Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers

Maité Rivollat1,2*, Choongwon Jeong1,3, Stephan Schiffels2, Isabel Küçükkalıçıcı2, Marie-Hélène Pemonge1, Adam Benjamin Rohrlich2,4, Kurt W. Alt5,6, Didier Binder7, Susanne Friederich8, Emmanuel Hgesquiére9,10, Detlef Gronenborn11, Luc Laporte10, Philippe LeFranc2,13, Harald Meller8, Hélène Réveillas1,14, Eva Rosenstock15,16, Stéphane Rottier1, Chris Scarre17, Ludovic Soler1,18, Joachim Wahl19,20, Johannes Krause2, Marie-France Deguilloux1, Wolfgang Haak2*

Starting from 12,000 years ago in the Middle East, the Neolithic lifestyle spread across Europe via separate continental and Mediterranean routes. Genomes from early European farmers have shown a clear Near Eastern/Anatolian genetic affinity with limited contribution from hunter-gatherers. However, no genomic data are available from modern-day France, where both routes converged, as evidenced by a mosaic cultural pattern. Here, we present genome-wide data from 101 individuals from 12 sites covering today’s France and Germany from the Mesolithic (N = 3) to the Neolithic (N = 98) (7000–3000 BCE). Using the genetic substructure observed in European hunter-gatherers, we characterize diverse patterns of admixture in different regions, consistent with both routes of expansion. Early western European farmers show a higher proportion of distinctly western hunter-gatherer ancestry compared to central/southeastern farmers. Our data highlight the complexity of the biological interactions during the Neolithic expansion by revealing major regional variations.

INTRODUCTION

The Neolithic transition, which broadly describes the shift from foraging to farming, is one of the most important events in human history. In western Eurasia, the Neolithic way of life has been shown to spread westward from the northern Levant and potentially bypassing inner Anatolia from the seventh millennium BCE onward (1). Farming expanded along two main, archaeologically well-defined streams: along the Danube River to central Europe and along the Mediterranean coastline to Iberia (2, 3).

Recent large-scale genomic studies have shown that the spread of farming was mediated by early farmers through demic diffusion (4–11), while studies with a regional focus hinted at more complex regional processes of admixture between incoming farmers and local indigenous hunter-gatherers (HG) (6, 7, 12). The continental route is relatively well documented from southeastern to central Europe and shows a rapid expansion with very limited initial biological interaction, especially during the Neolithic Linear Pottery culture [Linearbandkeramik (LBK)] (5–8), followed by continued coexistence and evidence for cultural exchange with local HG for more than one millennium (13, 14). The Neolithic expansion in southwestern Europe is linked to the Mediterranean route (3), where a similar pattern of increasing admixture at least one millennium after the initial settlement of first farmer groups is observed in Iberia (5, 6, 11, 12, 13). The bigger picture emerging from these studies suggests little admixture between first incoming farmers and local indigenous HG in all targeted regions, followed by an increase of HG ancestry during an advanced phase of the Neolithic (5, 6). The contributing HG component detected in later Neolithic phases has been shown to be of local origin for the Carpathian Basin and Iberia (6, 12, 15).

From an archaeological perspective, it is not trivial to define and demonstrate the various modes of interaction between late HG and early farmers, and it requires high-resolution data and precise analytical methods to generate a clear and accurate spatiotemporal framework. Along the continental route, archaeological evidence of contact between Mesolithic and Neolithic groups increases toward the West (2). Overall, the lithic assemblages of the last HG are technically and stylistically rather similar to those of the early farmers. Signals of contacts with HG are mostly microlith types with regional traditions in the late Mesolithic blade-and-trapeze horizon. Archaeological signals of contact between immigrant farmers and HG have been reported not only from the earliest LBK sites in Hesse but also from Vaihingen in southwestern Germany. Here, coexistence between

1Université de Bordeaux, CNRS, PACEA-UMR, 5199 Pessac, France. 2Max Planck Institute for the Science of Human History, Department of Archaeogenetics, Jena, Germany. 3Seoul National University, School of Biological Sciences, Seoul, Republic of Korea. 4ARC Centre of Excellence for Mathematical and Statistical Frontiers, University of Adelaide, Adelaide, South Australia, Australia. 5Danube Private University, Krems, Austria. 6Integrative Prähistorische und Naturwissenschaftliche Archäologie, Basel, Switzerland. 7Université Côte d’Azur, CNRS, CEPIAM-UMR, 7624 Nice, France. 8State Office for Heritage Management and Archaeology Saxony-Anhalt—State Museum of Prehistory, Halle (Saale), Germany. 9Inrap Grand Ouest, Bourguébus, France. 10Université de Rennes 1, CNRS, CREAH-UMR, 6566 Rennes, France. 11Römisch-Germanisches Zentralmuseum, Leibniz-Forschungsinstitut für Archäologie, Ernst-Ludwig-Platz 2, 55116 Mainz, Germany. 12Inrap Grand Est Sud, Strasbourg, France. 13Université de Strasbourg, CNRS, Archimède-UMR, 7044 Strasbourg, France. 14Centre Archéologie préventive de Bordeaux Métropole, Bordeaux, France. 15Freie Universität Berlin, Institut für Prähistorische Archäologie, Berlin, Germany. 16Freie Universität Berlin, Einstein Center Chronoi, Berlin, Germany. 17Department of Archaeology, Durham University, Durham, UK. 18Service départemental d’archéologie de Charente-Maritime, Saintes, France. 19State Office for Cultural Heritage Management Baden-Württemberg, Osteologie, Konstanz, Germany. 20Universität Tübingen, Mathematisch-Naturwissenschaftliche Fakultät, Tübingen, Germany.
*Corresponding author. Email: maite.rivollat@u-bordeaux.fr (M.R.); haak@shh.mpg.de (W.H.)
farmers, HG, and groups with ceramics such as the debated La Hoguette and Limburg wares is documented within LBK settlements (2, 16). While evidence of coexistence within settlements is no longer observable, particular HG objects of adornment have been observed from burials in western central Europe, suggesting a continuous tradition of coexistence of these different populations throughout the Middle Neolithic (17).

