	MartiniandPLoGenetics2017
11570	PT 2500 2000 CA	Verdeira dos Ruives, Vialonga, Lisboa	Portugal	2876	2491	4100	60	M	X2b+226	I	SW Iberia CA	OlaideNature2018
14229	PT 2500 2000 CA	Cova da Moura, Torres Vedras, Lisboa	Portugal	2290	2063	3775	25	M	U5a2+16362	I2a1a1	SW Iberia CA	OlaideNature2018

Table S7. qpAdm admixture modelling of each individual

ID	Group label (in this study)	Yamnaya proportion	ANF proportion	HG proportion	Yamnaya proportion (se)	ANF proportion (se)	HG proportion (se)	p-value	qpAdm_SNP
Bre445Y	FR 2500 2000 Bre	0.735	0.147	0.119	0.131	0.087	0.081	0.493598023	238045
Bre445F	FR 2500 2000 Bre	0.367	0.498	0.135	0.061	0.042	0.036	0.58362957	477024
Bre445H	FR 2500 2000 Bre	0.174	0.619	0.207	0.08	0.066	0.033	0.270925153	344246
Bre445A	FR 2500 2000 Bre	0	0.749	0.251	0	0.034	0.034	0.457258955	271380
Bre445B	FR 2500 2000 Bre	0	0.787	0.213	0	0.036	0.036	0.060515065	262739
Bre445C	FR 2500 2000 Bre	0	0.779	0.221	0	0.034	0.034	0.071216404	319221
Bre445D	FR 2500 2000 Bre	0	0.758	0.242	0	0.032	0.032	0.337905812	322096
Bre445E	FR 2500 2000 Bre	0	0.808	0.192	0	0.033	0.033	0.440742487	329382
SM6B54	CH 2800 2200 CWC	0.444	0.418	0.138	0.069	0.048	0.042	0.935126886	427280
MX188	CH 2800 2200 CWC	0.532	0.285	0.184	0.084	0.048	0.048	0.080657128	299495
MX191	CH 2800 2200 CWC	0.613	0.249	0.138	0.082	0.059	0.048	0.515390131	284377
MX196	CH 2800 2200 CWC	0.544	0.318	0.083	0.061	0.047	0.047	0.45917524	272099
MX197	CH 2800 2200 CWC	0.604	0.286	0.111	0.079	0.056	0.047	0.356667363	263483
MX198	CH 2800 2200 CWC	0.526	0.293	0.081	0.063	0.05	0.05	0.454648956	262783
MX199	CH 2800 2200 CWC	0.564	0.321	0.115	0.079	0.058	0.048	0.589765708	283774
MX195	CH 2800 2200 CWC	0.517	0.369	0.114	0.079	0.06	0.05	0.835256853	250170
MX192	CH 2800 2200 CWC	0.504	0.358	0.138	0.078	0.06	0.045	0.825715814	240618
MX190	CH 2800 2200 CWC	0.703	0.283	0.014	0.084	0.064	0.052	0.022208856	243155
Aes425	CH 2800 2200 CWC	0.529	0.293	0.081	0.063	0.059	0.05	0.197226905	297708
MX298	CH 2800 2200 CWC	0.38	0.404	0.216	0.074	0.053	0.045	0.790575161	284599
SK21	CH 2800 2200 FN	0	0.915	0.085	0	0.034	0.034	0.559050298	273258
SK26	CH 2800 2200 FN	0	0.815	0.185	0	0.036	0.036	0.339561422	233400
SK17	CH 2800 2200 FN	0	0.849	0.151	0	0.037	0.037	0.094121022	123745
SK10	CH 2800 2200 FN	0	0.791	0.209	0	0.036	0.036	0.380186939	208147
KNE002	CZ 2200 1700 Ufnetic_EBA	0.534	0.278	0.188	0.086	0.066	0.048	0.816427181	269764
KNE003	CZ 2200 1700 Ufnetic_EBA	0.485	0.356	0.159	0.076	0.055	0.045	0.667834621	351593
KNE006	CZ 2200 1700 Ufnetic_EBA	0.498	0.333	0.169	0.131	0.093	0.076	0.990712283	110177
KNE007	CZ 2200 1700 Ufnetic_EBA	0.449	0.363	0.168	0.083	0.063	0.05	0.338466113	246584
KO1006	CZ 2200 1700 Ufnetic_EBA	0.678	0.236	0.086	0.076	0.055	0.046	0.697532422	332551
KO1007	CZ 2200 1700 Ufnetic_EBA	0.63	0.212	0.158	0.076	0.056	0.046	0.717653971	282981
KO1008	CZ 2200 1700 Ufnetic_EBA	0.476	0.418	0.106	0.094	0.065	0.054	0.23140039	225839
KO1009	CZ 2200 1700 Ufnetic_EBA	0.482	0.364	0.155	0.08	0.062	0.046	0.04531609	289348
KO1010	CZ 2200 1700 Ufnetic_EBA	0.498	0.321	0.167	0.081	0.067	0.05	0.439127154	284309
KO1011	CZ 2200 1700 Ufnetic_EBA	0.587	0.282	0.131	0.097	0.07	0.055	0.174115285	225305
KO1013	CZ 2200 1700 Ufnetic_EBA	0.503	0.373	0.124	0.077	0.06	0.044	0.76382955	278955
KO1014	CZ 2200 1700 Ufnetic_EBA	0.508	0.293	0.199	0.085	0.063	0.049	0.344304547	248868
KO1015	CZ 2200 1700 Ufnetic_EBA	0.562	0.266	0.132	0.08	0.06	0.047	0.909932416	238825
KO1001	CZ 2200 1700 Ufnetic_EBA	0.549	0.288	0.163	0.085	0.063	0.049	0.362672132	261004
KO7002	CZ 2200 1700 Ufnetic_EBA	0.623	0.312	0.065	0.093	0.067	0.054	0.14111294	230223
MIB001	CZ 2200 1700 Ufnetic_EBA	0.564	0.269	0.167	0.087	0.061	0.053	0.52077908	234363
MIB002	CZ 2200 1700 Ufnetic_EBA	0.559	0.284	0.157	0.083	0.059	0.047	0.557878657	317800
MIB003	CZ 2200 1700 Ufnetic_EBA	0.538	0.338	0.145	0.115	0.08	0.069	0.326046589	292538
MIB004	CZ 2200 1700 Ufnetic_EBA	0.397	0.415	0.188	0.084	0.062	0.049	0.922203603	236575
MIB005	CZ 2200 1700 Ufnetic_EBA	0.421	0.335	0.244	0.096	0.072	0.051	0.28988763	202845
MIB024	CZ 2200 1700 Ufnetic_EBA	0.285	0.416	0.299	0.134	0.087	0.08	0.320920972	132256
MIB028	CZ 2200 1700 Ufnetic_EBA	0.615	0.302	0.083	0.074	0.053	0.045	0.780949639	352789
MIB034	CZ 2200 1700 Ufnetic_EBA	0.577	0.31	0.121	0.073	0.053	0.043	0.85528646	360881
MIG010	CZ 2200 1700 Ufnetic_EBA	0.577	0.271	0.153	0.112	0.079	0.061	0.587641053	169314
