

pool of the region during the development and spread of metallurgy across Europe, at the beginning of the Bronze and Iron Ages, are still unknown.

To investigate the demographic dynamics over the 7,000-y transect from the Mesolithic Period, before the onset of agriculture, to the Iron Age, we genetically analyzed 243 unique

individuals sampled from 54 different archeological sites, enriching for both complete mitochondrial genomes and a panel of 120 nuclear single-nucleotide polymorphisms (SNPs) including Y-chromosome SNPs, as well as a partially overlapping dataset of 58 low-coverage genomes. Our data from ancient western Europe fill a geographic gap that was so far missing,

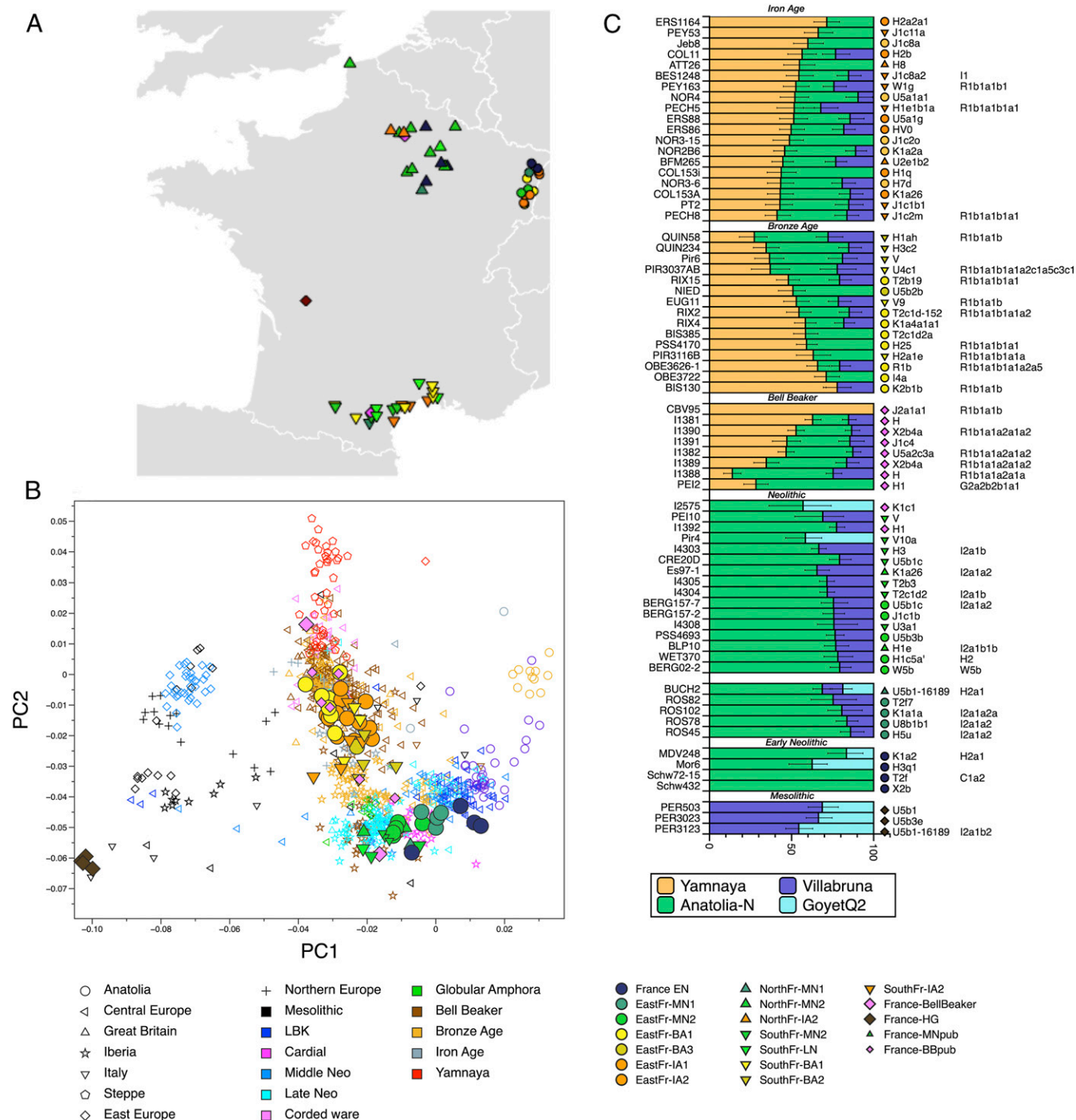


Fig. 1. Overview of the ancient French dataset. Filled symbols are used for ancient individuals from France whereas open symbols are used for other ancient western Eurasians. The shape of the symbol indicates the geographic origin and the color indicates the time period. (A) Location of the samples included in the study. (B) Principal-component analysis of ancient western Eurasians projected onto the variation of present-day genotypes, restricted to Europe. (C) Ancestry proportion for French individuals ranging from the Mesolithic to the Iron Age established using qpAdm (data from this study and from previously published studies). Each bar represents one individual with the associated mitochondrial DNA haplogroup and Y-chromosome haplogroup (Right). Error bars indicate ± 1 SE. Symbols used in A and B are indicated for each individual.

allowing a more global view of past population dynamics in Europe.