Looking at the Western Mediterranean route, the main uncertainties revolve around the origin of the dispersal of the Impresso–Cardial complex (ICC) in Italy. The late HG settlements are concentrated in the northeast of the peninsula, while the earliest farmers appear in the south, with little geographic overlap (18). In contrast, the discontinuity between the latest HG and earliest farmers is more pronounced in southern France, due to the marked stylistic differences of the latest HG blade and trapezes industries and the earliest Neolithic toolkits. These discrepancies reasonably argue for a cultural and population shift. However, only a few stratigraphic sequences describe for regions in central or southern Europe (3). The presence of characteristic HG mitochondrial haplogroups (notably U5) than that observed from today’s France. A mitochondrial DNA study conducted on coastal human remains from the Late Mesolithic (3000–2800 cal BCE) (23) appears only in the southern Alps, at least three centuries after the pioneer colonization of the Mediterranean coasts.

When reaching the northwestern parts of Europe, the dispersal of the Neolithic way of life became more complex, as shown by archaeological research (2, 18), resulting in a highly differentiated picture of interaction and exchange from the Mesolithic to the Neolithic (including regional variations). There is an even more pronounced chronological gap between the colonization of southeastern France by the ICC pioneer groups (terminus post quem 5850 BCE) (22) and of northeastern France by the LBK (5300 BCE) (23). It was suggested that both the concurrence of the two main streams (24,25) of the west European Neolithic expansion up to the Atlantic coasts and the degree of interactions with autochthonous Mesolithic societies (3, 18) may have created the mosaic pattern of diversity seen in the material culture across the region in the subsequent centuries (text S1). This pattern has been well described at the cultural level, but no genomic data have been available to date, in particular from today’s France. A mitochondrial DNA study conducted on farmers from the Paris basin has highlighted a higher amount of characteristic HG mitochondrial haplogroups (notably U5) than described for regions in central or southern Europe (26), suggesting different processes in action.

Our study aims to cover the key geographic region of modern-day France and neighboring regions in Germany to unravel the complexity and variability in cultural and biological interactions between human groups during the earliest stages of the Neolithic period. This study region is uniquely suited as it encompasses both (i) the convergence of early farming communities from central Europe (Danubian route) with those from southern France (Mediterranean route) and (ii) variable forms of interaction with indigenous late HG.

RESULTS
Here, we present newly typed genome-wide data from 101 individuals from 12 sites from modern-day France and Germany (3 Late Mesolithic and 98 Neolithic, 7000–3000 cal BCE; table S1). We enriched next-generation ancient DNA (aDNA) libraries for ~1.2 million single-nucleotide polymorphisms (SNPs) using targeted in-solution capture (8), as well as an independent in-house capture array for the complete mitogenome ([27 modified following (6)], and sequenced these on Illumina platforms to an average depth per site of 0.6× for 1240K and 249× for mitochondrial capture. Libraries were built with a partial uracil–DNA–glycosylase (UDG) treatment, allowing the assessment of postmortem deamination patterns (2 to 29%) expected for aDNA data. We estimated potential contamination in mitochondrial capture DNA using (28) (table S3) and the nuclear genome by testing for heterozygosity of polymorphic sites on the X chromosome in males (table S2) (29). We estimated kinship to exclude first-degree relatives from downstream analyses (text S6). We coanalyzed our new data with published ancient (n = 629) and modern (n = 2583) individuals from a panel of modern-day worldwide populations genotyped on the Affymetrix Human Origins (HO) panel (31, 32) as well as with ancient individuals genotyped on the 1240K panel. We excluded two individuals with low SNP coverage on the HO panel (<1,000) from downstream genome-wide analyses (table S1). To support the interpretation of our genetic results with a robust chronological context, we report new direct radiocarbon dates for 30 individuals (text S2 and table S1). A small set of 13 SNPs associated with phenotypes of interest were investigated. Results are presented in text S7 and show expected genotypes for Mesolithic and Neolithic European individuals.

We explored our new data qualitatively using principal components analysis (PCA) by projecting the ancient samples onto the genetic variation of a HO set of west Eurasians (Fig. 1C). Newly sequenced Late Mesolithic individuals from Bad Dürenberg (BDB001) and Bottendorf (BOT004 and BOT005; see text S6 and fig. S1) fall within the variability of western HG (WHG). Individual TGM009, attributed to the Late Neolithic site of Tangermünde from an Elb–Havel context [Trichterbecherkultur (TRB); texts 1 and 2], occupies an intermediate position between WHG and Neolithic farmers. Newly sequenced Neolithic individuals cluster with previously published data, forming two regional subgroups: the first including individuals from Neolithic central and southeastern Europe and the second individuals from Neolithic western Europe (Iberian Peninsula, France, and British Isles), which is shifted slightly toward WHG on PC1. Formal f-statistics of the form f3(Atl, European_HG; test, Anatolia_Neolithic), where test represents Neolithic groups, confirm these observations via different degrees of affinity with European HG (table S9).