MIG011	CZ 2200 1700 Ufnetic_EBA	0.441	0.395	0.164	0.081	0.061	0.051	0.510904554	221714
MIG012	CZ 2200 1700 Ufnetic_EBA	0.502	0.329	0.168	0.1	0.073	0.06	0.517850044	170786
MIS001	CZ 2200 1700 Ufnetic_EBA	0.601	0.269	0.09	0.071	0.052	0.042	0.480272771	303548
MIS002	CZ 2200 1700 Ufnetic_EBA	0.402	0.39	0.207	0.078	0.058	0.045	0.269345615	380307
MIS004	CZ 2200 1700 Ufnetic_EBA	0.523	0.367	0.11	0.107	0.077	0.065	0.4555386	157014
MIS006	CZ 2200 1700 Ufnetic_EBA	0.523	0.384	0.093	0.135	0.104	0.083	0.600109823	96321
PM003	CZ 2200 1700 Ufnetic_EBA	0.592	0.316	0.092	0.094	0.067	0.056	0.165137847	201288
PM004	CZ 2200 1700 Ufnetic_EBA	0.577	0.327	0.127	0.089	0.069	0.056	0.161166305	192508
PM006	CZ 2200 1700 Ufnetic_EBA	0.749	0.134	0.117	0.099	0.072	0.061	0.391170001	188675
PM009	CZ 2200 1700 Ufnetic_EBA	0.611	0.048	0.14	0.1	0.083	0.061	0.554101324	103417
PM010	CZ 2200 1700 Ufnetic_EBA	0.577	0.286	0.138	0.083	0.063	0.05	0.638449406	224117
PM012	CZ 2200 1700 Ufnetic_EBA	0.73	0.26	0.01	0.108	0.077	0.06	0.723883614	168825
ROU001	CZ 2200 1700 Ufnetic_EBA	0.452	0.354	0.087	0.084	0.061	0.049	0.548885792	258226
ROU005	CZ 2200 1700 Ufnetic_EBA	0.594	0.26	0.146	0.079	0.059	0.046	0.46526554	297974
VL039	CZ 2200 1700 Ufnetic_EBA	0.539	0.371	0.09	0.074	0.051	0.047	0.281523806	303721
VL042	CZ 2200 1700 Ufnetic_EBA	0.618	0.276	0.107	0.11	0.078	0.069	0.80945077	134345
VL045	CZ 2200 1700 Ufnetic_EBA	0.523	0.382	0.194	0.115	0.086	0.069	0.062727709	299775
VL046	CZ 2200 1700 Ufnetic_EBA	0.387	0.503	0.11	0.174	0.123	0.097	0.018818312	95952
VL047	CZ 2200 1700 Ufnetic_EBA	0.762	0.164	0.075	0.092	0.067	0.055	0.638139398	242464
VL050	CZ 2200 1700 Ufnetic_EBA	0.356	0.367	0.277	0.074	0.054	0.046	0.348968305	321182
VL058	CZ 2200 1700 Ufnetic_EBA	0.507	0.281	0.212	0.087	0.066	0.048	0.742495449	228441
VL060	CZ 2200 1700 Ufnetic_EBA	0.526	0.326	0.39	0.083	0.063	0.055	0.449154582	231106
VL061	CZ 2200 1700 Ufnetic_EBA	0.683	0.194	0.123	0.077	0.057	0.046	0.935441404	319082
VL065	CZ 2200 1700 Ufnetic_EBA	0.602	0.352	0.046	0.09	0.068	0.053	0.599139604	199781
4885	CZ 2500 2000 BBC	0.341	0.55	0.109	0.072	0.052	0.043	0.060898245	359421
4886	CZ 2500 2000 BBC	0.54	0.37	0.076	0.076	0.051	0.044	0.145026537	371264
4887	CZ 2500 2000 BBC	0.393	0.465	0.142	0.071	0.053	0.043	0.641669352	362513
4888	CZ 2500 2000 BBC	0.457	0.421	0.122	0.07	0.051	0.041	0.73624549	461185
4889	CZ 2500 2000 BBC	0.652	0.295	0.053	0.07	0.05	0.042	0.894206919	459955
4891	CZ 2500 2000 BBC	0.469	0.393	0.138	0.075	0.055	0.042	0.146119294	371293
4895	CZ 2500 2000 BBC	0.439	0.346	0.127	0.086	0.063	0.043	0.506848411	360819
4896	CZ 2500 2000 BBC	0.349	0.483	0.168	0.059	0.05	0.041	0.545037405	363522
5514	CZ 2500 2000 BBC	0.507	0.311	0.182	0.072	0.051	0.043	0.338692481	378025
7205	CZ 2500 2000 BBC	0.532	0.3	0.167	0.074	0.054	0.044	0.049522031	342924
7210	CZ 2500 2000 BBC	0.492	0.358	0.15	0.072	0.053	0.044	0.431854047	334858
7211	CZ 2500 2000 BBC	0.552	0.333	0.143	0.083	0.059	0.043	0.184073809	329226
7212	CZ 2500 2000 BBC	0.443	0.414	0.142	0.072	0.051	0.043	0.08762187	474913
7213	CZ 2500 2000 BBC	0.436	0.323	0.241	0.082	0.055	0.048	0.001802048	337528
7214	CZ 2500 2000 BBC	0.469	0.411	0.12	0.073	0.053	0.043	0.781962873	342585
7249	CZ 2500 2000 BBC	0.568	0.348	0.21	0.072	0.052	0.041	0.998155303	343478
7251	CZ 2500 2000 BBC	0.61	0.31	0.08	0.075	0.057	0.045	0.51442824	336920
7269	CZ 2500 2000 BBC	0.388	0.454	0.158	0.082	0.059	0.045	0.245015924	316374
7275	CZ 2500 2000 BBC	0.422	0.486	0.092	0.077	0.055	0.047	0.054723974	318194
7276	CZ 2500 2000 BBC	0.601	0.319	0.079	0.073	0.053	0.044	0.37000732	340607
7278	CZ 2500 2000 BBC	0.517	0.365	0.118	0.085	0.061	0.041	0.418187123	475057
7281	CZ 2500 2000 BBC	0.562	0.264	0.173	0.077	0.054	0.045	0.819255885	338558
7282	CZ 2500 2000 BBC	0.414	0.401	0.185	0.074	0.055	0.043	0.322373865	348317
7283	CZ 2500 2000 BBC	0.168	0.616	0.216	0.076	0.056	0.044	0.476066673	336470
7286	CZ 2500 2000 BBC	0.455	0.344	0.2	0.072	0.05	0.042	0.017501709	480899
7289	CZ 2500 2000 BBC	0.527	0.406	0.096	0.076	0.056	0.043	0.094045544	342133
7290	CZ 2500 2000 BBC	0.416	0.366	0.217	0.071	0.051	0.041	0.808908239	358818
KO1003	CZ 2500 2000 BBC	0.354	0.495	0.15	0.072	0.051	0.043	0.775274336	381716
KO1004	CZ 2500 2000 BBC	0.486	0.358	0.155	0.078	0.055	0.045	0.99839776	365188
KO1005	CZ 2500 2000 BBC	0.438	0.447	0.173	0.084	0.067	0.047	0.685376532	321212
KOP001	CZ 2500 2000 BBC	0.244	0.602	0.083	0.082	0.064</			

