Article:
Noraz, Rémi, Chauvey, Lorelei, Wagner, Stefanie et al. (2026) Ancient DNA reveals 4000 years of grapevine diversity, viticulture and clonal propagation in France. Nature Communications. 2494. ISSN: 2041-1723

https://doi.org/10.1038/s41467-026-70166-z

nature communications

Article

https://doi.org/10.1038/s41467-026-70166-z

Ancient DNA reveals 4000 years of grapevine diversity, viticulture and clonal propagation in France

Received: 9 August 2025

Accepted: 20 February 2026

Published online: 24 March 2026

Check for updates

Rémi Noraz[1,2], Lorelei Chauvey[1], Stefanie Wagner[1,3], Oscar Estrada[1], Nathan Wales[1], Emmanuelle Bonnaire[4,5], Manon Cabanis[6,7], Marie Derreumaux[4,8], Isabel Figueiral[2,6], Charlotte Hallavant[9], Philippe Marinval[10], Véronique Matterne[4], Jérôme Ros[2], Nuria Rovira[10], Margaux Tillier[10,11], Marilyne Bovagne[6,10], Jean Chausserie-Laprée[12], Clément Nicolas[13], Isabelle Daveau[6], David Delassus[14], Eric Gailledrat[6,10], Olivier Ginouvez[6], Bertrand Houix[6], Cécile Jung[6], Philippe Kuchler[15], Laure Leroux[9], Raphaël Macario[9], Elise Marlière[16,17], Michel Pasqualini[18], Gaël Piques[6,10], Hervé Pomaredes[6], Justine Robert[9], Cédric Roms[6], Muriel Roth-Zehner[5,19], Maxime Scrinzi[13], Maryanick Taras-Thomas[6], Josep Torres Costa[16,20], Guillaume Besnard[21], Jean-Frédéric Terral[2], Roberto Bacilieri[22], Laurent Bouby[2] ✉ & Ludovic Orlando[1] ✉

Wine production has deep historical roots in France, yet the biological foundations of early viticulture remain elusive. We report genome-wide ancient DNA data from 49 archaeological grape pips spanning ~4000 years, from the Bronze Age to the Medieval period, primarily focusing on France. Population genetic analyses reveal the genetic makeup of wild local grapevines in the Bronze Age and the early use of domesticated grapevines by ~625-500 BCE. Roman-era genomes reflect long-distance exchange from Iberian, Balkan, Levantine, and Caucasian domesticated varieties. Vegetative propagation, evidenced by genetically identical clones across sites and centuries, emerged by the mid-Iron Age and became a pillar of viticultural practice. Remarkably, one Medieval sample from Valenciennes is genetically identical to modern 'Pinot Noir', demonstrating clonal continuity over nearly 600 years. This study traces the introduction, integration and spread of diverse grapevine ancestries that have contributed to shape the varietal landscape of French viticulture.

Wine has played a central role in the cultural, social, and economic life of Europe since the first trace of viniculture ~8000 years ago. From the Protohistoric period onward, it became a highly valued commodity, symbolically charged and widely exchanged across long distances. The grapevine (*Vitis vinifera* subsp. *vinifera*, hereafter referred to as *V. vinifera*), thus, emerged as a key species not only in agriculture but also in shaping landscapes, cultural practices, and symbolic systems[1–5].

France, in particular, has long been a major hub in the development and diffusion of viticulture, positioned at the intersection of Mediterranean and continental European exchange networks[6]. Today, it remains one of the world's leading wine-producing countries, with viticulture forming a cornerstone of its cultural heritage and rural economy. The modern French wine industry contributes tens of billions of euros annually, supports hundreds of thousands of jobs, and shapes international perceptions of wine quality and tradition[7].

A full list of affiliations appears at the end of the paper. ✉e-mail: laurent.bouby@umontpellier.fr; ludovic.orlando@utoulouse.fr

Despite this deep legacy and ongoing economic importance, the biological and cultural foundations of French winemaking, especially in its earliest phases, remain poorly understood.

Over the past two decades, archaeology has made significant progress in reconstructing the history of viticulture and winemaking in France. A diversity of studies has shed light on wine trade routes, the spread of viticulture during Romanization, production sites and techniques, and changes in vineyard organization and planting systems[8–11]. These investigations have also informed on the socio-economic and cultural dimensions of wine production and consumption in ancient Gaul and neighbouring regions[4,5,12–14]. Recent work integrating archaeogenomics has further advanced our understanding of grapevine domestication and diversity in France. In particular, sequence data from over 10,000 single nucleotide polymorphisms (SNPs) comprised in the Vitis18kSNP genotyping array[15], obtained from 28 archaeological grape seeds, has started to trace the population dynamics of French grapevine cultivars from the Iron Age and through the Roman and Medieval periods[16]. The study identified broad genetic continuity with western cultivars used for making wine today and connected a historic ~1100 CE cultivar from Orléans (Loire valley) with its modern clonal descendant 'Savagnin Blanc'.

Despite these advances, fundamental questions remain unanswered, particularly regarding the diversity and biological characteristics of ancient grapevines and the nature of the wines produced before the Middle Ages. The viticultural practices that shaped regional varietal identities long before modern ampelographic classification systems, including the role played by vegetative propagation, remain also unknown. Knowledge of ancient viticulture largely relies on fragmentary and often biased historical sources, mostly written from elite urban centres rather than agricultural regions. Additionally, present-day or historical analogies are often used to infer the presence of particular cultivation practices in the deeper archaeological record; the validity of these analogies yet remains untested[13]. Given these limitations, archaeobotanical approaches focusing on ancient plant remains, thus, provide the gold standard approach for retracing the origins and evolution of viticulture.

In this study, we applied cutting-edge ancient DNA research to analyze whole-genome variation of 54 archaeological grape pips, including 49 newly sequenced samples from France and neighbouring Mediterranean territories. Our primary focus was on pips from Bronze Age to Middle Ages, allowing us to trace the genetic evolution of grapevines over approximately ~4000 years up to the present day. We identified the earliest occurrence of genetic ancestries typical of domesticated grapevines during the Iron Age, around ~625–500 BCE. While our analyses confirmed genetic affinities with modern western wine cultivars, we also detected genetic ancestries associated with grapevine lineages from the Levant and Caucasus region. The genetic landscape of grapevines in France became markedly diverse during the Roman period. Our findings highlight the importance of vegetative reproduction, particularly clonal propagation, as a central element of viticultural practices. We identified genetically identical clones across different regions and time periods, supporting the existence of long-distance exchange networks since at least the Iron Age. Finally, we document the presence of 'Pinot Noir' in Medieval pips from Valenciennes, providing direct evidence for the persistence of this iconic grape variety–now a cornerstone of the global wine industry and a major progenitor of many well-known European wine cultivars–since at least 1400–1500 CE.

Results

Genome sequencing

We assembled an extensive collection of 49 archaeological grape pips from France ($N = 47$) and Ibiza ($N = 2$), all originated from waterlogged environments exhibiting well-preserved morphology and organic material (Fig. 1; Supplementary Data 1). This material spans a broad temporal range from the Bronze Age (2300–2000 BCE, Before Common Era) to the end of the Medieval period (1400–1500 CE) (Fig. 1a). The majority of samples date to the Roman period (50 BCE-500 CE), and provide wide geographic coverage across present-day France (Fig. 1b). Two pips from Ibiza were also included to provide insights on a key trading hub of the ancient Mediterranean world[17].

Endogenous DNA content ranged from 3.72 to 66.51% (median = 13.42%), sufficient for cost-effective genome characterization through shotgun sequencing. As expected for ancient DNA, sequencing data showed characteristic patterns of post-mortem DNA damage, including an enrichment of purines at reference positions immediately adjacent to read termini (Supplementary Fig. 1a). Sequencing data also exhibited increasing C > T and G > A mis-incorporation rates toward the end of DNA fragments (Supplementary Fig. 1c; Supplementary Data 3). These signatures are consistent with post-mortem DNA fragmentation (Supplementary Fig. 1d; Supplementary Data 4) primarily driven by depurination, and cytosine deamination reactions concentrated within single stranded overhangs of DNA molecules[18]. Deamination rates were estimated between 5 and 57 times higher in single-stranded overhangs compared to double-stranded regions (Mean = 17.55; Supplementary Fig. 1e and Supplementary Data 5). Combined, these results support the authenticity of the ancient sequence data generated.

After removing PCR duplicates and low-quality alignments, the average depth-of-coverage across samples was ~1.80-fold (median = 1.60-fold, range = 0.07–10.43-fold, with 39 above 1-fold; Fig. 1d). The newly-generated sequence data were analyzed alongside a comprehensive genome panel of ancient[16,19] and modern accessions meeting minimum coverage requirements (Fig. 1e), and capturing global grapevine genetic diversity[20–27]. Data pseudo-haploidization and conditioning on transition SNPs[28] yielded a final data set of 5.79 million SNPs for downstream population genetic analyses, which included 1.05 million with allele frequencies superior to 5%.