Early Neolithic individuals from Stuttgart–Mühlhausen (SMH), Schwetzingen (SCH), and Halberstadt (HBS) in Germany (44 published and 42 newly reported, of which 5 low-coverage individuals (<18,000 SNPs mapping on the HO panel) are shifted toward the centroid) representing the LBK horizon (17 sites from Hungary, Austria, and Germany) form a homogeneous genetic group of early farmer populations in central Europe. Neolithic groups from France cluster with western Neolithic individuals, while southern French sites (Penimoun (PEN) and Les Bréguières (LBR), attributed to the ICC, are shifted even further toward WHG. This suggests a higher HG component than in any other early Neolithic individuals, and thus a different history of admixture compared to contemporaneous western early farmer groups from Iberia. For both of the sites PEN and LBR, we split the individuals with significantly different HG ancestry contributions into subgroups A (less) and B (more HG ancestry; see table S7). Individuals from southern France do not group with ICC.
individuals from the Adriatic region, who fall within the subgroup of southeastern and central European farmers (Fig. 1C).

Both central and western groups are still present during the fifth and early fourth millennium BCE, suggesting a geographic border along the Rhine River for regions that are situated north of the Alps. Individuals from the Middle Neolithic French sites of Gurgy (GRG), Prissé-la-Charrière (PRI), and Fleury-sur-Orne (FLR) in the northern half of France appear homogeneous (one PRI individual outside of the cloud has low SNP coverage), while the individuals from the site Obernai (OBN) form three groups: one with other contemporary western individuals (OBN A), one with a stronger HG component (OBN B), and a third one with central European farmers (OBN C), despite sharing similar cultural and chronological backgrounds (Fig. 1C and text S2).

**Quantifying universal HG ancestry in European farmer groups**

We used qpAdm (ADMIXTOOLS) (5) to estimate the proportion of European_HG ancestry in all new and published Neolithic individuals with directly associated radiocarbon dates to track changes through time (MODEL A; Fig. 2, text S9, and table S10). We observe a notably similar pattern with negligible amounts of HG ancestry during the first periods of attested farming in each region of central and southwestern Europe, followed by a steady increase after centuries of established farming. The last individuals without any traceable HG ancestry disappear around 3800–3700 cal BCE. Southeastern Europe shows a specific pattern in individuals linked to the Iron Gates region, and the British Isles show a sudden and constant higher HG component at the time of the arrival of the Neolithic in the region. Our new western LBK groups (east of the Rhine) confirm the previous estimates from the Middle-Elbe-Saale region and Transdanubia, which is in contrast to southeastern European sites such as Malak Preslavets in Bulgaria, for which higher and more variable estimates were reported (6). However, the situation is different in today’s France, where we observe not only the highest HG ancestry proportion overall when compared to other regions in Europe but it is also found in the oldest individuals from the southern sites of PEN and LBR. This observation is also supported by uniparental markers. Y chromosome lineages in western early farmers in the southern region are exclusively derived from HG (I2a; table S5 and text S5). In contrast, mitochondrial DNA results show a more universal Neolithic diversity profile, as previously reported [i.e., (4, 5)], with only two haplotypes (U5 and U8) that are potentially of HG origin (table S4 and text S4).

Two individuals from the Middle Neolithic OBN site located immediately west of the Rhine (Fig. 1A) also show a high proportion of the HG ancestry component (OBN B), in contrast to LBK sites east of the Rhine. To quantify the extra-HG proportion, we modeled LBK from Germany as a mixture of Anatolia_Neolithic and European_HG, and subsequently OBN subgroups as a more proximal mixture.
model of LBK and European_HG. We find this model well supported (table S11), and obtained up to 31.8% of excess HG ancestry for OBN B. Given the absence of a strong HG contribution in LBK groups east of the Rhine, we assume this to be a local contribution during the centuries following the arrival of first farmers. Moreover, male individuals from OBN carry exclusively the Y chromosome haplogroups I2a1a2 and C1a2b, attributed to HG groups (text S5), providing further evidence for a greater amount of the HG contribution in this region.

**Genetic structure in European HG**

The shared recent ancestry of early European farmers with Anatolian Neolithic farmers (32) and the rapid expansion across mainland Europe render it challenging to distinguish between different Neolithic routes at the genomic level as proposed by archaeology. Although we observe an increase in HG ancestry from east to west in published regions, this might simply be the result of geographic distance from the farming origin in line with Lipson et al. (6). To test for signals of different routes of expansion, we therefore made use of the increasingly emerging geographic structure in European HG ancestry (6, 7, 15, 33). Here, post-glacial European HG ancestry can be described by three main clusters (Fig. 3 and fig. S1) with clines of admixture between each: (i) GoyetQ2-like ancestry, (ii) Villabruna-related WHG ancestry, and (iii) eastern HG (EHG) ancestry. In addition, two clines between the clusters can be observed: one between GoyetQ2 and WHG formed by Iberian HG, and the other between EHG and WHG formed by southeastern European, Scandinavian, and Baltic HG (7, 12, 15).

Following this observation, we also performed $f_4$-statistics of the form $f_4$(Mbuti, test; Villabruna, EHG) testing all HG individuals (fig. S2A). Here, negative $f_4$-values for all European groups indicate shared ancestry with the Villabruna individual and thus the WHG cluster, whereas positive values show more attraction to EHG. Using qpAdm, we modeled HG individuals as a mixture of the distal sources Villabruna, EHG, and GoyetQ2 to establish the best-fitting subclades of European HG individuals as sources of ancestry (MODEL B; text S9, fig. S4, and table S12). We also added Anatolia_Neolithic to fit some of the HG groups who show signs of admixture with early farmers. We chose the most parsimonious model, i.e., fitting the data with the minimum number of sources, for each set of results. In addition to the GoyetQ2 component present in Iberian individuals, as shown recently (15), we confirm that individuals from the EHG-WHG cline, including southeastern European and Scandinavian HG, can be distinguished from WHG individuals, as they carry a substantial proportion of EHG ancestry (7). Figure 3 shows that the newly reported Mesolithic individuals from Germany (BDB and BOT) also form part of the WHG cluster.
In the following, we leverage the observation that, at the time of the arrival of farming, this broad ancestry cline between EHG and WHG covers key geographical regions in central/southeastern Europe (spanning today’s Germany, Hungary, Serbia, Romania, and Ukraine), which overlap with the proposed Danubian route of the Neolithic expansion. Consequently, we hypothesize that Early Neolithic farmers of the Mediterranean route associated with ICC Ware and especially those arriving in France and Spain show less ancestry from the EHG side of the EHG-WHG cline and instead carry an admixture signal dominated by Villabruna-related WHG and GoyetQ2-related ancestries.