ATI1_119	DE 2000 1500 EBA	0.465	0.466	0.069	0.079	0.059	0.049	0.676291653	265181
ATI1_2	DE 2000 1500 EBA	0.495	0.497	0.108	0.121	0.043	0.044	0.144653189	247590
ATI1_33	DE 2000 1500 EBA	0.306	0.324	0.17	0.074	0.056	0.043	0.015916121	311929
ATI1_43	DE 2000 1500 EBA	0.258	0.587	0.155	0.07	0.052	0.041	0.053630845	388959
ATI1_50	DE 2000 1500 EBA	0.528	0.331	0.141	0.11	0.083	0.058	0.101834549	175306
ATI1_70	DE 2000 1500 EBA	0.773	0.578	0.149	0.097	0.057	0.054	0.064442874	199528
ATI1_72	DE 2000 1500 EBA	0.292	0.579	0.073	0.073	0.054	0.044	0.001702293	327318
ATI1_98	DE 2000 1500 EBA	0.304	0.562	0.134	0.092	0.069	0.053	0.252556758	182778
DBKR_80	DE 2000 1500 EBA	0.536	0.403	0.061	0.076	0.054	0.046	0.442448616	340298
POST_28	DE 2000 1500 EBA	0.464	0.452	0.084	0.088	0.061	0.052	0.696791013	233517
POST_5	DE 2000 1500 EBA	0.433	0.421	0.046	0.041	0.046	0.04	0.078570917	118238
POST_99	DE 2000 1500 EBA	0.477	0.398	0.125	0.098	0.074	0.051	0.27730336	201994
WEHR_1415aJut	DE 2000 1500 EBA	0.461	0.385	0.154	0.087	0.059	0.052	0.034069992	248934
WEHR_1586	DE 2000 1500 EBA	0.489	0.375	0.137	0.084	0.062	0.053	0.314470236	200989
MK252	DE 2000 1500 EBA	0.412	0.5	0.087	0.086	0.063	0.05	0.177798439	209752
MK254	DE 2000 1500 EBA	0.3	0.543	0.158	0.082	0.059	0.046	0.057077245	287194
MK257	DE 2000 1500 EBA	0.533	0.392	0.075	0.098	0.071	0.056	0.256433595	204653
MK286	DE 2000 1500 EBA	0.452	0.399	0.149	0.084	0.062	0.049	0.054190953	246861
MK270	DE 2000 1500 EBA	0.494	0.42	0.085	0.1	0.073	0.058	0.448819128	165219
MK279	DE 2000 1500 EBA	0.335	0.397	0.268	0.092	0.071	0.055	0.413939399	179404
MK280	DE 2000 1500 EBA	0.418	0.385	0.198	0.082	0.052	0.045	0.324541435	275843
MK275	DE 2000 1500 EBA	0.42	0.448	0.132	0.09	0.067	0.049	0.012168476	290511
MK283	DE 2000 1500 EBA	0.203	0.607	0.19	0.076	0.054	0.046	0.223466124	297554
MK277	DE 2000 1500 EBA	0.331	0.574	0.095	0.085	0.065	0.049	0.636381866	202058
MUGO_188	DE 2000 1500 BBC	0.754	0.713	0.073	0.086	0.077	0.06	0.197008935	186554
UNTASS_68S12	DE 2500 2000 BBC	0.241	0.638	0.121	0.081	0.059	0.045	0.66712027	265258
UNTASS_1343	DE 2500 2000 BBC	0.529	0.353	0.118	0.073	0.053	0.043	0.303148544	365222
WEHR_11925KA	DE 2500 2000 BBC	0.512	0.346	0.142	0.078	0.056	0.045	0.01951286	419239
I0108	DE 2500 2000 BBC	0.46	0.367	0.174	0.08	0.056	0.048	0.998989597	337776
I0112	DE 2500 2000 BBC	0.441	0.341	0.127	0.076	0.054	0.04	0.388621288	387708
I0113	DE 2500 2000 BBC	0.326	0.438	0.236	0.085	0.061	0.049	0.166096993	275147
I3588	DE 2500 2000 BBC	0.441	0.468	0.091	0.091	0.064	0.053	0.042568221	208867
I3589	DE 2500 2000 BBC	0.378	0.398	0.224	0.092	0.062	0.052	0.021165599	253625
I3590	DE 2500 2000 BBC	0.4	0.455	0.144	0.091	0.063	0.053	0.288497215	205361
I3594	DE 2500 2000 BBC	0.27	0.576	0.154	0.075	0.054	0.043	0.343765533	287207
I3601	DE 2500 2000 BBC	0.54	0.336	0.124	0.076	0.055	0.045	0.878339486	302634
I5519	DE 2500 2000 BBC	0.437	0.353	0.21	0.071	0.052	0.043	0.344921141	324295
I5521	DE 2500 2000 BBC	0.581	0.31	0.11	0.081	0.056	0.046	0.106831202	309458
I5833	DE 2500 2000 BBC	0.487	0.371	0.143	0.075	0.054	0.04	0.136499522	307128
I5834	DE 2500 2000 BBC	0.425	0.48	0.094	0.079	0.059	0.045	0.131924876	283473
I6590	DE 2500 2000 BBC	0.409	0.391	0.2	0.073	0.054	0.044	0.375792487	388600
I6591	DE 2500 2000 BBC	0.394	0.472	0.134	0.069	0.05	0.041	0.596459913	474370
I9335	DE 2500 2000 BBC	0.441	0.53	0.129	0.068	0.048	0.041	0.891624759	488261
I12208	DE 2500 1200 BA	0.078	0.078	0.107	0.074	0.051	0.04	0.715220826	383775
ALM001	ES 2000 1200 BA	0.089	0.713	0.199	0.089	0.063	0.05	0.283866338	230927
ALM002	ES 2000 1200 BA	0.255	0.604	0.141	0.081	0.06	0.046	0.828592823	249136
ALM003	ES 2000 1200 BA	0.177	0.669	0.154	0.079	0.057	0.047	0.110502072	271866
ALM006	ES 2000 1200 BA	0.45	0.465	0.166	0.074	0.054	0.045	0.424809684	346234
ALM007	ES 2000 1200 BA	0.048	0.752	0.142	0.083	0.06	0.047	0.781881341	236425
ALM008	ES 2000 1200 BA	0.167	0.636	0.198	0.072	0.053	0.042	0.332690822	325432
ALM015	ES 2000 1200 BA	0.211	0.63	0.158	0.076	0.058	0.043	0.217793809	305793
ALM016	ES 2000 1200 BA	0.185	0.616	0.2	0.075	0.055	0.043	0.262361085	368645