The emergence and development of viticulture practices

Principal component analysis (PCA; Fig. 2), phylogenetic relationships and ADMIXTURE[29] ancestry profiles (Fig. 3; Supplementary Figs. 2, 3) of modern grapevine accessions reproduced the major genetic clusters previously identified. These included four distinct lineages of *Vitis vinifera* subsp. *sylvestris* (hereafter referred to as *V. sylvestris*): two predominantly found in central (Syl-W1) and western (Syl-W2) Europe, and two others common in the Levant (Syl-E1) and the Caucasus-South Caspian region (Syl-E2) (Fig. 2a). In addition to these wild lineages, six major *V. vinifera* clusters were also identified, corresponding to western Asian table grapevines (CG1), Caucasus wine grapevines (CG2), muscat grapevines from eastern Asian to western European (CG3), as well as Balkan (CG4), Iberian (CG5) and western European (CG6) wine grapevines (Fig. 2b). ADMIXTURE[29] genetic ancestry profiles, together with clustering based on phylogenetic relationships and PCA (Supplementary Figs. 3, 9; Supplementary Data 17, 21), revealed further substructure within the four *V. sylvestris* lineages and the CG1, CG3, CG4, CG5 and CG6 domesticated clusters, leading to the identification of 20 additional sublineages (Fig. 3a; Supplementary Figs. 4, 7, 10).

PCA projection onto the genetic variation observed in modern accessions revealed that the ancient grape pips spanned a broad range of genetic diversity. While some individuals clustered closely with major wild or domesticated lineages, others occupied intermediate positions, suggesting admixed ancestry (Fig. 2c).This material encompassed the Iron Age (L5, 500–475 BCE) site of Lattes (southern France), and a diversity of Roman sites from both southern (L8 Nîmes, 0–100 CE; L11 Valros, 100–200 CE and L15 Magalas, 300–400 CE) and northern France (L10 Troyes, 33–200 CE).

A total of 11 ancient samples clustered along Principal Component 1 (PC1) with the Syl-W1 and Syl-W2 wild lineages, characteristic of modern western Europe *V. sylvestris*. This group includes samples from

https://doi.org/10.1038/s41467-026-70166-z

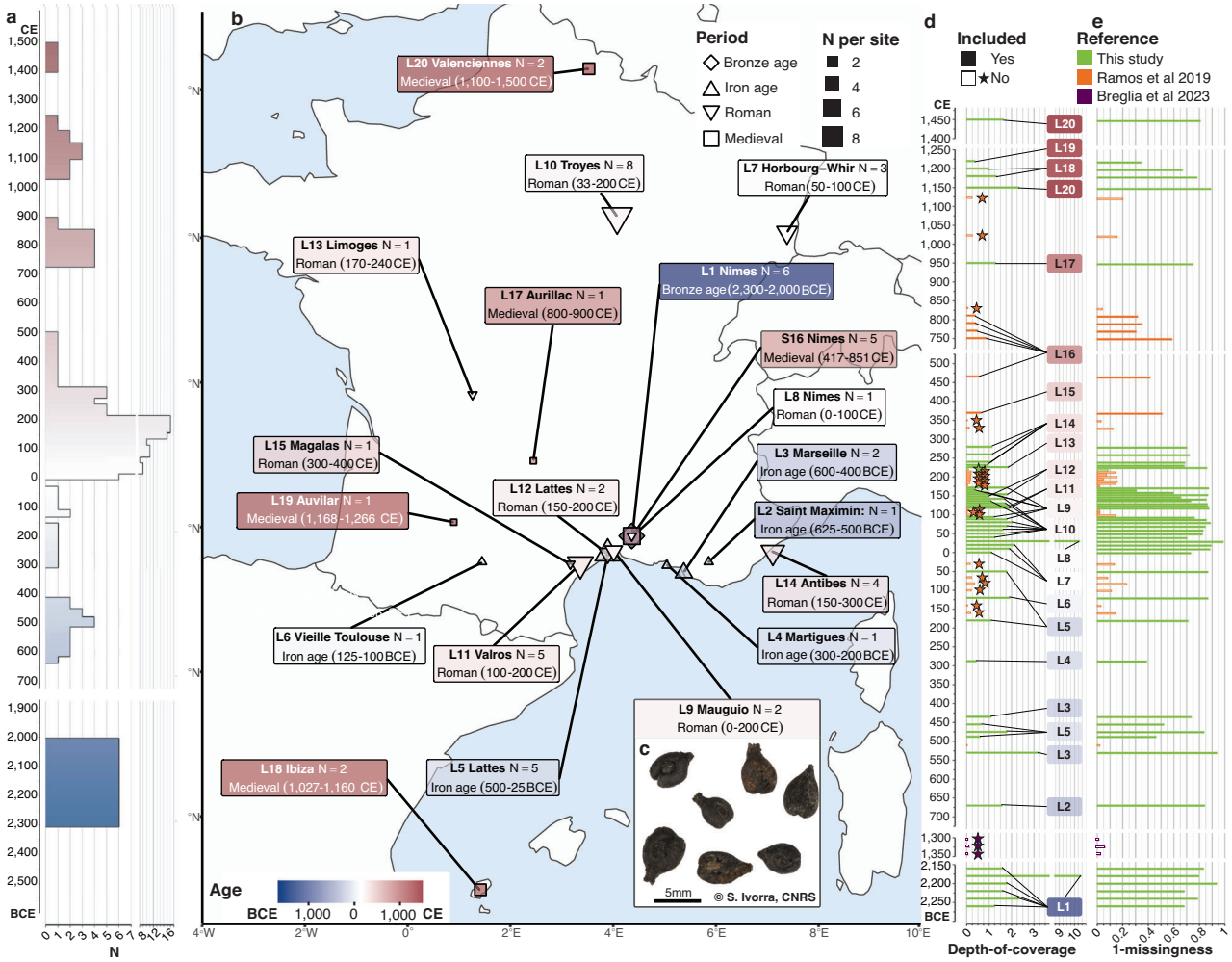

Fig. 1 | Archaeological contexts, ancient samples and genomes. a Temporal range of the 20 archaeological sites investigated. The colour gradient spans the Bronze Age (blue) to the Medieval period (red). **b** Map of the archaeological sites investigated, with colours reflecting their temporal range. Symbol shapes correspond to the Bronze Age (diamond), the Iron age (triangle), the Roman period (inverted triangle), and the Medieval period (square). Symbol sizes are proportional to the number of ancient grape pips analyzed. An unique label (LX) is assigned to each site and used for referencing throughout the text. **c** Ancient grape pips from Mauguio (S9, France). **d** Average depth-of-coverage. Samples sequenced for the first time in this study are coloured in green. One of these newly sequenced samples and most samples from Ramos-Madrigual et al.[16] and Breglia et al.[19], except six, were excluded due to insufficient coverage (i.e., < 0.1-fold). **e** Completeness. The values indicate the difference between one and proportion of missing data in the full matrix of 1,045,070 SNPs. All samples considered for downstream analyses show at most 69.64% missingness.

southern France, dating to the Bronze Age (S1 Nîmes, 2300–2000 BCE) and the Iron Age (L5 Lattes, 500–475 BCE; and L6 Vieille-Toulouse, 125–100 BCE), and from Roman northern France (L10 Troyes, 33–85 CE). ADMIXTURE[29] analysis revealed mixed ancestry profiles dominated by Syl-W2 (57.3–100.0%), sometimes also incorporating contributions from Syl-W1 (0–42.7%) (Fig. 3c). qpADM[30] modelling supported these findings, indicating Syl-W2 as the primary source mixed with Syl-W1 (Syl-W1_2 and Syl-W1_3) in samples from Bronze Age Nîmes (L1), Iron Age Lattes (L5), and Roman Troyes (L10). These results demonstrate that the genetic lineages represented by Syl-W2, today the most prevalent wild grapevine in France, were already established by 2300–2000 BCE. This highlights that western European wild grapes evolved in continuity and the absence of genetic influence from domestic grapevines for over 4000 years of viticulture. The remarkable genetic stability of these wild populations may reflect limited introgression from cultivated varieties, consistent with very low pollen-mediated gene flow from cultivated to wild populations in present-day southern France[31]. It also aligns with previous genomic work reporting that the vast majority of introgressive events occurred from wild grapevines into cultivated grapes rather than the other way around[32]. Combined, our findings depict reduced fitness for hybrids in natural forest habitats and/or intentional practices avoiding expanding cultivated forms outside of cultivated environments.