Tracing different HG ancestries in farmer groups

Analogous to the test with HG individuals, we calculated f-statistics of the form \(f_4(Mbuti, test; Villabruna, EHG)\) for all Neolithic groups (fig. S5A) to estimate the affinity of Neolithic groups (test) to the distal HG representatives of the east-west cline. Like the HG groups (fig. S2A), the \(f_4\) values from the Neolithic groups are mostly negative, indicating shared ancestry with Villabruna, but form a wide gradient reaching 0, which implies equal amounts of shared ancestry with EHG for some groups located in southeastern Europe. We used qpAdm to quantify these sources in MODEL B (text S9) for all Neolithic European groups. However, using these distal sources due to the very small amount of HG component in many early Neolithic groups, the model fails to reliably detect an EHG component (table S12).

We therefore restricted the analysis to more proximal sources. We chose the deeply covered Mesolithic individual Loschbour from Luxembourg as representative of WHG ancestry (west of the Rhine) and therefore as a proxy for the Mediterranean route (34). We selected the Hungarian individual KO1, which looks genetically like a HG, though found in a farming context, for the HG ancestry of the continental route (10). We first replicated the \(f_4\)-statistics of the form \(f_4(Mbuti, test; Loschbour, KO1)\) (fig. S5B). Here, positive \(f_4\)-values indicate shared ancestry with KO1 as a proxy for the EHG-WHG cline. In turn, negative \(f_4\)-values indicate excess shared ancestry with Loschbour and thus the WHG cluster. The results show a tendency for farmer groups east of the Rhine to share more ancestry with KO1 and farmer groups west of the Rhine with Loschbour. Notable exceptions are individuals from the Middle Neolithic Blätterhöhle group (6, 13), individual N22 from Poland (35), and our individual TGM009, all of which show a strong affinity to WHG-related individuals, although they are located east of the Rhine. On the basis of our geographic rationale, we then used qpAdm to model HG components in the Neolithic groups, including both HG sources that could be encountered en route for the proposed expansions (MODEL C; text S9, fig. S6, and table S13). We chose the best-fitting model given by qpAdm in cases where both HG sources were supported. Figure 4 summarizes these results, illustrating the diverse HG ancestries in time windows across major regions in Europe, adapting the models used to the targeted groups (table S14).
We caution that because the amount of HG ancestry is very low in many Neolithic groups (<10%), it remains difficult to characterize the ultimate source reliably. Nevertheless, our admixture patterns from supported models show clear geographic signals. Neolithic groups associated with the LBK in central Europe (Hungary, Austria, and Germany) carry a small HG proportion, which was likely derived from admixture with HG individuals of the EHG-WHG cline and could have occurred in southeastern Europe during a preceding phase of the Neolithic expansion around 6000–5400 BCE.

When using f-statistics of the form f4(Mbuti, test; BDB001, KO1) with our new Mesolithic genome (BDB001) from the Middle-Elbe-Saale region in central Germany as a geographically local HG proxy (instead of using Loschbour, which is located west of the Rhine), we do not find support for a local attraction for LBK groups, but the same pattern as for Loschbour (fig. S5C). This suggests that additional gene flow from neighboring Loschbour-like HG such as BDB001 in central Europe was negligible in the first Neolithic groups. However, the German Baalberge group (4000–3500 BCE) shows a marked increase of such HG ancestry, as well as individuals from the Blätterhöhle group, as has been suggested (5, 6), compared to a combination of both KO1-like and Loschbour-like ancestries for LBK groups (6). We can now show that this increase in WHG ancestry (up to 21.3 ± 1.5% in Baalberge; table S13) is driven by either local Loschbour-like ancestry or an expansion of farming groups from the west carrying this signal during the fifth millennium BCE, as suggested by archaeological data (36). For all studied Neolithic groups west of the Rhine, we observe a different pattern with a higher HG ancestry proportion, even for earliest groups that appears to be of a local (Loschbour-like) HG origin, consistent with archaeological data (2).

Late survival of HG ancestry in central Europe
In contrast to Middle-German Neolithic groups, the Late Neolithic individual TGM009 (~3300 BCE) found in an TRB/Elb-Havel group context shows a different pattern of HG ancestry. When modeled...
with distal sources, this individual carries 63.6 ± 5.2% Villabruna-related ancestry (see MODEL B; table S12 and Fig. 4). When we use proximal sources (MODEL C; table S13), we estimate the HG component to be split into 48.1 ± 6.4% KO1-related ancestry and 25.8 ± 6.1% Loschbour-like ancestry \( (P = 0.12) \). To test whether the subtle EHG-related signal observed in TGM009 (through KO1) could also come from Scandinavian HG, as suggested by the archaeological records of regional late contacts with the Scandinavian Mesolithic (37), we applied \( f \)-statistics of the form \( f_4 \) (Mbuti, test; Hungary_KO1, Sweden_Motala_HG). The \( f_4 \) value is significantly negative \( (f_4 = -0.0035, Z \text{ score} = -6.383) \), indicating that TGM009 shares more ancestry with southeastern Europe HG than with Scandinavian HG (table S16). Consequently, qpAdm models with Sweden_Motala_HG in place of Hungary_KO1 received a poor fit \( (P = 1.32^{-04}; \text{table S16}) \).