Results and Discussion

All analyzed individuals belong to well-defined archaeological contexts from the Mesolithic Period to the Iron Age and 40 were directly radiocarbon-dated for this study (Dataset S2). We built sequencing libraries for 243 individuals and enriched them for mitochondrial genomes and a panel of 120 selected SNPs across the nuclear genome (Dataset S1). Along with complete mitochondrial genomes of 223 individuals and Y-chromosome haplogroups of 62 male individuals, we report 58 ~0.05 to 0.5× genomes sequenced from uracil-DNA glycosylase (UDG)-treated libraries. We combined our genotype data with either complete genomes or 1,240k genome-wide polymorphisms in ancient individuals and genotypes of modern individuals from Europe, the Caucasus, and the Near East genotyped on the Affymetrix Human Origins array (3, 5, 9, 10, 12, 13, 21, 22). Complete mitochondrial genomes were retrieved from this ancient dataset and were grouped based on temporal and geographic proximity, often translating into cultural classifications (SI Appendix, Fig. S3-1). Contamination estimates derived from X-chromosome polymorphisms in male individuals were consistently low (between 0.68 ± 1.26 and $3.55 \pm 0.62\%$ depending on the method considered), with only one individual displaying more than 5% X-chromosome contamination (PIR3037AB) (Dataset S1).

The French Mesolithic Substrate. We generated genomic data for five Mesolithic individuals from Les Perrats cave (Agris, Charente) dated from 7177 to 7057 calibrated years BCE (calBCE) including low-coverage shotgun sequencing for three of them (0.111 to 0.134×). All possess haplogroup U5b, demonstrating in the territory of present-day France the presence of a Mesolithic substrate akin to that described elsewhere, where it was characterized by haplogroups U5b (85%) and, to a lesser extent, U5a (15%) (3, 23).

Mesolithic hunter-gatherers (HGs) discovered in France are located near the edge of a principal-component analysis (PCA) plot obtained from genome-wide genotypes, falling near the Mesolithic individuals from western Europe (5, 6, 9) (Fig. 1B and SI Appendix, Fig. S4-3). While the ancestry from an ~14,000-y-old individual from Villabruna in Italy was found to be dominant in Holocene HGs across western and central Europe (3), recently published studies conducted on ancient Iberians highlighted the survival of Magdalenian-associated ancestry (11). This dual ancestry is found in all Iberian HGs (ranging from 23.7 to 75.3%), while in most other regions only the Villabruna-associated ancestry remained. In order to investigate the proportions of ancestry in French HGs derived from these two Late Pleistocene lineages, we used qpAdm to model it as a mixture of two sources: ~15,000-y-old GoyetQ2 from Belgium, the least-admixed individual associated with the Magdalenian complex in Europe sequenced so far (11), and Villabruna. We found that French HGs from Les Perrats harbored relatively high proportions of GoyetQ2 ancestry ranging from 31.3 to 45.6% (Fig. 1C), comparable to proportions described in the La Braña or Canes1 Mesolithic individuals from Spain, suggesting a late survival of Magdalenian-associated ancestry in HGs outside the Iberian Peninsula.

Successive Waves of Migration and Admixture over the Course of the Neolithic in France. The arrival of an Anatolian Neolithic-associated ancestry component at the onset of the Neolithic is clearly visible in our dataset from both uniparentally inherited markers and at the genome-wide level. Maternal lineages from Neolithic farmers in France (from ~5300 BCE onward) are more diverse than those from HGs, and display variable frequencies across the time transect (SI Appendix, Fig. S3-2). Early and

Middle Neolithic individuals share more affinity with present-day southern Europeans (SI Appendix, Fig. S4-2), and their genetic variation is encompassed within that of contemporaneous European populations (Fig. 1B). Three of our individuals associated with the LBK culture, which stems from the Danubian Neolithic current, cluster genetically with other central European Early Neolithic (EN) individuals and share drift with other early farmers associated with the LBK cultural complex (SI Appendix, Fig. S4-6). Values of *D* (Mbuti, France_EN; western hunter-gatherers [WHGs], Anatolia_Neolithic) are consistent with those observed in LBK-associated individuals, with the exception of Mor6 (~7,100 calB.P.) in northeastern France (Fig. 2B). Despite a consistent dating and cultural assignment to the LBK, this individual falls within the genetic diversity of Iberian_EN individuals and harbors the highest proportion of shared alleles with WHGs. After testing the potential origin of this ancestry with qpAdm, we found the best fit to be with a two-source ancestry model between Anatolia_Neolithic and GoyetQ2 (*P* = 0.0911155), a situation so far only described in Early Neolithic individuals from the Iberian Peninsula. This is the northern-most description of such ancestry. To date, the distribution of the Magdalenian-associated ancestry in the HG population in western Europe that admixed with the Neolithic farmers in Europe is not clear because of the scarcity of genetic data from Mesolithic HGs. The observation of this ancestry in the Iberian Peninsula has been interpreted as a feature distinguishing the Cardial from the LBK Neolithic migrants (12).