Interestingly, Iron Age sites of southern France also yielded grape pips with genetic compositions characteristic of modern domesticated varieties (Fig. 2c), with ADMIXTURE[29] indicating substantial genetic affinities with CG3, CG4, and CG5 in ancient pips from Marseille (V127, 600–400 BCE) and Martigues (V5, 300–200 BCE), as well as mixed profiles dominated by CG6 and eastern ancestries (CG3, Syl-E1 and CG1; Figs. 3b, c). The site of Lattes also yielded grape pips with genetic makeup characteristic of modern domesticated varieties (Fig. 2c), with qpADM[30] a modelling (Fig. 4e; Supplementary Fig. 16) and ADMIXTURE[29] (Figs. 3b, c) pointing to various combinations of CG3-CG6 genetic affinities (V132, 575-450 BCE; V131, 475–450 BCE; and V99, 125–25 BCE). These findings indicate that Iron Age viticultural practices in southern France involved both the integration of local wild lineages and diverse cultivated varieties. This likely results from multiple intertwined exchanges throughout the Mediterranean, since genetic affinities link to the Levant (CG1-Syl-E1), southwest Asia (CG3), the Balkans (CG4) and Iberia (CG5; Fig. 2b).

https://doi.org/10.1038/s41467-026-70166-z

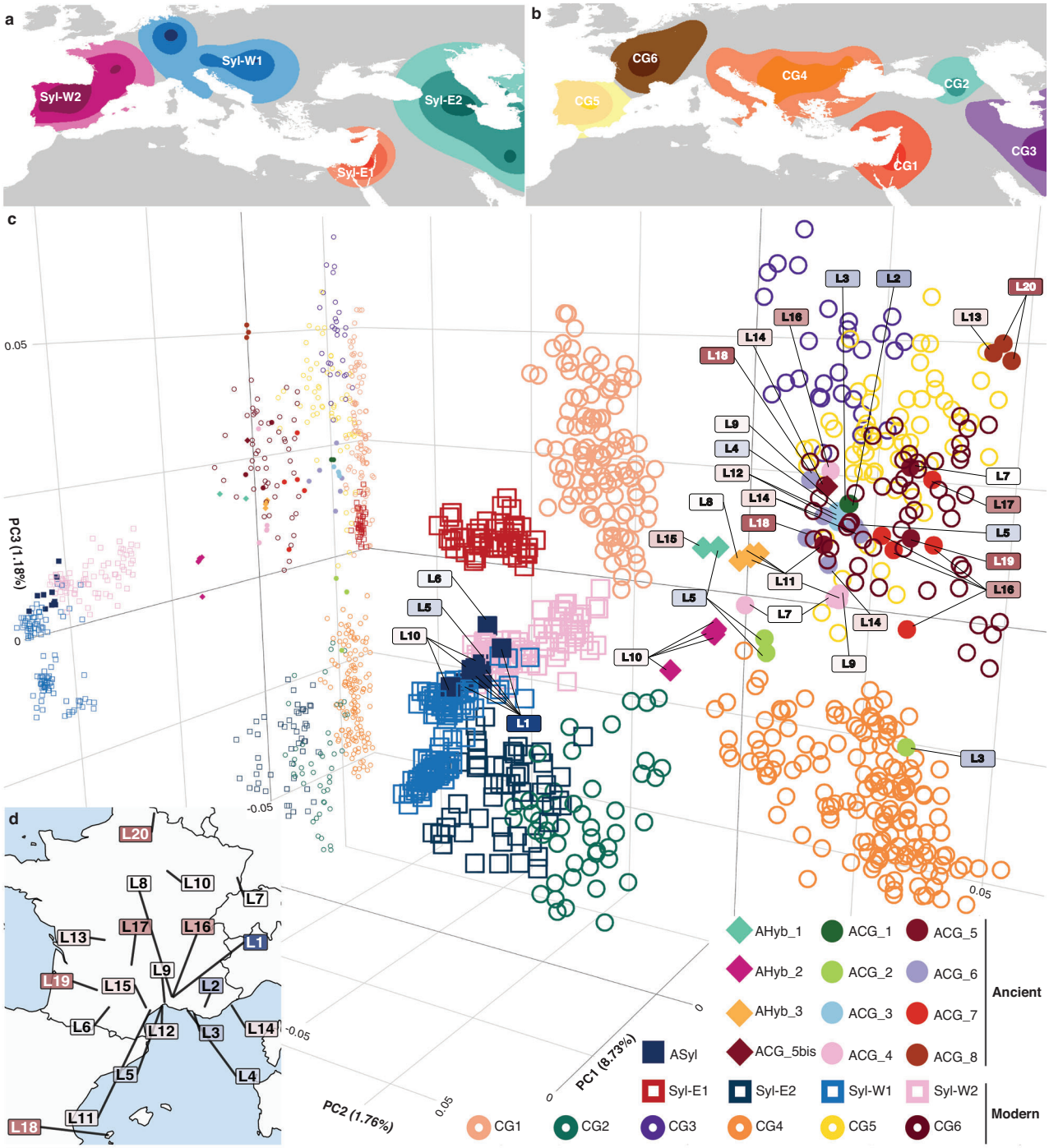

Fig. 2 | Modern population structure and Principal Component Analysis (PCA). **a** Map of the main genetic clusters of modern *Vitis sylvestris* accessions identified in Europe and Southwest Asia. **b** Same as (**a**), for modern *V. vinifera* accessions. The distribution ranges reflect the most frequent accession found in each individual area, as originally described and colour-coded by Dong and colleagues[20]. For clarity, the geographic extent of these clusters in regions where they are found in lower abundance is not shown. **c** 3D-PCA projecting ancient samples ($n = 54$) on the space defined by modern accessions of *V. vinifera* subspecies ($n = 743$). The three axes show the first three principal components, with their respective fraction of variance explained between parentheses. 2D-PCA is also shown on the plan formed by PC1 and PC3. Squares, circles and diamonds refer to *V. sylvestris*, *V. vinifera*, and their hybrids, respectively. Open and filled symbols indicate modern accessions and ancient samples, respectively. Symbols are coloured according to populations. Simplified labels point to the archaeological sites where samples were excavated, with colours reflecting their age using the same gradient as in Fig. 1a, b. **d** Summary map of the archaeological sites and their labels.

Viticultural practices were also conducive of wild admixture since one sample from Lattes (L5: V148, 500–475 BCE) exhibited a mixed genetic makeup including substantial (39.6%) wild ancestry from Syl-W2, with balanced contributions from both Syl-W2-1 and Syl-W2-2 sublineages (Figs. 3b, c). Syl-E1 ancestry, characteristic of modern Levant wild populations, was also found in small proportions in samples from Saint-Maximin (V134, 800–540 BCE), Marseille (V128, 500–400 BCE), and Lattes (V131, 475–450 BCE; V132, 200–150 BCE; and V99, 125–25 BCE) (Figs. 3b, c). This demonstrates that Iron Age