In the specific case of TGM009, we also tested whether individuals of the Pitted Ware Culture (PWC) group (~3200–2300 BCE) (38), considered to be “Neolithic HG,” would be a suitable contemporaneous proximal source with MODEL C set of outgroups (table S16). We find that the tests still support the three-way model with Anatolia_Neolithic, Luxembourg_Loschbour, and Hungary_KO1 as a third source \( (P = 0.12) \). However, when we add Sweden_Motala_HG to the outgroups, we find the best support for a four-way mixture model \( (P = 0.087) \) of 21.7 ± 2.4% Anatolia_Neolithic, 24.4 ± 6.2% Luxembourg_Loschbour, 12.6 ± 4.4% Sweden_PWC, and 41.3 ± 7.3% Hungary_KO1. The small but stable contribution of PWC groups at the northern fringes of the loess belt adds additional complexity to the modes of interaction between Late Mesolithic and the last HG groups in central Europe.

**Estimating admixture dates between HG and early farmers**

To gain further insights into the timing of the admixture of the HG ancestry proportion, we used the software package DATES to estimate the timing of admixture events between early farmers and HG, given in generations in the past (Fig. 5, table S17, and text S10). The resulting date estimates, when ordered temporally and geographically, echo patterns of both groups that are also visible in the PCA (east versus west of the Rhine). It confirms that this pattern is dependent not only on the amount of HG ancestry but also on the qualitative signature of the HG ancestry, as supported by fig. S7. Together, the date estimates for the four southern French groups from sites PEN A and B and LBR B suggest admixture with local HG relatively soon (100 to 300 years) after the arrival of Neolithic farmers about 5850 cal BCE, with an admixture date about 5740–5450 cal BCE for all groups (table S17). These date estimates agree with local archaeological data for the establishment of early farming in this region \( (18) \), although we cannot exclude a scenario in which admixture occurred on the Italian Peninsula shortly before.

In the rest of France, post-LBK groups show an older admixture date, which, together with a strong local HG component, place the admixture event with WHG-related HG during the first phase of the local Neolithic (Fig. 4). In contrast, the site OBN reveals a heterogeneous genetic signal forming three distinguishable groups with different HG proportions and corresponding admixture dates (tables S7, S13, and S17). The differences between the date estimates for the three groups suggest that the OBN group shows a genetic substructure rather than a recent ongoing admixture event.

**Tracing GoyetQ2 ancestry in France**

We also used qpAdm with GoyetQ2, Villabruna, and Anatolia_Neolithic as sources to track Magdalenian-associated GoyetQ2-like ancestry in the new western Neolithic groups (MODEL D; text S9 and table S15) \( (15) \). The GoyetQ2-related component is observed in the sites of PRI (6 ± 3%) in western France and GRG (3.7 ± 1.3%) from the Paris basin. Similar to Iberian Neolithic groups, the PRI group (4300–4200 BCE) can be modeled successfully with GoyetQ2 as one of the HG components (fig. S8 and table S15), suggesting either admixture with local HG in western France who retained this post-Last Glacial Maximum component or genetic contacts with Neolithic Iberia in later phases. Of note, individuals from the site PRI on the Atlantic coast indicate a slightly more recent admixture date with European_HG (about 5200 BCE), which is consistent with a later arrival of Neolithic groups in westernmost France (table S17). We plotted these results in fig. S9 to complement the summary map with the GoyetQ2-related component in Neolithic groups, which adds another layer of complexity to the overall genetic picture of Neolithic Europe.

**Connections with Britain and Ireland**

To address questions at broader scale on the western fringe of Europe, we also investigated the relationship between Britain and Ireland and the European mainland during the Neolithic. Here, we performed \( f_4 \)-statistics of the form \( f_4 \) (Mbuti, Britain/Ireland groups; PRI, test), which measures whether Neolithic groups from Britain and Ireland share more genetic drift with the French Atlantic group PRI than with a test population (see fig. S10). As previously described \( (39, 40) \), we confirm that British Neolithic groups share affinities with the Mediterranean Neolithic (LBR_A, France_MN, and Iberia_MN), which is also visible in the WHG-rich HG proportion (Fig. 4). However, on the basis of the results from our French Neolithic sites, we suggest that English, Welsh, and Scottish groups are connected to the Mediterranean Neolithic sphere not only via the Atlantic coast but more plausibly also via Normandy (FLR), the Paris Basin (GRG), and southern France (LBR_A, post-ICC groups).

**DISCUSSION**

The expansion of Anatolian Neolithic farmer–related ancestry across Europe has been described for many geographic regions \( (5–7, 9–11, 39, 40) \). A recurrent pattern of increasing HG gene flow in individuals associated with a farming lifestyle many centuries after initial contact and tentatively termed “HG resurgence” \( (5) \) has hitherto been observed in the Iberian Peninsula, northern/central Europe, the Carpathian Basin, and the Balkans \( (6) \). However, a comparison of genomic data in the light of the proposed main routes of Neolithic expansion has not yet been attempted. Our new results provide important insights for the region of modern-day France where both routes had intersected. This region has not been documented so far, but is ideally positioned to address these questions. The different Mesolithic genetic substratum in Europe \( (6, 7, 15, 33) \) enables us to track both Neolithic expansion routes in the form of the quantity and quality of the admixed HG component observed in Neolithic groups.