The observation that the ancestry of GoyetQ2 was found alongside Villabruna ancestry in HGs from Les Perrats in western France (~9,100 calB.P.), as well as in HGs from Rigney1 (central-eastern France; ~15,500 calB.P.) and from two caves in southwestern Germany, Hohlefelds (~15,000 calB.P.) and Burkhardtshöhle (~14,600 calB.P.), rather indicates that this mixed ancestry is characteristic of western European HGs (11). Thus, the admixture between HGs with GoyetQ2 ancestry and Neolithic migrants is not necessarily characteristic of the Cardial Neolithic migration wave, as it seems to have occurred as well in more northern locations in western Europe.

While the only male Mesolithic individual from our dataset was assigned to Y-haplogroup I, LBK individuals from France showed relative diversity, belonging to Y-haplogroups C1a2, G2a, and H2. G2a and H2 have been described in various European Early Neolithic contexts in both central Europe and Iberia (3, 6, 24). I2 haplogroups are likely to have been introduced into the Neolithic pool through admixture with hunter-gatherers.

In eastern France, the transition into the Middle Neolithic is represented by individuals from the Grossgartach culture, that derives from the Danubian sphere (~4700 to 4500 BCE; SI Appendix, Fig. S1-1). Their proximity to central European contemporaneous individuals on PCA plots generated from both mitochondrial haplogroup frequencies and genome-wide markers suggests genetic homogeneity within the Danubian current (Fig. 1B and SI Appendix, Fig. S3-3). Interestingly, the BUCH2 individual harbors mixed GoyetQ2 and Villabruna HG ancestry (Fig. 1C), suggesting again that the GoyetQ2 ancestry is not a specific signature of southwestern Europe. However, we can observe a subtle shift away from LBK individuals and toward WHGs, reflecting an increase in shared alleles with the latter. This echoes observations from other ancient populations across Europe (Figs. 1C and 2A), yet with a higher proportion. This phenomenon seems to appear in France very early and becomes stronger as the Neolithic progresses, since individuals from French Middle Neolithic sites share a higher proportion of ancestry with HGs (Fig. 2A) and display Mesolithic signature haplogroups (e.g., U5b) at variable frequencies across the territory (SI Appendix, Fig. S3-2). Predominant in the Champagne region of northeastern France, these haplogroups suggest