https://doi.org/10.1038/s41467-026-70166-z

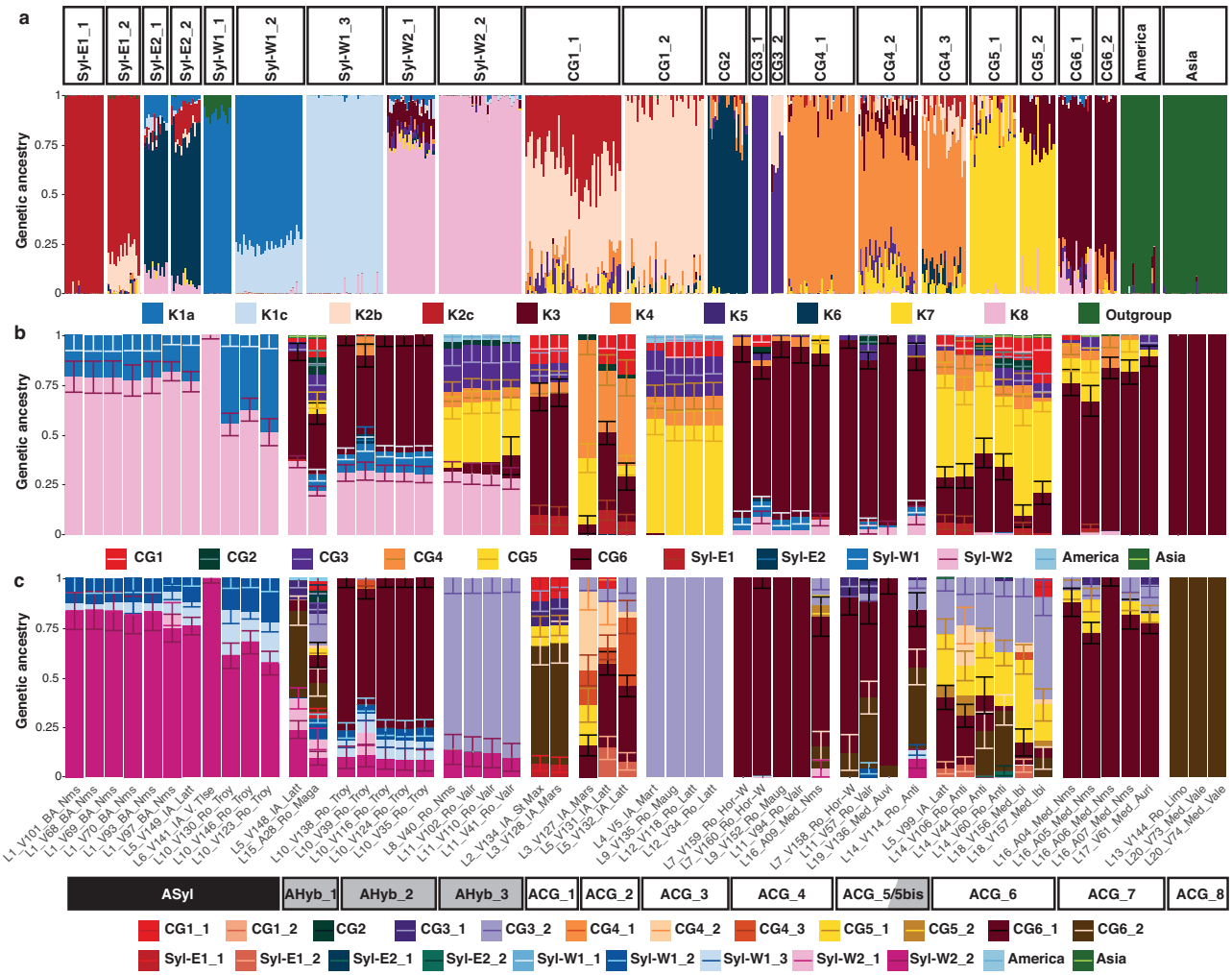

Fig. 3 | ADMIXTURE[29] genetic ancestry profiles. a Unsupervised ADMIXTURE[29] of modern accessions ($n = 559$) considering $K = 13$ ancestry components. Modern accessions are sorted according to their taxonomic status, with *Vitis sylvestris* to the left, outgroups to the right, and *V. vinifera* in the middle. Colours reflect the main genetic ancestry components first characterized by Dong and colleagues[20] with minor modifications. K1 ancestry is subdivided into K1a and K1b, which are predominantly found in the Syl-W1_1 and Syl-W1_2 subpopulations, respectively. Similarly, K2 is divided into K2b and K2c, primarily associated with the CG1 and Syl-E1 populations, respectively. Labels shown at the top indicate the subpopulation clustering considered. **b** Supervised ADMIXTURE[29] at the population level of ancient samples with 50 bootstrap pseudo-replicates ($n = 54$). **c** Same as (**b**) at sub-population level. Error bars represent the standard error of the mean (SEM) of ancestry coefficients across individuals, and bars are centred on the mean ancestry coefficient. Error bars are displayed only for ancestries with mean coefficients greater than their SEM. The labels provided within black, grey and white rectangles refer to the various ancient genetic clusters identified among wild, hybrid and domesticated accessions, respectively. The prefix 'A' is added to refer to ancient samples and avoid any confusion with those genetic clusters identified from modern accessions. Sample names reflect their original label, followed by their age category (BA = Bronze Age, IA = Iron Age, Ro = Roman, and Med = Medieval), and archaeological site.

viticulture practices of southern France involved admixture both from domesticated grapes and wild grapevines.

The Roman period, which is the most represented in our dataset, reveals a notable shift in the genetic composition of the grape pips investigated. A majority of samples from this period exhibited ancestry profiles dominated by the influence of lineages characteristic of present-day France (CG6) or Spain (CG5) (Fig. 4f; Supplementary Fig. 14c). The CG6 ancestry was found across both southern (Nîmes, 0–100 CE; Mauguio, 0–200 CE; Valros, 100–200 CE; and Antibes, 150–300 CE) and northern regions (Horbour-Whir, 0–50 CE; and Troyes, 33–200 CE). In contrast, CG5 ancestry, while present at Troyes (100–200 CE), appeared more limited in the north, accounting for only 2.4–5.9% of the genetic makeup in two samples (V35 and V124, 100–200 CE), according to qpADM[30] modelling. By comparison, CG5 contributed at least 67.7% genetic ancestry in samples from the south (Valros, V102-V110, 100–200 CE; and Antibes, V106-V44, 150–300 CE) (Fig. 4f). This pattern, although enhanced, was confirmed by ADMIXTURE[29], which showed relatively widespread CG6 ancestry, and CG5 ancestry confined to southern France (Fig. 3b; Supplementary Figs. 14, 16). CG1 ancestry, which is associated with cultivars from the Levant (Fig. 2b), was also present both in the south (Mauguio, V135, 0–200 CE; Lattes, V118–V34, 150–200 CE; Antibes, V106-V60, 150–300 CE; and Magalas, A28_R, 300–400 CE), and in the north (Horbourg-Whir, V160, 0–50 CE) (Supplementary Fig. 14c).

The Roman period also saw the continuation of viticultural practices involving the mixing of local wild grapevines with domesticated varieties, a trend already observed during the Iron Age. Substantial Syl-W1 ancestry (≥32.2%) was detected in samples from both northern and southern sites. The Syl-W2 lineage also contributed significantly in the south, with 12.6-16.9% ancestry according to qpADM[30] (Fig. 4f). ADMIXTURE[29] results indicate that Sylv-W2 ancestry was also present in northern regions, particularly in Troyes, where more than half of the detected wild introgressions could be attributed to Syl-W2 (Supplementary Figs. 14c, 15c).

https://doi.org/10.1038/s41467-026-70166-z

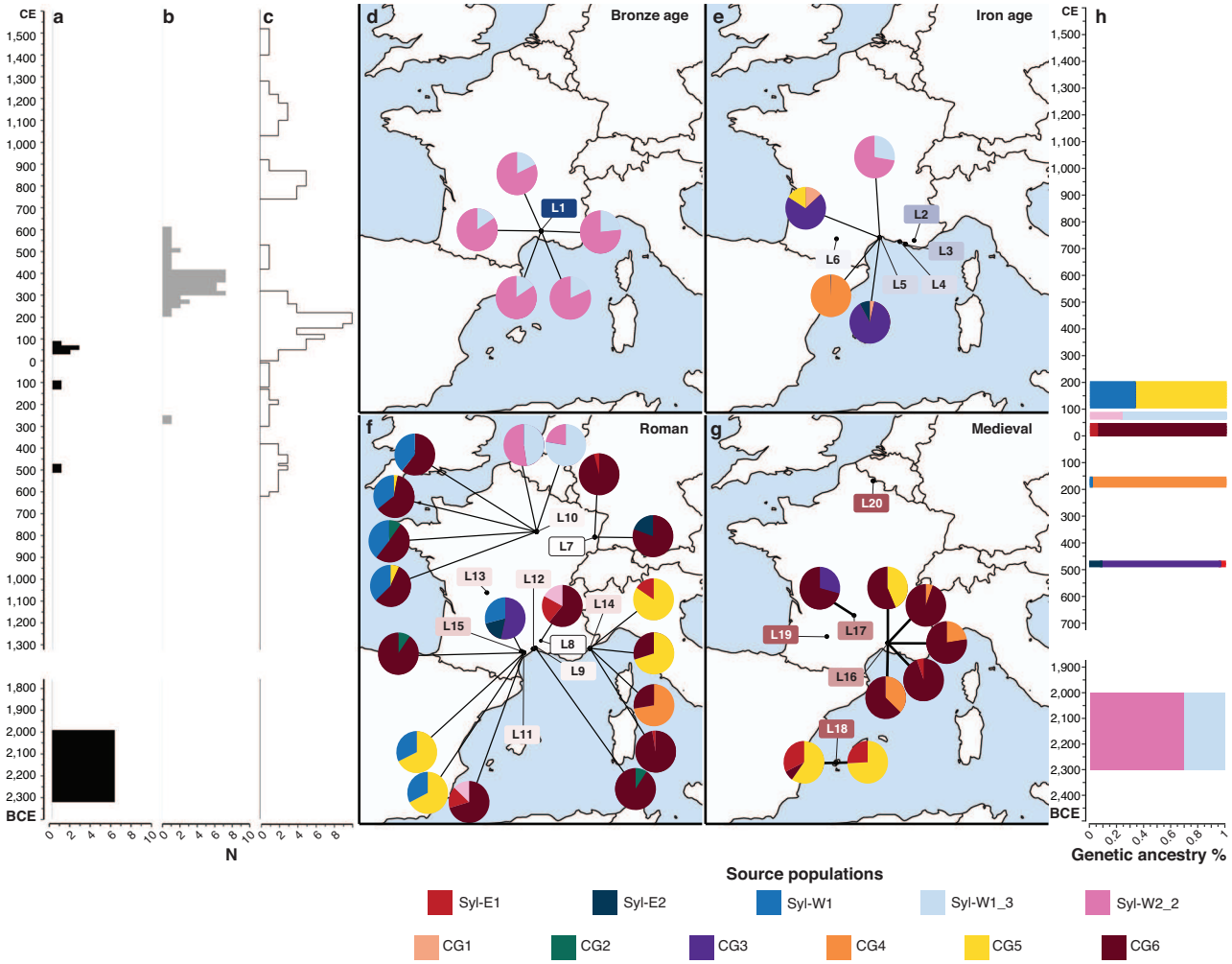

Fig. 4 | qpADM[30] modelling of ancient pips. a Temporal distribution of ancient *Vitis sylvestris* (ASyl). **b** Same as (**a**), for ancient hybrids (AHyb). **c** same as (**a**), for ancient *V. vinifera* (ACG). **d** Genetic makeup of the Bronze Age samples ($n = 5$). Models were evaluated using a χ^2-based goodness-of-fit test as implemented in qpADM[30], and only models with associated upper-tailed $p \geq 0.01$ are shown. Colours reflect the main source populations identified in models with highest p. **e** Same as d), for the Iron Age ($n = 4$). **f** Same as (**d**), for the Roman period ($n = 19$). **g** Same as (**d**), for the Medieval period ($n = 8$). **h** Earliest temporal occurrence for the diverse genetic sources identified.