Neolithic groups from southern France, which are a part of the Mediterranean/ICC route of expansion, show a different genetic profile when compared to early periods of expansion in other regions, with a substantially higher HG component than groups associated...
with the continental route (up to 56 ± 2.9% of HG component for PEN B; table S10). Although the sites PEN and LBR date ~400 years (i.e., ~16 generations) later than the first local coastal settlements in Liguria, Provence, and Mediterranean Languedoc, they suggest a recent local admixture event (between three and six generations earlier; Fig. 5 and table S17). Archaeological research has argued for increased interaction between incoming farmers and indigenous HG in the western Mediterranean during a second stage of the Neolithization process and especially in areas with higher HG population densities, e.g., the Tosco-Emilian Apennine and Po plain (18). We are now able to confirm that these contacts left a traceable biological signal during the Neolithic expansion in southern France. From an archaeological perspective, this suggests that HG have contributed to the clear changes observed within the material culture postdating the pioneer phase. Note that ICC individuals from the eastern Adriatic coast have only a very small amount of HG ancestry with a greater affinity to central European groups (see table S8). This fits with the hypothesis of a differentiation of technical traditions within material cultures observed from both sides of Apennine Mountains in Italy: an Adriatic tradition connected to the Balkans and a Tyrrhenian one whose origin is still unknown (41). It is tempting to associate such a strong HG component on the Tyrrhenian side with the characteristic/specific pottery traditions observed in this same region and to consider these original traditions the result of a HG reinterpretation (41). However, the scarcity of genomic data available from central and southern Italy currently does not allow this hypothesis to be tested directly. Moreover, ICC individuals from the Iberian Peninsula also carry less HG ancestry. Together, this rejects the hypothesis that ICC-associated individuals represent a uniform genetic horizon per se and argues for more regionally nuanced scenarios of interaction.

**Fig. 5. Admixture dating per group.** Admixture date estimation according to DATES software, obtained in generations. Radiocarbon date intervals (given with 2-sigma) for each site are black lines; blue diamonds are the estimated admixture date (SE = 1). Admixture is estimated with two sources: European_HG and Anatolia_Neolithic. Number of years calculated on the basis of 28 years for one generation. Admixture time calculated according to the oldest date of the radiocarbon interval for each group (table S17 and text S10).
Early central European farmers carried a very small percentage of HG component (about 5% on average), which is most likely nonlocal and instead derives from admixture in Transdanubia during an early stage of expansion. This not only matches general observations in the archaeological record (2) but also could explain why the earliest LBK lithic assemblages resemble those of the Late Mesolithic Blade and Trapeze Complex. It also confirms previous aDNA studies (5, 6), which argued for a fast spread of the first farmers across the German loess regions. LBK groups from southwestern and eastern Germany share more affinities with KO1 than with Loschbour (fig. S5). The admixture date estimates for southern German sites, SMH and SCH, 19.2 ± 3.8 and 12.3 ± 8.2 generations, respectively, are younger or contemporaneous to admixture date estimates from the Carpathian Basin and Austria (Fig. 5 and table S17). The temporal lag, and the subtle increase of shared KO1-like HG ancestry, allows us to trace LBK groups chronologically and matches well with the proposed model of LBK expansion from the core region in Transdanubia as suggested by archaeological research (14). However, the present picture does not explain the increasing amount of archaeological evidence of contact with HG and groups with southern influences in more western sites, particularly during the early LBK, such as the debated La Hogue and Limburg phenomena (16).

In contrast to the situation in central Europe, regions west of the Rhine show a different profile during the following fifth millennium BCE. Here, first farmers carry a higher local Loschbour-related HG component, later increasing to up to 33.3 ± 3% in several individuals in Alsace (OBN B). Mitochondrial data support this finding with a higher average proportion of HG-affiliated haplotypes (U5 and U8) in all French groups from the fifth millennium BCE (15.5%; table S4 and text S4) compared to LBK groups east of the Rhine (1.4%). Although we do not have genomes from LBK individuals from northern France, we can infer from this HG component that admixture processes happened locally after the first farmers arrived. When modeling the European HG component in the three groups from OBN with qpAdm, we observe an increase in the HG component of between 4.3 and 31.8%, which we attribute to local Loschbour-related sources (table S11). However, this approach cannot be directly applied to other contemporaneous French groups as potential additional movements linked to the Mediterranean route of Neolithic expansion that are supposed to have followed the first LBK farmers’ arrival in northern France could complicate the simple assumption of a two-way HG/farmer interaction (26, 42). The current genomic data do not allow us to distinguish whether the detected signal is confounded by a potential southern contribution. However, the estimated admixture dates for GRG and FLR in northern France indicate older admixture events that occurred more than 30 generations (840 to 930 years) earlier (Fig. 5 and table S17). In accordance with the established chronology of first Neolithic settlements in the French territory, the overlapping/synchronous date estimates obtained for southern ICC sites are consistent with the signal of a first HG contribution in the south of France, followed by a subsequent northward expansion of groups carrying this HG legacy (42).

The GoyetQ2-like HG component detected in individuals from the sites PRI and GRG suggests connections with the Iberian Peninsula where this post-LGM residual component is found, either by an admixture with HG (which could also be local) carrying this component or by contact with first farming communities from Iberia or because of exchanges with Neolithic groups from Iberia during the subsequent fifth millennium. To date, archaeological data are compatible with all three hypotheses (25, 43). However, while we can show affinities between PRI on the Atlantic coast and the Iberian Peninsula, the previously described affinities between the British Isles and Mediterranean Neolithic (39, 40) are currently best explained by sites mainly in Normandy and the Mediterranean area via the Paris Basin. English, Scottish, and Welsh groups show more genetic affinities to northern and southern France and Iberia than to western France (fig. S10). In contrast, Neolithic Ireland shows less affinities with the northern French coastline and the Mediterranean area than other British groups, which could be explained within an Atlantic framework. This whole pattern is in line with the hypothesis of two different phenomena and speeds of Neolithic expansion to the western and eastern British Isles, as proposed by archaeological data (44).