ALM017	ES 2000 1200 BA	0.198	0.669	0.199	0.077	0.055	0.042	0.076924701	307838
ALM019	ES 2000 1200 BA	0.155	0.689	0.156	0.069	0.052	0.04	0.936565661	372231
ALM020	ES 2000 1200 BA	0.015	0.774	0.211	0.081	0.058	0.045	0.799693324	293441
ALM021	ES 2000 1200 BA	0	0.686	0.314	0	0.043	0.043	0.001661511	321798
ALM025	ES 2000 1200 BA	0.112	0.723	0.165	0.099	0.071	0.056	0.804749407	201311
ALM026	ES 2000 1200 BA	0.088	0.681	0.281	0.085	0.061	0.046	0.265662666	241563
ALM028	ES 2000 1200 BA	0.177	0.624	0.199	0.074	0.054	0.044	0.040599666	288151
ALM029	ES 2000 1200 BA	0.077	0.733	0.19	0.102	0.072	0.056	0.376924825	196314
ALM032	ES 2000 1200 BA	0.095	0.753	0.191	0.095	0.069	0.052	0.043160087	222415
ALM035	ES 2000 1200 BA	0.039	0.712	0.249	0.08	0.057	0.046	0.242332112	307838
ALM036	ES 2000 1200 BA	0.151	0.667	0.182	0.078	0.058	0.046	0.136919564	289648
ALM038	ES 2000 1200 BA	0.008	0.829	0.163	0.082	0.065	0.05	0.296927003	193680
ALM039	ES 2000 1200 BA	0.112	0.673	0.215	0.077	0.057	0.045	0.373598275	296021
ALM040	ES 2000 1200 BA	0.026	0.798	0.189	0.106	0.076	0.046	0.156185528	146863
ALM041	ES 2000 1200 BA	0.014	0.823	0.263	0.073	0.051	0.04	0.456843295	266110
ALM042	ES 2000 1200 BA	0.221	0.583	0.195	0.074	0.052	0.044	0.580484867	313426
ALM044	ES 2000 1200 BA	0.072	0.7	0.229	0.114	0.078	0.062	0.749048038	167502
ALM047	ES 2000 1200 BA	0.183	0.725	0.092	0.087	0.067	0.05	0.638809543	204168
ALM050	ES 2000 1200 BA	0.112	0.64	0.258	0.08	0.06	0.04	0.129769748	401473
ALM051	ES 2000 1200 BA	0.076	0.706	0.219	0.07	0.052	0.046	0.085054008	275800
ALM053	ES 2000 1200 BA	0.173	0.688	0.139	0.081	0.059	0.048	0.410674289	247940
ALM055	ES 2000 1200 BA	0.183	0.669	0.148	0.094	0.068	0.058	0.715633447	174136
ALM057	ES 2000 1200 BA	0	0.728	0.272	0	0.054	0.054	0.934589506	182206
ALM063	ES 2000 1200 BA	0.111	0.7	0.189	0.081	0.057	0.048	0.263737037	357186
ALM064	ES 2000 1200 BA	0	0.755	0	0	0.043	0.043	0.040793662	325929
ALM067	ES 2000 1200 BA	0.162	0.695	0.143	0.084	0.061	0.046	0.152222913	286536
ALM068	ES 2000 1200 BA	0.01	0.76	0.23	0.073	0.054	0.044	0.867864274	319603
ALM069	ES 2000 1200 BA	0.091	0.691	0.218	0.083	0.061	0.05	0.655168569	248106
ALM071	ES 2000 1200 BA	0.102	0.707	0.197	0.089	0.062	0.05	0.80976868	170233
ALM075	ES 2000 1200 BA	0.139	0.668	0.193	0.102	0.074	0.059	0.017470467	185076
ALM076	ES 2000 1200 BA	0.146	0.654	0.2	0.072	0.054	0.043	0.643781527	318275
ALM078	ES 2000 1200 BA	0.163	0.649	0.189	0.1	0.07	0.057	0.336629542	184135
ALM081	ES 2000 1200 BA	0.178	0.604	0.218	0.074	0.054	0.045	0.082097504	304726
ALM084	ES 2000 1200 BA	0	0.739	0.261	0	0.044	0.044	0.773671565	318288
ALM086	ES 2000 1200 BA	0.122	0.706	0.206	0.075	0.058	0.042	0.327601691	299925
ALM087	ES 2000 1200 BA	0.039	0.663	0.298	0.095	0.069	0.058	0.509125582	154412
ALM088	ES 2000 1200 BA	0.089	0.753	0.158	0.084	0.06	0.05	0.467915586	249550
BAS002	ES 2000 1200 BA	0.144	0.641	0.215	0.077	0.054	0.045	0.175163552	290339
BAS022	ES 2000 1200 BA	0.086	0.758	0.156	0.075	0.055	0.043	0.970452347	326556
BAS023	ES 2000 1200 BA	0.183	0.675	0.141	0.079	0.056	0.048	0.6762968	285653
BAS024	ES 2000 1200 BA	0.122	0.648	0.23	0.076	0.054	0.044	0.678788979	324837
BAS025	ES 2000 1200 BA	0.076	0.581	0.144	0.083	0.068	0.048	0.952522497	109254
BAS027	ES 2000 1200 BA	0.069	0.712	0.219	0.07	0.05	0.051	0.063123418	285009
CBR004	ES 2000 1200 BA	0.169	0.655	0.176	0.066	0.047	0.041	0.696157746	487732
CMO001	ES 2000 1200 BA	0.076	0.695	0.229	0.097	0.068	0.057	0.433505953	184370
EF4011	ES 2000 1200 BA	0.231	0.66	0.109	0.102	0.072	0.057	0.850685089	182593
LHO001	ES 2000 1200 BA	0.062	0.732	0.215	0.077	0.054	0.044	0.0604506	332022
LHO003	ES 2000 1200 BA	0.131	0.66	0.209	0.078	0.056	0.045	0.53796099	317977
PLZ001	ES 2000 1200 BA	0.252	0.551	0.197	0.068	0.051	0.042	0.977479521	342592
PUC001	ES 2000 1200 BA	0.199	0.59	0.212	0.07	0.05	0.043	0.699219207	337955
PUC002	ES 2000 1200 BA	0.086	0.724	0.19	0.069	0.05	0.044	0.70391687	342142
PUC003	ES 2000 1200 BA	0.136	0.653	0.212	0.077	0.054	0.046	0.007881149	316061
PUC004	ES 2000 1200 BA	0.199	0.623	0.178	0.077	0.057	0.043		