Furthermore, Syl-E1 ancestry, typical of wild grapevines from the Levant, was identified in one sample from Horbourg-Whir (0–50 CE), as well as in southern samples from Nîmes (0–100 CE), Valros (100–200 CE) and Antibes (150–300 CE) (Fig. 4f, Supplementary Figs. 14c, 15c). These findings suggest the influence of non-local wild grapevines into cultivation. Similarly, Syl-E2 ancestry, characteristic of present-day wild grapevines from the Caucasus (Fig. 2a), was detected in samples from Horbourg-Whir (0–50 CE), Magalas (300–400 CE) (Fig. 4f; Supplementary Fig. 14c), Nîmes (0–100 CE),Valros (100–200 CE) and Antibes (150–300 CE) (Supplementary Fig. 14c). CG2 ancestry–a lineage most commonly found today in cultivars from the Caucasus (Fig. 2b)–was also detected in two samples from southern France (Mauguio, 0–200 CE; Valros, 100–200 CE) and in one sample from Troyes (33–65 CE), according to qpADM[30] (Fig. 4f). Combined, these findings point to continuing long-distance contacts during the Roman period, with seeds or cuttings likely transported, either directly from the Levant and the Caucasus, or indirectly through Greece or Italy, into France. These introductions appeared to have been successfully integrated into local viticultural systems through sexual reproduction, contributing to the genetic diversity of cultivated grapevines in Roman Gaul.

Importantly, the genetic composition of Medieval grapes closely resembled that observed during the Roman period (Fig. 4g). The two major domesticated lineages found in present-day France and Spain (CG6 and CG5) dominated the ancestry of all individuals that could be successfully modelled using qpADM[30]. ADMIXTURE[29] also revealed the widespread presence of CG6 ancestry in France and Ibiza, although CG5 ancestry predominated in Ibiza and southern France, aligning with alternative significant qpADM[30] models describing pips from Nîmes (A09, 417–515 CE; A04 and A07, 731–851 CE) as CG6 and CG5 mixtures (Supplementary Fig. 16). Although qpADM[30] modelling failed for all individuals from Valenciennes in northern France, ADMIXTURE analyses[29] revealed that two individuals (V73, 1100–1200 CE; and V74, 1400–1500 CE) carried ancestry mainly derived from the CG6 lineage, specifically the CG6_2 sublineage (Fig. 3b, c). At the same time, qpADM[30] models show substantial influences from non-local lineages, including wild ancestries such as Syl-E1 (5.7% in one sample from Nîmes, A07_M, 731–851 CE; and 25.8–32.3% in two samples from Ibiza, V156 and V157, 1027–1160 CE), and domesticated components such as CG3 (28.9% in one sample from Aurillac, V61, 800–900 CE), which is characteristic of the Levant and Central Asia (Fig. 4g). Similar patterns were observed in ADMIXTURE analyses[29], with CG3 ancestry detected not only in Medieval Aurillac, but also in Medieval Ibiza and Nîmes. These findings suggest that oriental influences persisted into the Medieval period, with continued integration of non-local genetic material into local grapevine populations.

Finally, samples identified as carrying pure wild ancestries displayed a distinct temporal distribution compared to those with purely domesticated or hybrid ancestries (Fig. 4a–c; two-sided Mann–Whitney U test $p < 1.19e{-3}$). Samples with exclusively wild genetic profiles were indeed present from the Bronze Age (2300–2000 BCE) to the Roman period (60–85 CE). In contrast, hybrid samples showing mixtures of wild and domesticated ancestries were first detected in the mid-Iron Age (500–475 BCE), and persisted into the Roman period (300–400 CE). Samples with genetic profiles consisting of ancestries representative of modern domesticated lineages appeared from the Iron Age (625–500 BCE) to the late Medieval period (1400–1500 CE). This temporal pattern supports the emergence of viticulture in the region between 2000 BCE and 500 BCE, consistent with archaeological evidence pointing to the introduction of grapevine cultivation in southern France during the early 6th century BCE[4,5,11]. Over time, agricultural practices increasingly favoured the propagation of grapevines with genetic composition strictly combining ancestries found in modern cultivars at the expense of wild grapevines.

Vegetative propagation

Grapes can be propagated either sexually, through seeds, or vegetatively, through cuttings that produce genetically identical clones. To investigate the emergence and extent of vegetative propagation, we searched for genetically identical or closely related samples in our data set using READv2[33], a methodology originally developed to assess relatedness in ancient human genomes[33]. We first calibrated the approach through simulations-based analyses using modern grapevine accessions with known pedigrees, establishing normalization thresholds for pairwise genetic distances that minimize false assignment rates (see "Methods"). The calibrated thresholds produced a very low false positive rate, with only 0.4% of unrelated individuals misclassified as related up to the second degree (Supplementary Fig. 17a). They also reliably identified identical domesticated clones and identical wild clones in respectively 99.3% and 94.60% of cases (Supplementary Figs. 17e, 18d). No unrelated domesticated individuals, nor closely related individuals, were identified as identical, confirming the robustness of our approach to identify grapevines that are maintained through vegetative propagation (Supplementary Figs. 17a–d). The method also showed good sensitivity for close kin, correctly identifying 86.9% of first-degree relatives as related up to the third degree (Supplementary Fig. 17d). In contrast, accuracy decreased for identifying second- and third-degree relatives, with 69.8% and 93.0% of such pairs incorrectly classified as unrelated, respectively (Supplementary Figs. 17b, c). To maintain high confidence in our inferences, we therefore restricted our analyses to clone detection and first- or second-degree relatives only, excluding more distant assignments. Due to the overlap in misclassification rates between first- and second-degree relatives (48.8% and 1.7%, respectively), we also grouped these two categories into a single class of closely related individuals.

We first screened for pairs of clones among the ancient grape pips. Strikingly, a substantial fraction of the wild Bronze Age grape pips from Nîmes were genetically identical (Fig. 5a). This finding is surprising, as vegetative propagation is limited in wild grapevines. Given that these pips were recovered in close spatial proximity, in a palaeochannel where seeds may have been deposited naturally through water or gravity, we interpret them as likely deriving from a single grapevine.

More remarkably, we identified genetically identical clones across distinct archaeological sites dating to roughly similar time periods, suggesting the use of vegetative propagation and the intentional movement of cuttings or, less likely, pips–possibly through commercial exchanges, such as the trade of fruits or preparations in amphorae (Fig. 5b). For instance, clonal samples were found in Iron Age contexts at Saint-Maximin (625-500 BCE) and Marseille (V128, 500–400 BCE), two sites located in southern France and separated by ~40 km. In the 6th c. BCE, Saint Maximin presented all the characteristics of local indigenous settlements found in the region. It especially stood out from those sites found on the outskirts of Marseille, a large neighbouring city with intense seaborne connection with the Greek world. The genetic proximity documented may therefore represent the earliest direct evidence of vegetative material and viticultural knowledge exchanged between coastal trade hubs and rural inland sites. Another case, from the Roman period, involves a pip from Horbourg-Whir (V159, 0–50 CE) in northeastern France, which was genetically identical to two others pips from the South of France (Mauguio, V152, 0–200 CE; and Valros, V94, 100–200 CE), located 560 and 610 km away, respectively. Evidence of long-term vegetative reproduction was also detected across both time and space. The identification of identical clones in two sites separated by a time gap necessarily implies the practice of vegetative propagation, as maintaining self-pollination in hermaphrodite cultivated plants through centuries appears especially difficult. In southern France, identical samples were recovered from Iron Age Martigues (V5, 300–200 BCE) and Roman-period Mauguio (V135, 0–200 CE) and Lattes (V118–V34, 150–200 CE). In central and northern France, clonal continuity was observed over a thousand years between Roman-era Limoges (V144, 170–240 CE) and Medieval Valenciennes (V73, 1100–1200 CE), two sites located approximately 530 kilometres apart. Combined, these findings indicate that vegetative reproduction was a widespread and consistent feature of grapevine cultivation across the entire region investigated from at least the Iron Age (625–500 BCE), with some grapevine clones persisting over one millennium. This enduring clonal continuity echoes historical and archaeological evidence showing the prevalence of layering techniques for propagating grapevines in southern Gaul and the wider Roman world[34,35].