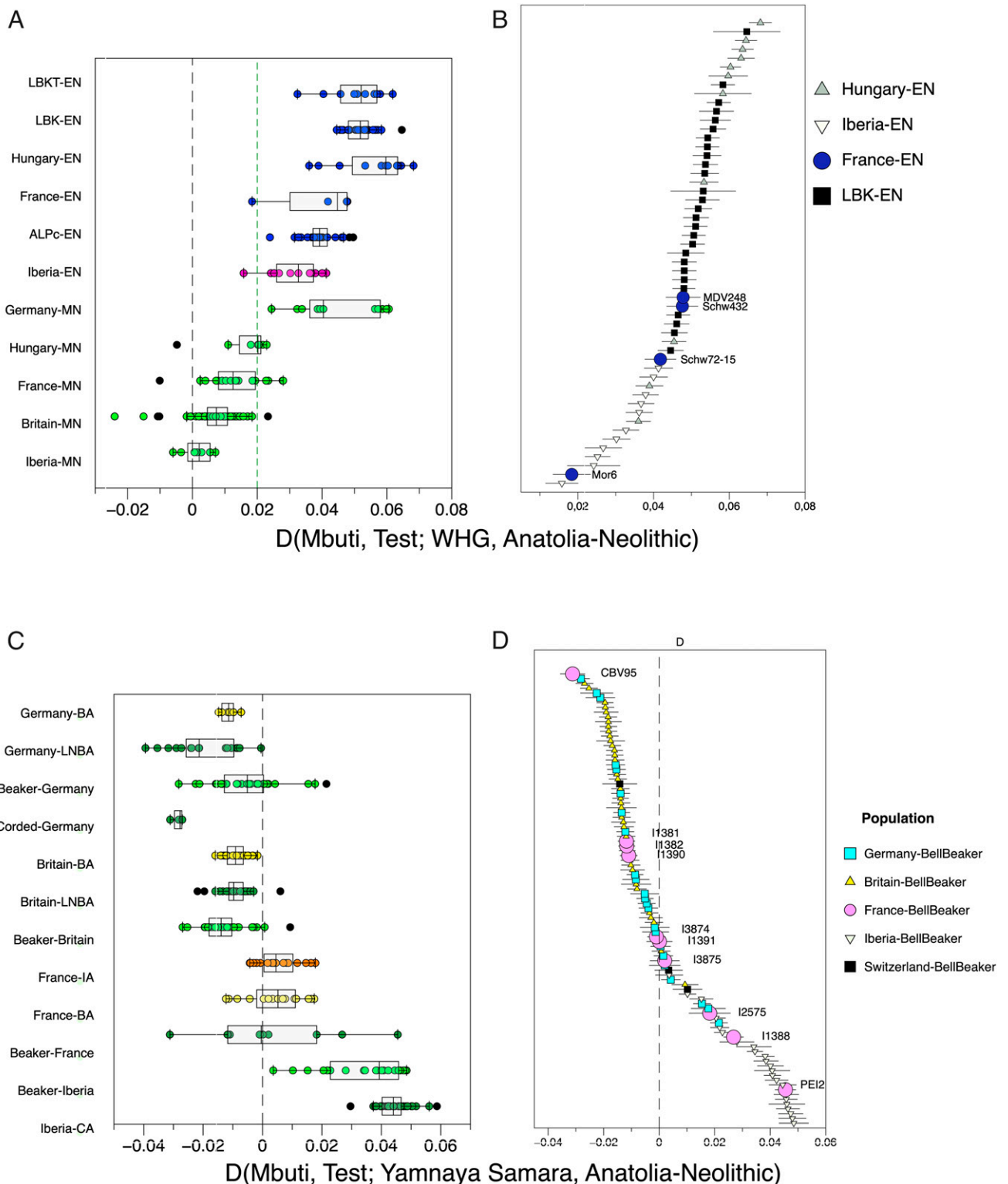


Fig. 2. Genetic affinity between ancient French and either western hunter-gatherers or steppe herders. (A) Boxplot of the D -statistic values of the form $D(\text{Mbuti, Test; WHG, Anatolia-Neolithic})$ for the Early and Middle Neolithic European population with WHGs consisting of Loschbour, ElMiron, GoyetQ2, Villabruna, KO1, and LaBrana to represent better the GoyetQ2 and Villabruna ancestry. (B) D statistics of the form $D(\text{Mbuti, Test; WHG, Anatolia-Neolithic})$ for Early Neolithic individuals of present-day Germany, France, and Iberia. Error bars represent ± 1 SE. (C) Boxplot of the D -statistic values of the form $D(\text{Mbuti, Test; Yamnaya_Samara, Anatolia-Neolithic})$ for Late Neolithic and Bronze and Iron Age populations. (D) D statistics of the form $D(\text{Mbuti, Test; Yamnaya-Samara, Anatolia-Neolithic})$ for individuals associated with the Bell Beaker cultural complex. Error bars represent ± 1 SE. ALPc, Eastern (Alföld) Linear Pottery Culture; BA, Bronze Age; IA, Iron Age; LNBA, Late Neolithic and Bronze Age; CA, Chalcolithic or Copper Age.

variable proportions of admixture between autochthonous hunter-gatherers and farmers. During the second half of the Middle Neolithic (MN2, ~4500 BCE), individuals from northern and southern France are poorly distinguished on the PCA plot (Fig. 1B).

Unlike in Iberia, the migration wave from which the Neolithic ancestry in ancient populations from France originated could not clearly be established using *D* statistics comparing the ancestry proportion between Iberia_EN and LBK_EN, as most individuals display nonsignificant values (*SI Appendix*, Fig. S4-6) (12). Using qpAdm, we investigated the ancestral source of the HG ancestry in individuals from French MN2 sites. Only individuals from southern France harbor varying levels of GoyetQ2-like ancestry, suggesting admixture with different populations of hunter-gatherers (Fig. 1C).

The Bell Beaker Period in France. At the end of the Neolithic, the Bell Beaker complex (named after the particular shape of its ceramics) appears and becomes widespread across Europe. The Beaker phenomenon is peculiar because of its large geographic distribution and because of its coexistence with local Late Neolithic and Copper Age cultures.

We report two Bell Beaker-associated individuals (CBV95 and PEI2), that we coanalyzed with previously reported contemporaneous individuals from Europe, including France (10). French Beaker-associated individuals display a wide range of steppe-ancestry proportions (Figs. 1C and 2D). CBV95 in northern France derives the highest proportion of alleles from the Yamnaya in our dataset, and belongs to Y-chromosome haplogroup R1b (Figs. 1C and 2D and *SI Appendix*, Fig. S4-5), providing the earliest clear evidence of the presence of this haplogroup in France around 2500 BCE (*Dataset S10*). This lineage was associated with the arrival of migrants from the steppe in central Europe during the Late Neolithic, and was described in other parts of Europe and in Bell Beaker-associated individuals from southern France, while being almost absent in Iberia prior to the Bronze Age (10, 13). PEI2, a male unearthed from a collective burial site near Carcassonne in southwestern France with artifacts of the Bell Beaker complex, falls within the genetic diversity of Neolithic individuals in the PCA. Modeling admixture

proportions between three source populations, Anatolia_Neolithic, Villabruna, and Yamnaya_Samara, we could, however, detect 28.3% of steppe ancestry in PEI2 (Fig. 1C). These observations are consistent with previous findings and confirm that steppe ancestry appeared later and with a lower impact in southwestern Europe than in other parts of the continent (10, 25). Surprisingly, the admixture model did not include Villabruna as an ancestry source for either CBV95 or PEI2, which differs from previously known Late Neolithic individuals.

Relative Continuity between the Bronze Age and the Iron Age.

Bronze Age individuals in France display new mitochondrial haplogroups such as U2, U4, and I, although at a low frequency. First described in eastern European populations and in Late Neolithic central Europe (13), these haplogroups suggest gene flow with a population deriving maternal ancestry from the Pontic steppe herders at the onset of the Bronze Age. As foretold by its early occurrence in Beaker-associated CBV95, a drastic Y-chromosome turnover occurs during the Bronze Age, where R1b replaces the preexisting diversity of Neolithic lineages in our sampling (Bronze Age, 11/13 and Iron Age, 7/10 individuals having R1b; *SI Appendix*, Fig. S5-1 and *Datasets S10 and S11*).