11390	FR_2500_2000_BBC	0.573	0.253	0.174	0.09	0.067	0.051	0.10384708	216670
11391	FR_2500_2000_BBC	0.431	0.119	0.15	0.187	0.117	0.11	0.152107478	81920
CBV95	FR_2500_2000_BBC	0.679	0.217	0.104	0.136	0.103	0.079	0.215743362	120687
GBVPK	FR_2500_2000_BBC	0.392	0.469	0.138	0.073	0.056	0.045	0.303760751	310698
PE12	FR_2500_2000_BBC	0	0.764	0.236	0	0.047	0.047	0.241637233	144052
13874	FR_2500_2000_BBC	0.381	0.469	0.151	0.075	0.045	0.045	0.48488123	302766
13875	FR_2500_2000_BBC	0.517	0.343	0.14	0.076	0.056	0.044	0.938616638	351547
116247	FR_E_2800_2200_LN	0	0.755	0.245	0	0.035	0.035	0.265386326	222245
5X32	FR_E_2800_2200_LN	0.595	0.27	0.134	0.097	0.068	0.058	0.043695565	192556
11392	FR_E_2800_2500_BBC/FR_E_2800_2200_LN	0	0.772	0.228	0	0.032	0.032	0.424153787	293029
113710	GB_2200_1200_BA	0.474	0.396	0.13	0.11	0.079	0.069	0.601528104	136699
113714	GB_2200_1200_BA	0.486	0.388	0.125	0.076	0.053	0.046	0.359539003	364130
113716	GB_2200_1200_BA	0.286	0.554	0.16	0.051	0.053	0.042	0.118961507	385778
116403	GB_2200_1200_BA	0.509	0.297	0.193	0.091	0.068	0.05	0.144436942	213628
119857	GB_2200_1200_BA	0.385	0.514	0.1	0.073	0.053	0.042	0.821642198	313550
119858	GB_2200_1200_BA	0.428	0.371	0.157	0.079	0.06	0.043	0.201892744	326512
119859	GB_2200_1200_BA	0.455	0.417	0.128	0.077	0.054	0.047	0.066734515	318163
119860	GB_2200_1200_BA	0.474	0.339	0.187	0.082	0.059	0.051	0.513090985	258884
119913	GB_2200_1200_BA	0.501	0.38	0.118	0.079	0.055	0.047	0.119301985	312419
119915	GB_2200_1200_BA	0.565	0.26	0.175	0.073	0.053	0.043	0.881866461	354566
12448	GB_2200_1200_BA	0.538	0.26	0.203	0.086	0.062	0.051	0.851527678	240612
112776	GB_2200_1200_BA	0.486	0.377	0.136	0.075	0.055	0.042	0.731000723	360328
112786	GB_2200_1200_BA	0.439	0.37	0.19	0.075	0.056	0.046	0.899278492	305115
114543	GB_2200_1200_BA	0.508	0.303	0.189	0.075	0.055	0.045	0.061393561	343384
114553	GB_2200_1200_BA	0.675	0.17	0.175	0.1	0.078	0.06	0.14811262	348484
116394	GB_2200_1200_BA	0.12	0.542	0.338	0.243	0.2	0.175	0.655839986	62926
116395	GB_2200_1200_BA	0.427	0.35	0.222	0.187	0.142	0.088	0.012296369	97654
116396	GB_2200_1200_BA	0.536	0.162	0.301	0.132	0.097	0.075	0.178995839	118685
116400	GB_2200_1200_BA	0.347	0.443	0.21	0.095	0.069	0.052	0.09036037	223850
116424	GB_2200_1200_BA	0.462	0.359	0.179	0.047	0.062	0.04	0.644407186	357681
118606	GB_2200_1200_BA	0.449	0.338	0.213	0.073	0.056	0.043	0.419480638	314310
119220	GB_2200_1200_BA	0.472	0.34	0.188	0.131	0.101	0.075	0.280196273	117459
120997	GB_2200_1200_BA	0.488	0.352	0.16	0.083	0.057	0.048	0.12291553	295222
121386	GB_2200_1200_BA	0.402	0.378	0.22	0.09	0.064	0.051	0.41148836	240453
12413	GB_2200_1200_BA	0.471	0.320	0.208	0.172	0.141	0.124	0.09849115	87343
12598	GB_2200_1200_BA	0.592	0.386	0.022	0.136	0.078	0.078	0.280479494	114698
114200	GB_2500_2000_BBC	0.433	0.432	0.135	0.121	0.086	0.069	0.562662966	114034
11767	GB_2500_2000_BBC	0.475	0.398	0.127	0.076	0.053	0.047	0.630759632	327629
12417	GB_2500_2000_BBC	0.708	0.117	0.175	0.07	0.057	0.045	0.600727356	361550
12418	GB_2500_2000_BBC	0.511	0.302	0.187	0.073	0.054	0.043	0.33823929	305656
12446	GB_2500_2000_BBC	0.403	0.423	0.175	0.085	0.061	0.051	0.885941246	236300
12447	GB_2500_2000_BBC	0.752	0.164	0.084	0.076	0.054	0.044	0.756604635	396128
12450	GB_2500_2000_BBC	0.757	0.153	0.09	0.075	0.054	0.043	0.53044603	401173
12452	GB_2500_2000_BBC	0.429	0.329	0.172	0.069	0.047	0.045	0.913725763	397681
12453	GB_2500_2000_BBC	0.505	0.328	0.166	0.084	0.06	0.049	0.425531608	281620
12454	GB_2500_2000_BBC	0.542	0.324	0.134	0.073	0.055	0.042	0.066732939	368471
12556	GB_2500_2000_BBC	0.603	0.262	0.135	0.072	0.053	0.043	0.83819007	369941
14950	GB_2500_2000_BBC	0.462	0.345	0.143	0.08	0.047	0.045	1.608	344377
15379	GB_2500_2000_BBC	0.429	0.447	0.124	0.077	0.057	0.045	0.031881949	315545
15512	GB_2500_2000_BBC	0.651	0.224	0.125	0.075	0.053	0.045	0.091628532	385504
15513	GB_2500_2000_BBC	0.606	0.223	0.171	0.078	0.054	0.045	0.037406965	359202
16775	GB_2500_2000_BBC	0.518	0.34	0.142	0.077	0.056	0.044	0.763559661	303559
16776	GB_2500_2000_BBC	0.708	0.117	0.094	0.07	0.057	0.042	0.10588155	31786
16777	GB_2500_2000_BBC	0.531	0.344	0.125	0.076	0.056	0.045	0.546002303	346505
16778	GB_2500_2000_BBC	0.455	0.384	0.161	0.081	0.057	0.046	0.706851545	274512
12443.5G	GB_2500_2000_BBC	0.646	0.255	0.099	0.068	0.049	0.042	0.6204904	470297
12445.5G	GB_2500_2000_BBC	0.55	0.312	0.169	0.088	0.047	0.043	0.44188868	489001
13255.5G	GB_2500_2000_BBC	0.758	0.184	0.057	0.178	0.041	0.041	0.382299364	371608
124342	HR_1800_1600_MBA	0.48	0.465	0.055	0.067	0.05	0.042	0.955129164	375686
126726	HR_1800_1600_MBA	0.378	0.59	0.032	0.071	0.049	0.042	0.448286963	399646
126773	HR_1800_1600_MBA	0.426	0.569	0.005	0.076	0.055	0.042	0.584086773	334408
125774	HR_1800_1600_MBA	0.25	0.653	0.097	0.125	0.053	0.041	0.692252407	371787
126893	HR_1800_1600_MBA	0.46	0.482	0.058	0.07	0.05	0.041	0.846436619	385631
JAG06	HR_1800_1600_MBA	0.449	0.36	0.191	0.078	0.055	0.048	0.262906071	309140
JAG58	HR_1800_1600_MBA	0.453	0.293	0.254	0.078	0.055	0.049	0.790430045	275007
JAG78	HR_1800_1600_MBA	0.421	0.375	0.204	0.078	0.057	0.045	0.024164581	339355
JAG82	HR_1800_1600_MBA	0.283	0.438	0.278	0.178	0.128	0.112	0.52761119	388712
JAG85	HR_1800_1600_MBA	0.433	0.438	0.129	0.079	0.059	0.048	0.852097439	266391
JAG93	HR_1800_1600_MBA	0.35	0.506	0.145	0.082	0.052	0.052	0.96411802	291182
12364	HU_2500_2000_BBC	0.292	0.658	0.05	0.074	0.053	0.044	0.348260871	341142
12365	HU_2500_2000_BBC	0.512	0.365	0.068	0.06	0.053	0.041	0.74886571	308027
12741	HU_2500_2000_BBC	0.036	0.791	0.173	0.067	0.05	0.041	0.729416045	381735
12786	HU_2500_2000_BBC	0.643	0.281	0.075	0.076	0.055	0.045	0.413212072	352162
12787	HU_2500_2000_BBC	0.794	0.146	0.06	0.082	0.059	0.048	0.445646695	303752
13529	HU_2500_2000_BBC	0.463	0.402	0.135	0.084	0.062	0.049	0.378162775	265682
14176	HU_2500_2000_BBC	0.529	0.291	0.18	0.082	0.061	0.046	0.423027411	270865
17044	HU_2500_2000_BBC	0.344	0.547	0.109	0.071	0.054	0.041	0.318698056	347709
17045	HU_2500_2000_BBC	0.091	0.74	0.169	0.073	0.057	0.043	0.883389866	343498
BR0003	IT_2200_1500_BA	0.201	0.662	0.136	0.074	0.053	0.043	0.277573423	321328
BR0007	IT_2200_1500_BA	0.337	0.573	0.09	0.17	0.119	0.089	0.340499888	104338
BR0010	IT_2200_1500_BA	0.144	0.668	0.188	0.112	0.086	0.06	0.848311244	185117
111442	IT_2200_1500_BA	0.89	0.11	0	0	0.043	0.043	0.128139709	334984
13122	IT_2200_1500_BA	0.134	0.865	0.002	0.073	0.052	0.044	0.306963669	295377
13123	IT_2200_1500_BA	0.096	0.776	0.128	0.071	0.052	0.041	0.238225977	391697
13124	IT_2200_1500_BA	0.208	0.653	0.031	0.07	0.052	0.041	0.995099095	390827
13125	IT_2200_1500_BA	0.122	0.878	0	0	0.052	0	0.292636276	394122
13878	IT_2200_1500_BA	0.081	0.888	0.03	0.069	0.052	0.04	0.394926896	372187
18561	IT_2200_1500_BA	0.239	0.606	0.156	0.072	0.052	0.042	0.122671947	394289
13123	IT_2200_1500_BA	0.132	0.748	0.124	0.064	0.047	0.038	0.52139174	488360
11979	IT_2500_2000_BBC	0.211	0.636	0.043	0.102	0.063	0.049	0.809655156	191311
12477	IT_2500_2000_BBC	0	0.833	0.167	0	0	0.04	0.877670265	352656
12478	IT_2500_2000_BBC	0.279	0.592	0.128	0.078	0.056	0.044	0.3617737	286006
111629	NL_1600_1100_BA	0.636	0.295	0.068	0.122	0.084	0.069	0.030743746	154737
111971	NL_1600_1100_BA	0.552	0.317	0.131	0.076	0.056	0.044	0.309535359	331727
111972	NL_1600_1100_BA	0.467	0.218	0.175	0.075	0.051	0.041	0.135385226	377884
111973	NL_1600_1100_BA	0.477	0.373	0.149	0.07	0.05	0.042	0.944153076	377285
112081	NL_1600_1100_BA	0.491	0.359	0.151	0.075	0.053	0.042	0.281917771	354044
112082	NL_1600_1100_BA	0.61	0.29	0.1	0.075	0.054	0.042	0.187240326	346375
112083	NL_1600_1100_BA	0.502	0.385	0.113	0.075	0.053	0.045	0.149594015	350230
117019	NL_1600_1100_BA	0.417	0.317	0.266	0.085	0.062	0.051	0.140875449	230729
120062	NL_1600_1100_BA	0.594	0.257	0.149	0.073	0.053	0.043	0.181097674	364787
126828	NL_1600_1100_BA	0.581	0.28	0.139	0.073	0.053	0.044	0.025099323	388687
128829	NL_1600_1100_BA	0.416	0.461	0.122	0.071	0.051	0.042	0.642428822	372181
128830	NL_1600_1100_BA	0.504	0.317	0.169	0.074	0.056	0.043		