Wild grape pips exhibiting first/second degree genetic relatedness were only detected within the same archaeological sites (Fig. 6a). In contrast, such close genetic relationships were identified across both space and time among grape pips showing a domesticated genetic makeup (Fig. 6b). In the southern regions, these relationships occurred over short distances (< 100 km) as well as long distances (> 700 km), extending beyond present-day France–for example, between the French Mediterranean coast and Ibiza (Spain). Notably, close genetic relationships also reached further north, connecting Roman Horbourg-Whir to Medieval Valenciennes. The chronological context of each site provided a unique opportunity to trace the cultivation history of individual grapevines across the country. For instance, sample V5 from Iron Age Martigues (300–200 BCE) showed close genetic ties to three samples from the Roman site of Antibes (V60, V106, and V156, 150–300 CE), and two samples from Medieval Ibiza (V156 and V157, 1027–1160 CE). Given that such genetic proximity indicates first- or second-degree relatedness, this pattern suggests that these grapevines underwent no more than four cycles of sexual reproduction over the course of more than a millennium. This aligns with the picture emerging from genetically identical clones as well as archaeological and historical evidence for organized vineyard management, and the wide circulation of cultivars across the western Mediterranean at the time[12,13]. Combined, these findings support that vegetative propagation was a common viticulture practice during Roman and Medieval periods, highlighting the role of extensive exchange networks in disseminating grapevine cuttings through space and time.

Given the long-term maintenance of certain cultivars through vegetative propagation, we next examined our dataset for instances of clonal matches or close genetic relatives between ancient samples and modern accessions (Fig. 6c). Strikingly, two Medieval samples were found to be genetically identical to modern cultivars. Sample V156 from Medieval Ibiza (1027–1160 CE) matched with the 'Folha de Figueira' cultivar, a white grape variety still grown in Portugal today. Sample V74 from Medieval Valenciennes (1400–1500 CE) was also

https://doi.org/10.1038/s41467-026-70166-z

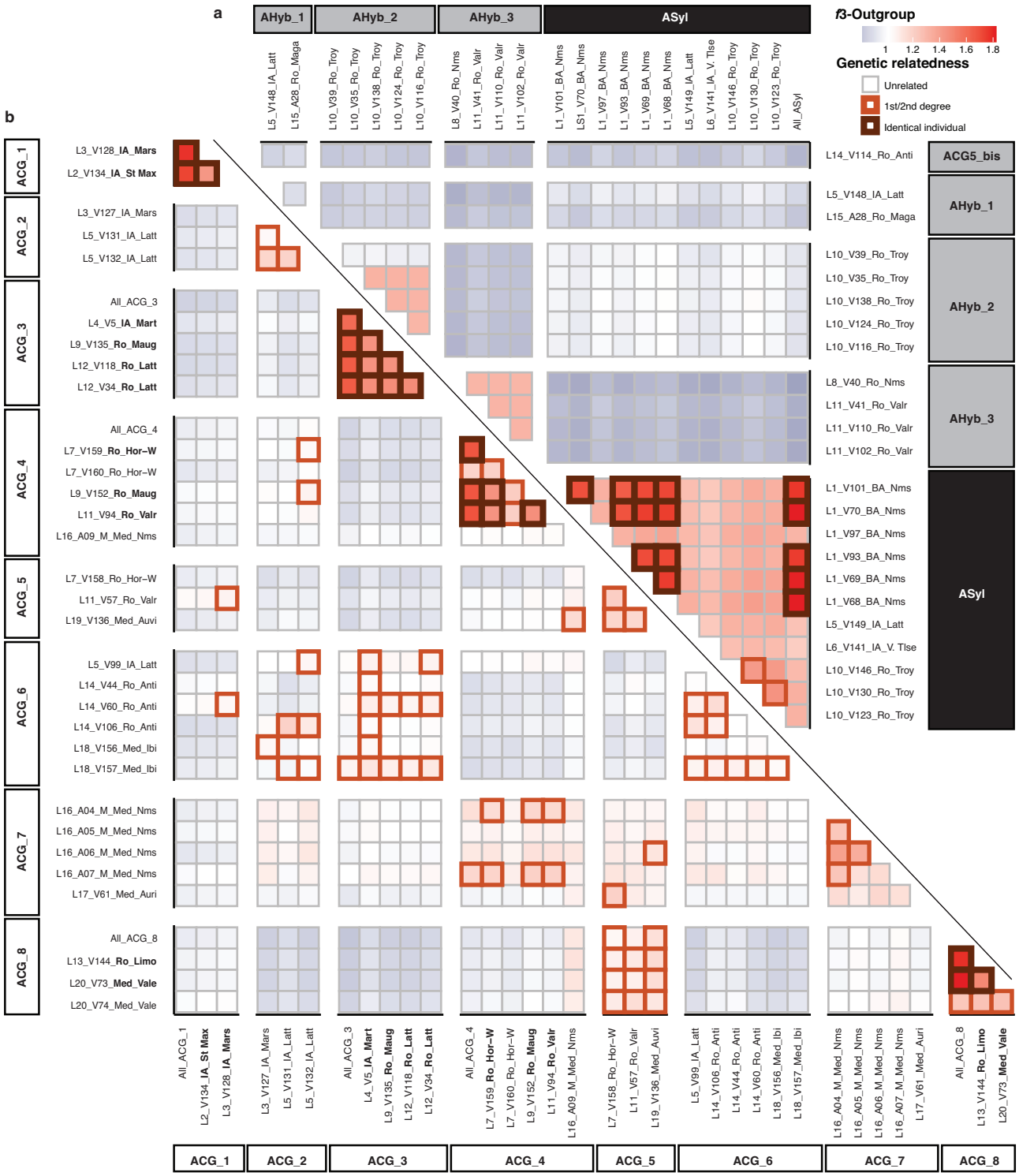

Fig. 5 | *f*3-outgroup statistics[63] and genetic relatedness between pairs of ancient samples. a Ancient *Vitis sylvestris* and/or ancient hybrids. Samples are grouped according to their genetic clusters. *f*3-ougroup statistics, using non-*V. vinifera* species as outgroups, measure the extent of shared genetic drift between pairs of samples. The colour gradient reflects lower (blue) to higher (red) shared genetic affinities. Squares of wider stroke indicate closer genetic relatedness up to the second degree. **b** Same as (**a**), for ancient *V. vinifera*. The labels provided within black, grey and white rectangles refer to the various genetic clusters identified among wild, hybrid and domesticated accessions, respectively. The prefix 'A' is added to refer to ancient samples and avoid any confusion with those genetic clusters identified from modern accessions. Groups identified with a "hyb_" prefix correspond to the data merger from several samples identified as genetically identical.

identical to the renowned red grape variety 'Pinot Noir'–one of the oldest known cultivated grapevines, believed to have originated in Burgundy, northern France. Our results indicate that 'Pinot Noir', a grape variety critical for the global wine industry at the origin of some of the world's most prestigious wines, was genetically already fully established no later than the end of the 15th century CE. This variety also had a massive influence on modern viticulture, with many cultivars representing first- or second-degree relatives (Fig. 6c).