Bronze and Iron Age France share a common space in the PCA plot, both shifted toward modern central Europe and falling within the genetic diversity of Bronze Age Britain and central Europe (Fig. 1B), with a homogenization of the steppe component (Figs. 1C and 2C). In contrast to what was described for central Europe (9), there is no further shift toward eastern Eurasian genotypes during the Iron Age (9). Instead, the steppe component, heterogeneously distributed between individuals during the Bronze Age (ranging from 30 to 70%), becomes homogeneous (Figs. 1C and 2C), and individuals from the Hallstatt and La Tène culture in the French territory display similar affinities toward both modern and ancient populations. This could indicate that the transition from the Bronze Age to the Iron Age in France was mostly driven by cultural diffusion, without major gene flow from an external population. This would be consistent with an archeological and linguistic hypothesis proposing that the Celts from the second Iron Age descended from populations

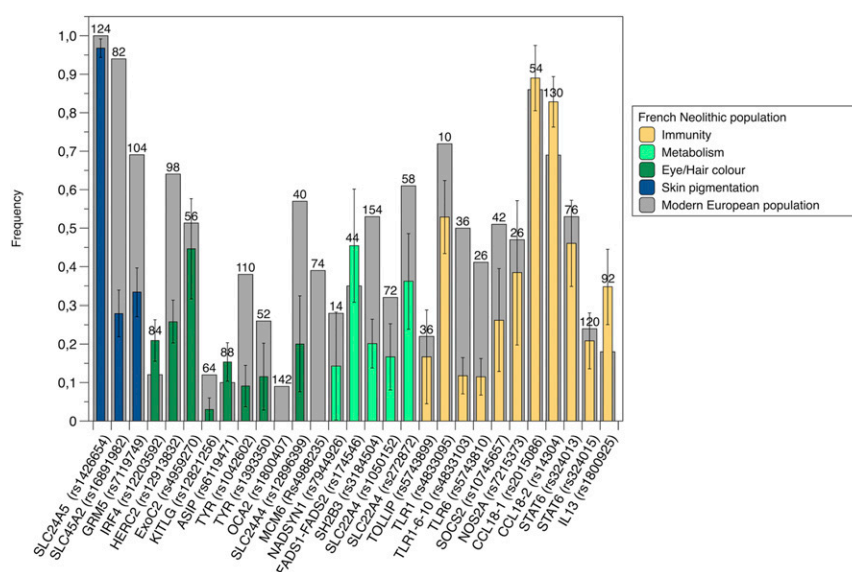


Fig. 3. Frequency of the derived allele for several types of genetic markers in Neolithic French and present-day Europeans. Numbers indicated above the bars correspond to the total number of observations of the derived allele for each variant in the Neolithic population. The color code reflects the function associated with the considered locus. The frequency of the alleles in present-day European populations is indicated in gray. Error bars indicate ± 2 SE.

already established in western Europe, within the boundaries of the Bell Beaker cultural complex (26–29). It is important to mention, however, that due to the relative genetic homogeneity among European populations by the Bronze Age, subsequent migrations between different parts of Europe could easily remain unnoticed at this level of coverage (10, 25, 30).

Analysis of Genetic Markers Associated with Phenotypes. To follow the evolutionary timelines of genetic adaptation to changing lifestyles and environments, we analyzed 73 autosomal loci, associated with both physical and physiological traits, genotyped across 149 individuals (Fig. 3; see [Dataset S3](#) for details on the targeted positions). Among the phenotypic traits of interest, we focused on genetic variants involved in eye and skin pigmentation (such as SLC24A5, SLC45A2, GRM5, HERC2, IRF4, and TYR). It has been previously shown that Mesolithic HGs and Anatolian Neolithic farmers differ in their pigmentation alleles, with the latter carrying derived alleles responsible for a light skin that are near fixation in Europe today (22), while HGs are described as having a rather dark skin and light eye color (6, 31). The Mesolithic individuals from whom we retrieved reliable genotype information after stringent filtering (one to three individuals) possessed the ancestral pigmentation variant for loci SLC45A2 and GRM5, which had been associated with darker skin. In contrast, the derived alleles of these two loci are found at significantly higher frequency in the French Neolithic population (SLC45A2: 24.9%; GRM5: 38.4%), although they remain far from the frequency reached by these alleles in present-day Europe, where they are near fixation in the population (93.8 and 68.9%, respectively). SLC24A5, described as the principal mutation associated with light skin pigmentation, is found in 96% of our Neolithic population, close to the 99% found in present-day Europe. These results suggest different evolutionary timelines for these mutations in France, where a similar phenotype (light skin pigmentation) possibly resulted from several selection events, although the genetic determinism of these traits is likely more complex and not encoded by these three loci alone (32). The Mesolithic individual who displayed coverage of variants involved in the pigmentation of eye color (HERC2 and IRF4) carried the alleles associated with blue eye color. At the time of the Neolithic, the frequency of these two variants reaches 29.6 and 24.4%, respectively. Interestingly, these two variants seemingly have different trajectories after the Neolithic. While HERC2 rises in frequency again from 37.5% during the Bronze Age to finally reach more than 63% in present-day Europe, IRF4 further decreases from 23.81% during the Bronze Age to reach a frequency of 11% at present. None of the genotyped Neolithic individuals carried the mutation responsible for the persistence of lactase in Europeans, consistent with a later origin for this mutation (13, 30). We also analyzed several innate immunity-related variants that show signs of recent positive selection in modern European populations (6, 33–35). For a number of genes involved in the immune response (NOS2A, TOLLIP, CCL18, STAT6, and IL3), the frequency of derived alleles was comparable to that found in present-day Europeans, indicating that selection on these loci predates the Neolithic, whereas frequencies of variants in the TLR1–6–10 locus, the Toll-like receptor family involved in the innate immune response, are much lower than those found in present-day Europeans (36). The TLR1–6–10 gene cluster is possibly associated with resistance to leprosy, tuberculosis, or other mycobacteria (36). Our data suggest that the onset of agriculture and the proximity to domesticated animals did not exert a selective pressure on this locus. Likewise, variants associated with celiac disease in SH2B3 or SLC22A4 that could have been the target for positive selection owing to the role they play in inflammation and the elimination of a wide array of environmental toxins (33) did not yet reach the frequency of modern populations.