Table S7. fastGLOBETROTTER inference of Yamnaya and Late Neolithic surrogate groups for ancient individuals

Target Group	Generation (\pm 95% CI)	pd fit of a sin	Date calBCE	Admixture date BCE
FR_N_2500_2000_BBC_SMGB54	12,42 \pm 5,51	0,48	2339 - 2139	2586,82 \pm 154,30
FR_N_2500_2000_Bre445FK	18,18 \pm 3,06	0,48	2509 - 2371	2948,95 \pm 85,68
FR_S_2500_2000_BBC	19,03 \pm 7,33	0,60	2435 - 2351	2925,94 \pm 205,22
NL_2500_2000_BBC	22,03 \pm 9,73	0,62	2523 - 2345	3050,87 \pm 272,35
NL_2500_2000_BBC	35,24 \pm 5,41	0,50	2119 - 1967	3029,84 \pm 151,38
NL_1800_1600_BA	35,12 \pm 12,22	0,64	1837 - 1699	2751,25 \pm 342,19
GB_2500_2000_BBC	19,63 \pm 7,62	0,45	2091 - 1967	2578,53 \pm 213,28
FR_S_1700_1200_MBA	38,18 \pm 16,84	0,26	1687 - 1611	2718,16 \pm 471,54
FR_N_2200_1700_EBA	14,91 \pm 6,29	0,48	2133 - 2029	2498,43 \pm 176,15
FR_E_2200_1700_EBA	31,06 \pm 4,50	0,57	1878 - 1772	2694,72 \pm 126,04
FR_S_2200_1700_EBA	28,06 \pm 6,17	0,42	1944 - 1880	2697,60 \pm 172,78
IE_Rathlin_2000_1500_EBA	24,35 \pm 4,93	0,57	1996 - 1900	2629,92 \pm 138,08
CZ_2500_2000_BBC	21,85 \pm 11,39	0,54	2433 - 2335	2995,80 \pm 318,89
CZ_2500_2000_BBC (n=5)	27,69 \pm 2,45	0,83	2326 - 2218	3047,40 \pm 68,51
PT_1800_1400_MBA*	34,58 \pm 4,30	0,46	1800 - 1400	2568,22 \pm 120,31
ES_2000_1200_BA	24,61 \pm 4,22	0,56	1914 - 1816	2554,07 \pm 118,06
FR_N_2500_2000_Bre445YY	16,73 \pm 3,50	0,73	2609 - 2445	2995,55 \pm 97,96