https://doi.org/10.1038/s41467-026-70166-z

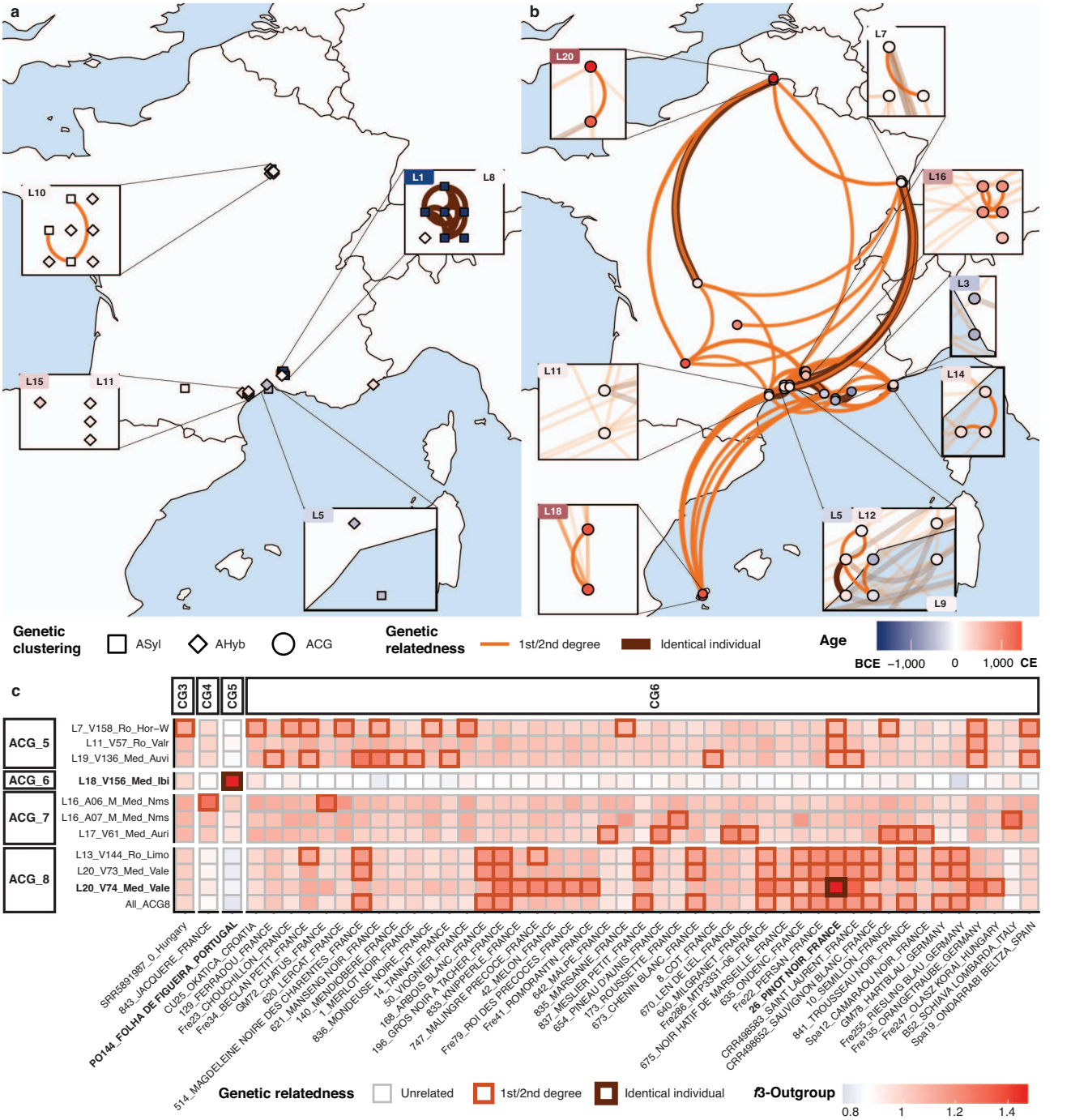

Fig. 6 | Genetic relatedness. a Amongst ancient *Vitis sylvestris* and hybrid samples. Contemporary archaeological sites are zoomed in, whether or not they comprised samples genetically related up to the 2nd degree. Transparent lines within the zoomed-in areas indicate the genetic relatedness between the samples from the highlighted sites and those from other locations. **b** Same as (**a**), for ancient *V. vinifera* samples. The colour gradient spans the Bronze Age (blue) to the Medieval period (red), following Fig. 1a, b, and site labels follow Fig. 2d. **c** Between pairs of ancient samples (vertical axis), and modern accessions (horizontal axis). Samples are grouped according to their genetic clusters. *f*3-outgroup statistics, using non-*V. vinifera* taxa as outgroups, measure the extent of shared genetic drift between pairs of samples. The colour gradient reflects lower (blue) to higher (red) shared genetic affinities. The figure only shows those pairs showing at least one genetic relatedness up to the 2nd degree. Lines with a wider stroke indicate closer genetic relatedness.

Discussion

This study substantially builds upon previous work exploring the genetic landscape of ancient grapevines in France between 510–475 BCE and 1100–1300 CE[16]. First, we relied on whole-genome shotgun sequencing instead of in solution hybridization capture for 10,000 SNPs. This approach enabled a comprehensive characterization of genetic variation across the entire genome, enhancing resolution for population structure, relatedness, and admixture analyses. Second, we more than doubled the number of ancient grape pip genomes analyzed, substantially increasing both the spatial and temporal coverage of grapevine cultivation in the region. Notably, our dataset extends the chronological range back to the Bronze Age (2300–2000 BCE) and forward into the Late Medieval period (1300–1500 CE). In addition, half of the specimens derive from Roman era contexts, offering an unprecedented window into grapevine diversity and cultivation practices during this formative period. All of the analyzed pips were

preserved in waterlogged environments and displayed excellent biomolecular preservation, with moderate to high endogenous DNA content. This pattern is consistent with earlier findings from ancient wood preserved under similar conditions, which also yielded DNA suitable for cost-effective whole-genome sequencing[36,37].

France provides an ideal setting for exploring ancient grape genetic relationships to wild populations and modern varieties, and trace their spatio-temporal diffusion. Its geographical position is key to examine the long-term development of viticulture within the broader context of Mediterranean and European interactions. Our ancient grape pip assemblages encompassed local wild lineages, particularly Syl-W2, dominant in France since at least 2300–2000 BCE. From the beginnings of viticulture, they also included diverse domesticated ancestries, such as CG5 and CG6 (modern Spanish and western European cultivars), and CG3 and CG4 (associated with eastern Asia and the Balkans). The earliest genome exclusively carrying ancestries from domesticated grapevines was identified in the Iron Age site of Saint Maximin (625–500 BCE), providing a minimum *ante quem* date for the emergence of viticulture in the region in concordance with the generally admitted arrival of viticulture in France when Greek settlers founded the city of Marseille at the beginnings of the sixth century BCE[8,38]. No genetic trace of domesticated cultivars was detected in earlier periods, which aligns with archaeobotanical evidence reporting wild, not cultivated[11], grapevines in western Europe since the Early Holocene (11,000-10,000 BCE)[10,39]. Our finding challenges the hypothesis from Dong et al.[20], which posits an earlier diffusion of viticulture into western Europe during the Neolithic.

The presence of CG5-related ancestry in some Iron Age samples from France raises questions about the actual origin of this genetic group. This group is characteristic of present-day Iberia and is generally considered to have originated locally, as viticulture is attested there a few centuries earlier than in Southern France[40]. Extending genetic mapping of ancient grape pips across the full Mediterranean range may, however, reveal alternative interpretations, with plant material originally diversified elsewhere and later introduced into Iberia.

Iron Age assemblages, such as those from Lattes (500–25 BCE), show the coexistence of local wild grapevines and introduced domestic lineages, consistent with the presence of wild morphotypes in most Iron Age and Roman sites from France, including at Lattes[11]. Our work also provides evidence of admixture between local wild grapes and domesticated forms as early as the Iron Age (e.g., V148, 500–450 BCE), and into the Roman period, especially in the North (Fig. 4b). Such examples may reflect deliberate efforts to develop new cultivars better adapted to local environmental conditions through adaptive introgression, allowing for the expansion of viticulture from the Mediterranean into the broader Gallic territory at the time of growing demands for viticulture products[8,11,41]. This interpretation aligns with historical accounts by Dion[42] and Roman writings of the time[43] as well as the model proposed by Bonhomme et al.[44] and the continuous introgression from wild to cultivated grapes reported by Xiao et al.[32].

Patterns of genetic affinities during the Iron Age and Roman period suggest connections with the Balkans, the Iberian Peninsula, and western Asia, indicating long-distance contact. This aligns with the study from Ramos-Madrigual et al.[16] who identified genetic affinities between an early Roman pip from Nîmes (France) and modern cultivars from the Balkans and western Asia. Archaeobotanical data further support external influence, particularly during the early stages of viticulture in southern Gaul, when Greek and Etruscan merchants engaged with local indigenous populations[10,40]. Combined, archaeobotanical and palaeogenomic evidence suggest that the genetic landscape of ancient viticulture in France was shaped by dynamic cultivation practices and exchange networks.

The detection of identical clones and closely related individuals across sites and centuries highlights the early and widespread adoption of vegetative propagation. From the Iron Age onward, certain grapevine lineages were maintained over centuries and more, in some cases linking Iron Age and Medieval contexts through only a few cycles of sexual reproduction. The identification of two Medieval samples genetically identical to modern cultivars–'Folha de Figueira' and 'Pinot Noir'–adds to the formerly reported match between one ~1100 CE landrace from Orléans and modern 'Savagnin Blanc'[16], a parent of 'Pinot Noir'[45,46]. Such links may extend even further back in time as two Roman pips were previously found highly related to 'Pinot Noir'[16]. Combined, these results demonstrate not only the longevity of specific clonal lines but also their foundational role in the development of modern viticulture. In particular, 'Pinot Noir' is the fourth most widely cultivated variety worldwide, with over 105,480 hectares recorded in 2016[47]. It represents the flagship cultivar of Burgundy, and is known for giving rise to some of the finest red wines since the end of the Middle Ages[48] and to many modern cultivars, which emerged as first to third degree relatives[38]. These findings point to viticultural traditions heavily relying on clonal reproduction to ensure inheritance of desirable traits across generations and geographies.

The genome collection characterized in this study provides a valuable foundation for investigating the origins and historical trajectories of grapevine lineages that are essential to the global wine industry. The genetic ancestries found in the ancient grape pips investigated show complete overlap with those of modern European and Southwest Asian accessions. It indicates some level of genetic continuity over time, including until the present-day, but also suggests key periods of transition in cultivation practices. Future work must, however, expand the temporal breadth of the assemblages investigated, especially for pre- and post-Roman contexts, to gain fine-grained resolution into dynamics of cultivation practices and diversity changes of grapevines over time. Strikingly, the observed overlap between modern and ancient accessions also offer favourable conditions for the application of statistical imputation techniques, which leverage dense modern genomic reference panels[20–22] to infer genotypes and haplotypes from low-coverage ancient sequence data[49]. Such approaches are particularly promising for increasing resolution into the changing population structure of grapevines in relation to major cultural, ecological or social changes. Moreover, imputation may also enable the prediction of phenotypic traits, including for quantitative traits, such as berry colours[50], shape, chemical composition, and flavour[51], thereby opening avenues for exploring the evolution of agronomic and sensory characteristics in ancient viticulture. Additionally, targeting genes associated with phenological traits–such as flowering time or ripening–may provide insights into past adaptations to local climatic conditions[37] and shifts in cultivation practices over time.