In summary, our study highlights a diverse pattern of cultural and biological interactions between first farmers and indigenous HG along the western Mediterranean coastline and west of the Rhine, which confirms a high variability of processes during the Neolithic expansion, distinct in the proportion of the HG component as well as processes and duration. The genetic structure among HG groups allowed us to track local admixture in early farmers, which not only is higher west of the Rhine compared to central and southeastern Europe but also is largely attributed to local and distinctly WHG-related sources. We show that nuanced sampling as well as increasing cohort numbers of both HG and early farmers can help to unravel the regional dynamics of the Neolithic transition and aid in refining our understanding of the underlying processes and developments over time, at both micro- and macro-regional scales. On the basis of our observations, we find that broad-brush models are increasingly less likely to reconcile the full spectrum and details of the farmer-forager interactions and thus advocate the use of models with more regional focus in future studies.

**MATERIALS AND METHODS**

**Study design**

**Archaeological samples**

We processed and analyzed samples from 101 individuals, 3 Mesolithic and 98 Neolithic (7000–3300 cal BCE), from 12 sites in modern-day France and Germany. A detailed description of each site is provided in the Supplementary Materials, including contextual data and available radiocarbon dates (text S2 and table S1).

We sampled, extracted, and prepared DNA for next-generation sequencing in three different dedicated aDNA facilities. Individuals from PRI and PEN were sampled in the clean room of the Laboratoire PACEA, Bordeaux University, France. Samples coded “XN” from the site SMH were processed at the Interfaculty Centre for Archaeology at the University of Tübingen, Germany, and all other samples were processed at the Max Planck Institute for the Science of Human History (MPI-SHH) in Jena, Germany.

**Sampling**

We sampled mainly petrous bones and tooth and, for a few individuals, also other type of skeletal remains (femur and phalanges; table S1). Samples were first irradiated with ultraviolet light for 30 min on all sides. All teeth were cleaned with a low-concentration bleach solution (3%). Teeth prepared in Tübingen and Jena were cut along the cementum/enamel junction, and powder was collected by drilling into the pulp chamber. Teeth prepared in Bordeaux were ground to...
fine powder completely. Petrous bones were either cut in half and powder drilled from the denser regions around the cochlea (GRG) or drilled from the outside (FLR, LBR, OBN, and PEN). Before that, a layer of the bone surface was mechanically removed.

**DNA extraction**
From the extraction step onwards, all samples were processed at MPI-SHH in Jena, except the samples labeled XN, which were processed at the University of Tübingen. At MPI-SHH, DNA was extracted following the protocol described by Dabney et al. (45), with the High Pure Viral Nucleic Acid Kit (Roche). The same protocol was used in Tübingen with MinElute columns (Qiagen).

**Library construction**
One hundred and four double-stranded libraries were built with unique index pairs (46, 47). We applied the partial UDG (UDG half) protocol to remove most of the aDNA damage while preserving the characteristic damage pattern in the terminal nucleotides (48). A separate set of non-UDG libraries prepared in Tübingen was used for the mito-capture of SMH samples (table S3).

**Shotgun screening and capture (1240K and mitochondrial)**
We first screened all indexed libraries via shotgun sequencing of 5 million reads on an Illumina HiSeq 4000 sequencer using either a single [1 × 75–base pair (bp) reads] or double end (2 × 50–bp reads) kit. We used EAGER (49) to process the raw data and to select libraries with >0.1% endogenous human DNA and those showing characteristic damage aDNA patterns for downstream SNP capture. Selected libraries were hybridized in-solution to different oligonucleotide probe sets synthesized by Agilent Technologies to enrich the consensus mitogenome of the ancient sample to a panel of 311 unique index pairs (text S8).

**aDNA data processing**

**Read processing and aDNA damage**
After demultiplexing, raw sequence data were processed using EAGER. This included clipping adaptors with AdapterRemoval (50), mapping with BWA (Burrows-Wheeler Aligner) v0.7.12 (51) against the Human Reference Genome hs37d5, and removing duplicate reads with the same orientation and start and end positions. After using mapDamage v.2.0.6 to observe characteristic aDNA damage patterns, we used BamUtil (https://genome.sph.umich.edu/wiki/BamUtil:_trimBam) to clip two bases at the ends of each read for each sample to remove residual deaminations and 10 bases for reads from the non-UDG libraries performed for mito-capture on SMH samples (tables S2 and S3).

**Sex determination**
Following Mittnik et al. (52), we determined the genetic sex by calculating the number of reads mapping to each of the sex chromosomes with respect to the autosomes. We set a threshold of $Y$ ratio < 0.05 for a female and $Y$ ratio > 0.4 for a male (text S3 and table S2).

**Contamination estimation**
We used the ANGSD (Analysis of Next Generation Sequencing Data) package to test for heterozygosity of polymorphic sites on the X chromosome in male individuals, applying a contamination threshold of 5% (table S2) (29). For mito-captured samples, we estimated contamination levels using contamMix 1.0.10 (53) by comparing the consensus mitogenome of the ancient sample to a panel of 311 worldwide mitogenomes as a potential contamination source (table S3). Fourteen individuals from FLR and SMH gave a rate of contamination over 5%, but while checking the read assembly visually, we found no consistent pattern of contaminating mitochondrial DNA lineages. On the basis of low contamination estimates for males on the X chromosome—when available—we opted to include these individuals in the downstream analyses. Libraries used for mito-capture for SMH samples are non-UDG–treated, and hence, damage is also biasing our contamination estimates (53).