*Archaeological context dates were used as radiocarbon dates

Table S10. Comparison of the estimation of the number of generations since the admixture between Neolithic- and steppe-ancestry individuals using either fastGlobeTrotter or DATES and varying source populations

fastGLOBETROTTER	WesternLN x Yamnaya	EasternLN x Yamnaya	EuropeLN x Yamnaya
SMGB54	11,92 \pm 7,32	15,50 \pm 4,98	12,42 \pm 5,51
BRE445FK	15,45 \pm 5,23	18,34 \pm 4,84	18,18 \pm 3,06
GBVPK	17,95 \pm 3,76	18,56 \pm 4,55	19,03 \pm 7,33

DATES v4010	WesternLN x Yamnaya	EasternLN x Yamnaya	EuropeLN x Yamnaya
SMGB54	7,15 \pm 18,99	56,80 \pm 77,33	-21,40 \pm 98,32
BRE445FK	12,16 \pm 7,63	7,10 \pm 4,57	8,40 \pm 6,01
GBVPK	10,14 \pm 3,14	16,81 \pm 8,17	11,52 \pm 4,24

To allow comparison of the two methods, the shotgun whole genomes of the same individuals v

Supplementary References

1. P. Reimer *et al.*, The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). *Radiocarbon* **62**, 725-757 (2020).
2. C. Domenech-Jaulneau, A. Corona, in *InterNéo 12, Journée d'information du 6 novembre 2018*. (Association pour les Etudes Interrégionales sur le Néolithique et Société Préhistorique Française, Paris, Saint-Germain-en-Laye, 2018), pp. 159-169.
3. S. Brunel *et al.*, Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 12791-12798 (2020).
4. L. Hachem *et al.*, in *Les sépultures individuelles campaniformes en France*, L. Salanova, Y. Tchéremissinoff, 978-2-271-07124-8. , Eds. (Gallia Préhistoire, Paris, 2011), vol. XLIIe Supplément, pp. 21-35.
5. B. L. Beard, C. M. Johnson, Strontium isotope composition of skeletal material can determine birth place and geographic mobility of humans and animals. *J. Forensic Sci.* **45**, 1049-1061 (2000).
6. T. D. Price, J. H. H. Burton, R. A. Bentley, The characterization of biologically available strontium isotope ratios for the study of prehistoric migration. *Archaeometry* **44**, 117-135 (2002).
7. J. E. Ericson, Strontium isotope characterization in the study of prehistoric human ecology. *Journal of Human Evolution* **14**, 503-514 (1985).
8. R. A. Bentley, Strontium isotopes from the earth to the archaeological skeleton: a review. *J. Archaeol. Meth. Theor.* **13**, 135-187 (2006).
9. M. Willmes *et al.*, Mapping of bioavailable strontium isotope ratios in France for archaeological provenance studies. *Applied Geochemistry* **90**, 75-86 (2018).
10. L. J. Raggatt, N. C. Partridge, Cellular and molecular mechanisms of bone remodeling. *The Journal of biological chemistry* **285**, 25103-25108 (2010).
11. T. D. Price, C. Knipper, G. Grupe, V. Smrčka, Strontium isotopes and prehistoric human migration: the bell beaker period in Central Europe. *Eur. J. Archaeol.* **7**, 9-40 (2004).
12. M. M. Ash, S. J. Nelson, *Wheeler's dental anatomy, physiology, and occlusion*. (Saunders, W.B. , Philadelphia, 2003).

General Conclusion

The recovery of DNA and the generation of genomic data from ancient specimens significantly transformed our understanding of the genetic history of human populations. As a result of methodological developments of both molecular biology and bioinformatics, the contribution of paleogenomics to the fundamental questions of evolutionary biology and human history is drastically increasing. Although the first large-scale studies covering vast geographical regions and time periods provided the first glimpse of world-wide peopling processes, especially in western Eurasia, regional small-scale studies at higher resolution are required to reconstruct the human past. These studies are hampered by the low coverage of a large part of the genomic data. This is of particular concern for most genomes from older periods, such as the Upper Paleolithic, due to the scarcity and low accessibility of well-preserved materials, although exceptions exist (e.g., the genomes obtained from the remains preserved in the Denisova cave in southern Siberia). In the Holocene, however, coinciding with the start of the Neolithic period, the abundance of human remains increases due to the increase in population size, the gathering in settlements as a consequence of the sedentary lifestyle. The petrosal part of the temporal bone has been shown to be the skeletal part from which the highest endogenous DNA yields are obtained. This discovery boosted the paleogenomic field as it led to higher data quality and decreased costs. For the projects of the present thesis project, we chose different study designs with the use of state-of-the-art methods tailored to the abundance of the remains analyzed, the degree of DNA preservation, the quantity and quality of published data for the comparative analyses and the scientific questions addressed.