Conclusion

The study of 7,000 y of demographic history through the study of ancient genomes throughout France reveals successive migration events mostly accompanied by major cultural changes and admixture events whose traces are still found in current European populations.

Our genomic data obtained for three Mesolithic individuals from western France show that the late survival of the Upper Paleolithic Magdalenian-associated ancestry (represented by GoyetQ2) was not restricted to the Iberian Peninsula. This finding expands the region where this ancestry is found and raises the question of where this admixture occurred. More Mesolithic individuals from western Europe need to be studied to better characterize the history of this ancestry.

While the first Neolithic people who inhabited the territory of present-day northern France descended from early farmers in Anatolia, therefore forming a group distinct from the local hunter-gatherers, we could detect later admixture with hunter-gatherers supporting their progressive integration within farming communities. Although widespread, the process of admixture with hunter-gatherers shows regional variability and origins. However, the scarcity of the data available for the south of France does not allow us to conclude whether contact occurred between the LBK and Cardial complexes, nor to identify any legacy that may be present in the Neolithic cultures that followed. Genomic data from other French regions in the north and the south will be necessary to confirm the hypotheses on the origin and the relations of the different cultural groups of the Middle and Late Neolithic.

The study of a selection of nuclear markers revealed differences between Neolithic people and present-day Europeans. Derived allele frequencies indicate selection on loci involved in pigmentation, diet, and immunity that are consistent with adaptation to high latitudes and changes in diet. Finally, neither Mesolithic HGs nor Neolithic farmers in the territory of present-day France carried the allele allowing for lactase persistence during adulthood, in agreement with studies of Neolithic populations farther east that reported the presence of this allele only by the end of the Bronze Age (13, 30).

An important outcome of our study is the genetic continuity between Bronze and Iron Age individuals of France but with less heterogeneity between individuals from the Iron Age. In such a context, tracing migrations from one part of Europe to the other becomes more challenging and will require a deeper temporal and geographical resolution.

Materials and Methods

Processing of Archaeological Samples. A total of 243 individuals from 54 archaeological sites across France passed screening for the preservation of ancient DNA and were processed for sequencing ([SI Appendix, Fig. S1-1](#) and [Dataset S1](#)). These individuals covered a time period ranging from the Mesolithic to the Iron Age, with the majority of samples covering the Neolithic Period (~5500 to 3500 BCE). As direct dates were not available for all samples, they were labeled and grouped for analyses based on geographical proximity and cultural assignment ([SI Appendix, Fig. S1-1](#)).

Targeted DNA Enrichment and Sequencing. DNA was extracted from petrous bones and teeth and used to build double-stranded Illumina libraries with thorough decontamination procedures and optimized protocols in a dedicated facility at Institut Jacques Monod ([SI Appendix, Text S2](#)) (37). Sequence enrichment for both the whole mitochondrial genome and nuclear regions was performed through a capture approach using biotinylated RNA baits obtained through in vitro transcription of PCR-amplified target regions (37) (see [SI Appendix, Text S2](#) for details on the procedure). Both enriched and whole-genome shotgun libraries were sequenced on an Illumina MiSeq or Illumina NextSeq at Institut Jacques Monod and Institut de Recherche Biomédicale des Armées, respectively.

Data Processing and Authentication. Adapters were trimmed and overlapping paired-end reads were merged using ILLUMINA (38). Merged reads were then quality-trimmed with cutadapt and only reads longer than 30 bp were retained. Mapping was performed with BWA version 1.2.3 (11) against the Cambridge reference sequence (39) for human mitochondrial DNA with a duplication of its first 100 bases at the end to ensure mapping of the reads overlapping the junction resulting from the virtual linearization of the circular mitogenome and to the human reference genome hs37d5. In both cases, PCR duplicates were removed using Picard's MarkDuplicate (<https://broadinstitute.github.io/picard/>). Ancient DNA authentication was assessed using mapDamage (40) and contamination was estimated using both the mitochondrial genome and X chromosome (41).