**Genotyping**
We genotyped our bam files with pileupCaller (https://github.com/ stschiff/sequenceTools/blob/master/src/SequenceTools) by randomly calling one allele per position considering the human genome as pseudo-haploid genome. We called the SNPs according to the Affymetrix HO panel (~600K SNPs) (31, 32) and the 1240K panel (8). Numbers of SNPs covered at least once are given in table S1.

**Mitochondrial and Y chromosome haplogroup assignment**
To process mitochondrial DNA data, we extracted reads from the aDNA data processing when available or from 1240K data using samtools v1.3.1 (54) and mapped these to the revised Cambridge reference sequence. We called consensus sequences using Geneious R8.1.974 (55) and used HaploGrep 2 to determine mitochondrial haplotypes (text S4 and table S4) (56).

We used pileup from the Rsamtools package (57) to call the Y chromosome SNPs of the 1240K SNP panel from all male individuals (mapping quality ≥30 and base quality ≥30). We manually assigned Y chromosome haplogroups using pileups of Y-SNPs included in the 1240K panel that overlap with SNPs included on the ISOGG SNP index v.14.07 (text S5 and table S5).

**Kinship analysis**
We estimated the degree of genetic relatedness between our individuals by applying Relationship Estimation from Ancient DNA (READ) (text S6) (30).

**Phenotypic and functional analysis**
We investigated individual genotypes for 13 SNPs associated with phenotypes of interest (8) by calculating genotype likelihoods based on the number of reads from our bam files (quality <30) for each specific position to determine the presence of the ancestral (non-effect) or derived (effect) alleles. We restricted the analysis to individuals with >200,000 SNPs mapping to the HO panel (text S7).

**Population genetic analysis**

**Merging new dataset with published data**
Two datasets were created for genome-wide analysis. We first merged our new data with published ancient data to the HO panel (~600K SNPs) (31, 32). This dataset was only used for PCA (see below). We then merged our data with published ancient data to the 1240K SNP panel (8) including 300 present-day individuals from 142 populations sequenced to high coverage (58). This second dataset, restricted to the autosomes, was used for all other population genetic analyses.

**Group labels**
Table S6 describes analysis groups and labels, for new, as well as published ancient samples.

**Principal components analysis**
We performed PCA using the “smartpc” program v10210 (EIGENSOFT) on the HO dataset (Fig. 1C) (59). We computed principal components from 777 present-day west Eurasians on which ancient individuals were then projected using the options lsqproject: YES and shrinkmode: YES. We excluded individuals with less than 10,000 covered SNPs. A zoomed-out version of the PCA is presented in text S8.
f-Statistics

Outgroup f3-statistics were calculated using qp3Pop and f4-statistics using qpStat with the f4 mode from ADMIXTOOLS (31) on the 1240K SNP panel. To investigate HG diversity, we performed outgroup f3-statistics of the form f3(test, test; outgroup) to create a similarity matrix, which was then used to generate the heatmap using the heatmap.2 function of the R-package gplots (fig. S1) (60). By calculating the values of 1-f3, we created a dissimilarity matrix from our f3-statistics. We performed multidimensional scaling on two dimensions for the dissimilarity matrix and then plotted the principal dimensions (Fig. 3). We computed SEs using the default block jackknife approach. We report and plot three standard errors for the f4-statistics test in supplementary figures (s2, s3, s5, and s9).

qpAdm modeling

We used qpAdm and the 1240K SNP panel to estimate ancestry proportions (ADMIXTOOLS) (5) with respect to a basic set of 11 outgroups: Mbuti, Papuan, Onge, Han, Karitiana, Mota, Ust_Ishim, MA1, Czech_Vestonice, Caucasus_HG, and Israel_Natufian. We then added additional outgroups according to the tests performed (see text S9 and tables S10 to S13 and S15 for further details).

Admixture date estimation

We used the method DATES v.753 to leverage patterns of ancestry covariance to estimate the date of admixture between European_HG and Anatolia_Neolithic (see text S11 and table S17 for further details).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://advances.sciencemag.org/cgi/content/full/6/22/eaa5344/DC1

REFERENCES AND NOTES


Downloaded from https://advances.sciencemag.org on February 5, 2021


94. G. Vernet, *Le Néolithique ancien (épicaudial) et le site Néolithique moyen/ancien sur le site du Brézet à Clermont-Ferrand (Massif central, France). Rencontres Méridionales de Préhistoire Récente, 85–90 (2004).*


105. J. Grünberg, *The Mesolithic burials of the Middle Elbe-Saale region, in Mesolithic Burials—Rites, Symbols and Social Organisation of Early Postglacial Communities, International Conference Hallé (Saale), Germany, 18 to 21 September 2013.*

106. M. Stecher, J. Grünberg, K. W. Alt, *Bioarchaeology of the Mesolithic individuals from Bottendorf (Thuringia, Germany), in Mesolithic Burials—Rites, Symbols and Social Organisation of Early Postglacial Communities, International Conference Halle (Saale), Germany, 18 to 21 September 2013.*


Submitted 17 September 2019
Accepted 23 March 2020
Published 29 May 2020
10.1126/sciadv.aaz5344

Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers

Maité Rivollat, Choongwon Jeong, Stephan Schiffels, Isil Küçükkalipçi, Marie-Hélène Pemong, Adam Benjamin Rohrlach, Kurt W. Alt, Didier Binder, Susanne Friederich, Emmanuel Ghesquière, Detlef Gronenborn, Luc Laporte, Philippe Lefranc, Harald Meller, Hélène Réveillas, Eva Rosenstock, Stéphane Rottier, Chris Scarre, Ludovic Soler, Joachim Wahl, Johannes Krause, Marie-France Deguilloux and Wolfgang Haak

Sci Adv 6 (22), eaaz5344, DOI: 10.1126/sciadv.aaz5344