Methodological developments of the DNA extraction, purification and library construction protocols enabled us to generate low coverage genome sequences from the cranial fragments of two Early Upper Paleolithic Buran Kaya III individuals from Crimea, in which DNA is very poorly preserved (Chapter 2, Fig S10-12). To maximize the amount of comparable

data between these individuals and other Paleolithic genomes, we exploited the SNP panel coming from Human Genome Diversity Project (HGDP). We applied population genetics tools that can deal with missing data. The combined analysis of these genomes with other recent ancient genomic findings from Europe allowed us to create a broad and updated model of population movements, interactions and turnovers during the first peopling waves of Europe by *Homo sapiens* in the Early Upper Paleolithic. Both individuals from Crimea were found to be more similar to subsequent European populations than to earlier inhabitants of Europe associated with the Initial Upper Paleolithic (IUP). However, since the Buran Kaya III individuals carried European IUP-associated ancestry, we can conclude that some of this ancestry persisted following the climatic crisis during the severe Heinrich 4 stadial and the CI super-eruption that devastated southern and eastern Europe at the beginning of the Heinrich 4 event. Thus, various levels of admixture events must have occurred between local and early migrating populations where migrating populations may be more numerous than the local ones, a phenomenon paralleled by the later Holocene expansions.

The two studies of the Holocene population dynamics in the Paris Basin presented in this thesis covers a period from the Late Mesolithic to the beginning of the Bronze Age, between 8,200 and 4,200 years ago, The ancient genomes produced have sufficient coverage to allowed us to impute and phase them together with other contemporary genomes using the reference panel of the 1000 Genomes Project with recently developed methods. Our sampling strategy made the preparation of a dataset possible suited for a deep characterization of Neolithic farmer populations in western Europe. By generating genomic data from the earliest farmer population that arrived in the Paris Basin, the last stage of the early Neolithic expansion, we were able to discriminate the genomes of the people moving along the continental and Mediterranean migration routes although they diverged from a common ancestor in southeastern Europe just around one thousand years ago. The early Neolithization of Europe is paralleled by a continuous fine-scale differentiation process of the genome of the early farmers. This differentiation became more apparent in later populations when the early expansion of farming came to an end and the population size of the different groups increased. The newly established local ancestries were preserved even during the Bronze Age as we were able to distinguish distinct Neolithic ancestries after the arrival of Yamnaya-related steppe-ancestry in western Europe. The detection of such prehistoric fine-scale structures in Europe allowed us to locate the possible origins of individuals or groups whose ancestors were moving and to reveal

interactions that could not be detected by the archaeological data alone. For instance, we detected the ICC-related ancestry in an individual buried in a particular burial in an LBK site in the Paris Basin, possibly after being sacrificed, which might indicate the notion of different identities during the Early Neolithic. We have shown complex gene flow processes by analyzing spatiotemporal patterns of admixture events between encountering populations from transition periods of the Mesolithic to the Neolithic and from the Late Neolithic to the Bronze Age. These analyses revealed correlations between cultural and genomic transformations as we detected different admixture dates pointing to known prehistorical events such as the expansion of farming via the ICC and LBK cultures, the end of Early Neolithic, the emergence of the Corded Ware and Bell Beaker cultural complexes. On the other hand, by analyzing the genomes of individuals with highly elevated hunter-gatherer ancestry we detected that the ancestors of these individuals were hunter-gatherers, who received recent gene flow from Neolithic groups yet preserved their hunter-gatherer ancestry. This means that they did not receive further gene flow from the Neolithic individuals but rather mated inside of their group. The individuals and groups with high hunter-gatherer ancestry were not identifiable archaeologically but seem to have been integrated into the Neolithic villages. The admixture dating results indicate that the admixture between hunter-gatherers and Neolithic farmers took place in the first half of the 4th millennium BCE lending particular importance to this period for future studies. Strikingly, we also observed similar cases of preferential mating between individuals of related ancestry when studying the spread of the steppe-ancestry in Europe at the onset of the Bronze age. The pattern of preservation of a high level of steppe-ancestry during the westward migration into Europe and the observation that most of the admixture with local Neolithic farmers could be dated to an initial major pulse of admixture when arriving in eastern Europe is also indicative of preferential intra-community mating of people with related ancestry. Indeed, we detected in real-time the emergence of a second admixture pulse when analyzing a Neolithic woman who had a son with an individual with steppe ancestry. We dated this second pulse to a period that saw the emergence of a new culture, the advanced Bell Beaker Complex, a situation that mirrors the previous emergence of the Corded Ware Culture in Central Europe. The analyzed individuals dated after 2500 BCE were descendants of this second major admixture pulse. They have kept for about one thousand years a relatively high level of steppe-ancestry indicating that they must have preferred intra-community mating over exogamy. Thus, even though mating between people of different ancestries occurred in pulses, often associated with the emergence of new cultural practices, as seen for the LBK, the CWC and the BBC, the pattern of preferential intra-community mating appears to be widespread even when two communities

share the same territory. This mating preference is probably mainly connected to culture, and most of all to language, and a major trait of human behavior in most cultures up to the present. Another case of segregation associated with ancestry could also be detected in the two Chouilly hypogea, as the dead appear to have been segregated in time (although the timing overlaps) but also based on their origin, with one hypogeum entombing individuals with Northern France ancestry, and one hypogeum entombing people of Southern France ancestry. Indeed, the belonging to a community, either via biological or cultural affiliation, or both, is of such importance to humans that they preserve the community links beyond death.

The imputation of ancient genomes allowed us to exploit different haplotype-based methods to answer various questions by providing further resolution to our analyses compared to more commonly used independent SNP-based methods. Therefore, to better characterize the populations and understand their genetic variation it was our strategic choice to analyze more but less well-covered genomes rather than a few high coverage genomes. It must be noted, however, that this was possible only since we analyzed the ancestors of present-day populations who are well-represented in the reference panels. Indeed, it would be highly challenging to characterize past populations when their various ancestries are not represented in the present-day populations, as in the case of early AMHs, or when the reference panels are poorly represented, such as for the Sub-Saharan African populations. Although the imputation and phasing of ancient genomes using present-day reference panels can provide sufficient power for a high-resolution analysis of past populations, it is still challenging to infer the exact haplotypes in ancient genomes. Since in the present-day populations the ancestral haplotypes are broken by recombination through many generations, the best solution would be to use ancient panels specific to certain periods and regions to access the accurate ancestral haplotypes. However, this would require the generation of many high coverage genomes from ancient individuals and family trios, which is technologically highly challenging and costly despite the fact that the number of high coverage genomes has been increasing recently.