Uniparental Haplogroups. Consensus sequences of the mitochondrial genome were obtained using the software ANGSD version 0.910, relying only on reads with base quality above 20 and mapping quality above 30, and with more than 3× coverage. Mitochondrial haplotypes were determined based on PhyloTree phylogenetic tree (42) build 17 and by using the HAPLOFIND web application (43) and Phy-Mer (44). Unexpected or missing mutations were visually inspected in Geneious version R6 (45) to check if they could be the results of misincorporations in low-coverage regions. Sequences with more than 5% undetermined sites were removed. Y-chromosome haplogroups were retrieved using Yleaf software (46).

Nuclear Markers. Target regions were summarized in a file, allowing their extraction using the mpileup tool from the SAMTools toolkit (47). Proper variant-calling files were obtained using bcftools call and filtered for a base quality of 20 and a minimum coverage of five reads using the SelectVariants tool from the Genome Analysis toolkit (48). Allelic count and frequency estimation was carried out with PLINK (49).

Reference Datasets. We merged our dataset with published ancient individuals from western Eurasia as well as genotyping data from the modern populations of the Human Origins panel (50). We only used published data obtained from UDG-treated libraries. Genotype was called for the 1.2 million

SNPs broadly used in ancient DNA in-solution capture procedures (22). A single allele was drawn at random for each position (minimum mapping and base quality of 30) using pileupCaller (<https://github.com/stschiff/sequenceTools>), rendering the individuals from the dataset homozygous for each locus. For genotyping data from modern individuals of the Human Origins panel, heterozygous sites in the eigenstrat file were randomly recoded as homozygous for the reference or the alternative allele using an awk script. The two files were then merged using PLINK version 1.9 (49).

Population Genetic Analyses. We conducted PCA on ancient individuals, projecting them onto the genetic variation of present-day Eurasians using the lsqproject option from smartpca (51). We computed *D* and *f* statistics using Admixtools to estimate gene flow and shared drift between populations and to test tree topologies (52). We used qpAdm to estimate the ancestry proportions attributable to early Anatolian farmers, hunter-gatherers, and steppe herders in the individuals of this study (13).

Data Availability. Sequencing data can be found in the European Nucleotide Archive (ENA) at EMBL-EBI under accession no. PRJEB38152.

ACKNOWLEDGMENTS. We thank the Service Régional de l'Archéologie from Occitanie, Grand Est, Hauts-de-France, ANTEA-archéologie, and Institut National de Recherche en Archéologie Préventive (INRAP) for providing bone materials. We are grateful to Olivier Gorgé for making shotgun sequencing possible at Institut de Recherche Biomédicale des Armées in the Division Défense NRBC (nucléaire, radiologique, biologique et chimique), Département des Services, Unité de Biologie Moléculaire. This work was supported by the Agence National de la Recherche (Grant ANR15-CE27-0001), INRAP (PAS [Projet d'Activités Scientifiques] Ancestra) and CNRS. The ARTEMIS program (French Ministry of Culture) funded the C14 datings. The paleogenomic facility of Institut Jacques Monod (IJM) obtained support from the University Paris Diderot within the program "Actions de Recherches Structurantes." The sequencing facility of IJM is supported by grants from University Paris Diderot, Fondation pour la Recherche Médicale (DGE20111123014), and Région Ile-de-France (11015901).

1. P. Bellwood *et al.*, First Farmers: The Origins of Agricultural Societies, by Peter Bellwood. Malden (MA): Blackwell, 2005. *Camb. Archaeol. J.* 17, 87–109 (2007).
2. K. Kristiansen, T. B. Larsson, *The Rise of Bronze Age Society: Travels, Transmissions and Transformations*, (Cambridge University Press, 2005).
3. Q. Fu *et al.*, The genetic history of Ice Age Europe. *Nature* 534, 200–205 (2016).
4. A. Keller *et al.*, New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nat. Commun.* 3, 698 (2012).
5. I. Lazaridis *et al.*, Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* 513, 409–413 (2014).
6. I. Olalde *et al.*, Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* 507, 225–228 (2014).
7. F. Sánchez-Quinto *et al.*, Genomic affinities of two 7,000-year-old Iberian hunter-gatherers. *Curr. Biol.* 22, 1494–1499 (2012).
8. P. Skoglund *et al.*, Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* 336, 466–469 (2012).
9. C. Gamba *et al.*, Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* 5, 5257 (2014).
10. I. Olalde *et al.*, The Beaker phenomenon and the genomic transformation of north-west Europe. *Nature* 555, 190–196 (2018).
11. V. Villalba-Mouco *et al.*, Survival of Late Pleistocene hunter-gatherer ancestry in the Iberian Peninsula. *Curr. Biol.* 29, 1169–1177.e7 (2019).
12. C. Valdiosera *et al.*, Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia. *Proc. Natl. Acad. Sci. U.S.A.* 115, 3428–3433 (2018).
13. W. Haak *et al.*, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522, 207–211 (2015).
14. A. Beau *et al.*, Multi-scale ancient DNA analyses confirm the western origin of Michelsberg farmers and document probable practices of human sacrifice. *PLoS One* 12, e0179742 (2017).
15. M. Lacan *et al.*, Ancient DNA suggests the leading role played by men in the Neolithic dissemination. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18255–18259 (2011).
16. M. Rivollat *et al.*, When the waves of European Neolithization met: First paleogenetic evidence from early farmers in the southern Paris Basin. *PLoS One* 10, e0125521 (2015).
17. M. Rivollat *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.* 161, 522–529 (2016).
18. J. Guillemin, D. Garcia, Eds., *La protohistoire de la France [in French]*, (Hermann, 2018).
19. J.-P. Demoule, *La révolution néolithique en France [in French]*, (La Découverte, 2007).
20. C. Constantin, D. Vachard, Anneaux d'origine méridionale dans le Rubané récent du Bassin parisien [in French]. *Bull. Soc. Préhist. Fr.* 101, 75–83 (2004).
21. M. Lipson *et al.*, Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* 551, 368–372 (2017).
22. I. Mathieson *et al.*, Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* 528, 499–503 (2015).
23. C. Posth *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).
24. E. R. Jones *et al.*, Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* 6, 8912 (2015).
25. I. Olalde *et al.*, The genomic history of the Iberian Peninsula over the past 8000 years. *Science* 363, 1230–1234 (2019).
26. P. Brun, "L'origine des Celtes. Communautés linguistiques et réseaux sociaux" in *Celtes et Gaulois. L'archéologie face à l'histoire. La préhistoire des Celtes [in French]*, D. Vitali, Ed. (Glux-en-Glenne, Bibracte, 2008), pp. 29–44.
27. J.-P. Demoule, Les Indo-Européens, un mythe sur mesure [in French]. *Recherche* 308, 40–47 (1998).
28. J.-P. Demoule, *Mais où sont passés les Indo-Européens?: Le mythe d'origine de l'Occident [in French]*, (Éditions du Seuil, 2014).
29. C. F. C. Hawkes, S. C. Hawkes, *Greeks, Celts, and Romans: Studies in Venture and Resistance*, (Rowman and Littlefield, 1973).
30. M. E. Allentoft *et al.*, Population genomics of Bronze Age Eurasia. *Nature* 522, 167–172 (2015).
31. S. Brace *et al.*, Population replacement in Early Neolithic Britain. *bioRxiv*:10.1101/267443 (18 February 2018).
32. P. G. Hysi *et al.*; International Visible Trait Genetics Consortium, Genome-wide association meta-analysis of individuals of European ancestry identifies new loci explaining a substantial fraction of hair color variation and heritability. *Nat. Genet.* 50, 652–656 (2018).
33. K. A. Hunt *et al.*, Newly identified genetic risk variants for celiac disease related to the immune response. *Nat. Genet.* 40, 395–402 (2008).
34. H. Yasuda *et al.*, Association of single nucleotide polymorphisms in endothelin family genes with the progression of atherosclerosis in patients with essential hypertension. *J. Hum. Hypertens.* 21, 883–892 (2007).
35. L.-T. Zhou, L. Qin, D.-C. Zheng, Z.-K. Song, L. Ye, Meta-analysis of genetic association of chromosome 9p21 with early-onset coronary artery disease. *Gene* 510, 185–188 (2012).
36. L. B. Barreiro *et al.*, Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense. *PLoS Genet.* 5, e1000562 (2009).
37. D. Massilani *et al.*, Past climate changes, population dynamics and the origin of bison in Europe. *BMC Biol.* 14, 93 (2016).
38. G. Renaud, U. Stenzel, J. Kelso, ILLUMINA: Adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res.* 42, e141 (2014).
39. R. M. Andrews *et al.*, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* 23, 147 (1999).

40. A. Ginolhac, M. Rasmussen, M. T. P. Gilbert, E. Willerslev, L. Orlando, mapDamage: Testing for damage patterns in ancient DNA sequences. *Bioinformatics* **27**, 2153–2155 (2011).
41. G. Renaud, V. Slon, A. T. Duggan, J. Kelso, Schmutzi: Estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* **16**, 224 (2015).
42. M. van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* **30**, E386–E394 (2009).
43. D. Vianello *et al.*, HAPLOFIND: A new method for high-throughput mtDNA haplogroup assignment. *Hum. Mutat.* **34**, 1189–1194 (2013).
44. D. Navarro-Gomez *et al.*, Phy-Mer: A novel alignment-free and reference-independent mitochondrial haplogroup classifier. *Bioinformatics* **31**, 1310–1312 (2015).
45. M. Kearse *et al.*, Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
46. A. Ralf, D. Montiel González, K. Zhong, M. Kayser, Yleaf: Software for human Y-chromosomal haplogroup inference from next-generation sequencing data. *Mol. Biol. Evol.* **35**, 1291–1294 (2018).
47. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
48. A. McKenna *et al.*, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
49. S. Purcell *et al.*, PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
50. I. Lazaridis *et al.*, Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).
51. A. L. Price *et al.*, Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
52. N. Patterson *et al.*, Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).