Methods

Ethics and samples

All samples were collected and analyzed under the appropriate permits issued by the relevant authorities in the regions covered by this study. Each sample included in the analyses is associated with an archaeologist who is a co-author of the article and who was granted official authorization to excavate and analyze archaeological material, issued in France by the Ministry of Culture and in Ibiza by the Consell Insular d'Eivissa, and who can attest to the provenance of the material. We confirm that this research was conducted in compliance with all ethical standards applicable to archaeology and archaeogenetics, as reviewed and approved by the scientific committees of the MICA and Viniculture ANR (Agence Nationale de la Recherche) projects.

Ancient DNA extraction

Samples were processed in the state-of-the-art ancient DNA facilities at the CAGT, UMR5288 CNRS (Toulouse, France). The procedure for extracting ancient DNA followed the methodology from Wales et al.[52],

with minor modifications. Individual pips were first crushed with a hammer before being incubated at 37 °C for 20 hours in a 1.5 mL digestion buffer consisting of: 10 mM Tris-HCl (pH 8.0), 10 mM NaCl, 2% (w/v) sodium dodecyl sulphate (SDS), 5 mM CaCl2, 2.5 mM EDTA (pH 8.0), 40 mM dithiothreitol (DTT), and 0.5 mg/mL proteinase K. The full volume used for digestion was purified through two cycles of mixing and centrifugation (15,871 x *g*), with equal volumes of phenol-chloroform-isoamyl-alcohol, and one cycle of mixing and centrifugation with an equal volume of chloroform. The hydrophylic fraction collected was further mixed with ten volumes of PB buffer and purified using the MinElute PCR purification kit (Qiagen). This consisted of a first centrifugation step for 1 min at 6010 x *g*, followed by a washing step in which the column was washed with 600 µL of PE buffer and centrifuged twice for 1 min at 6010 x *g* to remove any remaining trace of alcohol. Ancient DNA was finally eluted adding 60 µL of elution buffer supplemented with Tween 0.5% to the columns before incubation for 10 min at 37 °C, and final centrifugation for 2 min at 15,871 x *g*. Extraction blanks were included every 8 samples to monitor contamination.

Library preparation and sequencing

Double-stranded blunt-end DNA libraries were prepared following the protocol from Rohland et al.[53]. DNA libraries were PCR amplified during 12 cycles in a total reaction volume of 25 µL, using AmpliTaq Gold reaction mix and 1U of AmpliTaq Gold DNA Polymerase Enzyme, together with 3 µL of DNA library, 0.8 mg/mL of BSA, 0.2 mM of inPE1 primer, and 0.2 mM of indexing primer containing an unique 6-bp "external" barcode. Amplified DNA libraries were quantified on a TapeStation 4200 instrument using a High sensitivity D1000 ScreenTapeAssay (Agilent technologies), and Qubit HS dsDNA assay (Invitrogren). Each DNA library was diluted to between 1 nM and 10 nM for shallow sequencing on Illumina MiniSeq instrument (Paired-end mode, 2×81 cycles) at CAGT. Those DNA libraries showing sufficient DNA preservation were sequenced more extensively on the Illumina HiSeq4000 and/or the NovaSeq6000 (S2 and S4) at Illumina platform Genoscope, Evry (Paired-end mode, 2 × 150 cycles) (Supplementary Data 2).

Sequence data processing, mapping and filtering

Raw sequencing data were processed using the Paleomix pipeline (v1.2.13)[54] for adapter trimming, read demultiplexing and collapsing, mapping against the grape reference nuclear genome (12xV2)[55], alignment quality (≥ 25) and PCR duplicate filtering. Alignment was carried out using Bowtie2 (v2.3.4.1)[56], with the sensitive option and end-to-end mode as recommended by Poullet and Orlando[57]. Final average depth-of-coverage was estimated using paleomix coverage –ignore-readgroups. The presence of signatures of post-mortem DNA damage was assessed using mapDamage2 v2.0.9-2-gd251da0[58], with default parameters and downsampling sequence data to a random set of 100,000 high-quality unique alignments (Supplementary Fig. 1; Supplementary Datas 3-5).

The same processing, mapping and filtering procedure was repeated on previously published sequence data, including from a total of $N = 80$ ancient grapes[16,19,59], and $N = 991$ modern grape accessions. The latter included $N = 908$ genomes of *V. vinifera* subspecies divided into 10 populations according to Dong and colleagues[20], selected to encompass the whole range of genetic ancestries present in modern populations[20–22], as well as $N = 80$ genomes of 48 other *Vitis* species, and 3 genomes of other Vitaceae species, that were used as outgroups[23–27] (Supplementary Datas 6–8).

To accommodate uneven sequencing depths bias between ancient and modern data, all genomes were pseudo-haploidized by randomly sampling one allele per site at those sites covered in at least 50% of the individuals. Only transversion SNPs were retained to limit the impact of post-mortem DNA damage on downstream analyses, leaving a total of $N = 5{,}785{,}152$ sites that were further filtered using PLINK (v1.90)[60] to prepare two complementary data sets. The first data set (hereafter referred to as the "core data set") retained a total of $N = 932{,}511$ unlinked SNPs (r2 < 0.5 in 50-SNP sliding windows, step=5) showing a minimum allele frequency of 5% (MAF ≥ 0.05). The core data set was used for PCA and ADMIXTURE[29] analysis. The second data set (hereafter referred to as the "extended data set") retained all SNPs showing MAF ≥ 0.05, including $N = 1{,}045{,}070$ SNPs (Supplementary Data 9). This second data set was applied for phylogenetic reconstruction, ancestry modelling and genetic relatedness analysis.

To improve the definition of genetic clusters, crucial for supervised analysis, both data sets were declined into two data sets representing two levels of resolution. First, modern accessions represented by more than 50% missingness, and ancient accessions showing over 75% missing data were removed, leaving in both data sets a total of $N = 54$ ancient grapes and $N = 983$ modern individuals. Then, only modern accessions with at least 50% of their ancestry similar to the main ancestry components of their population were retained, leaving a total of $N = 796$ modern individuals. These individuals were used for analyses defining the population level, retaining respectively 913,705 and 1,045,990 SNPs for the core and extended data sets (Supplementary Data 9). Finally, we applied k-means clustering to the genetic data of the 796 modern individuals selected above in order to assess potential substructure within populations. Clustering was performed on seven different input datasets, each representing a distinct combination of genetic information: (1) PCA coefficients alone, (2) genetic ancestries from ADMIXTURE[29] alone, (3) genetic distances alone, (4) PCA coefficients combined with genetic ancestries, (5) PCA coefficients combined with genetic distances, (6) genetic ancestries combined with genetic distances, and (7) the combination of all three (PCA, ADMIXTURE[29], and genetic distances). For each population, the optimal number of genetic clusters was determined using the consensus-based algorithm implemented in the *NbClust* R package (v3.0.1)[61] (Supplementary Data 10). Only individuals consistently co-grouped in 5 of the 7 clustering approaches were kept, and 20% of the accessions with the highest mean genetic distance were removed in each subgroup (Supplementary Data 11). These thresholds were chosen to maximise genetic diversity and sub-population identities, leaving a final set of $N = 559$ modern individuals from 21 different subpopulations. These individuals were used for analyses carried out at the sub-population level, including a total of 881,071 and 1,056,315 SNPs for the core and extended data sets, respectively (Supplementary Data 9).

Phylogenetic reconstruction and population demographic structure

Neighbour-joining trees were reconstructed considering either the whole data set (Supplementary Fig. 2a), or the populational (Supplementary Fig. 3a) and sub-populational levels (Supplementary Fig. 4a). Trees were obtained using FastME (v2.1.6.2)[62], and a total of 100 bootstrap pseudo-replicates to assess node robustness. Pairwise f3-outgroup statistics[63] were calculated using Admixtools (v7.0.2)[63] with the inbreed option, considering *non-V. vinifera* taxa as outgroups (Supplementary Fig. 5; Supplementary Data 12). PCA analyses were carried out using SMARTPCA (v1600)[64], turning on the autoshrink option to take into account the various levels of missingness between ancient and modern genetic data. Eigenvectors were first inferred from the modern accessions, and before projecting ancient samples. PCA was calculated considering either the whole data set, or the population and subpopulation levels, including or excluding *non-V. vinifera* subsp. (Fig. 2c; Supplementary Figs. 6-10; Supplementary Datas 13–18). PCA analyses were also repeated restricting genomic positions to those from Ramos-Madrigal et al.[16] found in the core data set ($N = 4363$; Supplementary Fig. 19; Supplementary Data 19), in order to illustrate the gain in resolution resulting from whole-genome data.

Unsupervised ADMIXTURE (v1.3.0)[29] genetic ancestry profiles were obtained from K = 2 to K = 15 genetic ancestries considering either the whole data set (Supplementary Fig. 11a), or the populational (Supplementary Fig. 12a) and sub-populational levels (Supplementary Fig. 13a). Suboptimal values of K were selected based on the minimal cross-validation (CV) error, and the ability to estimate ancient genetic ancestries solely from modern components: K = 13 for the whole data set (Supplementary Fig. 11b; Supplementary Data 20), K = 13 at the populational level (Supplementary Fig. 12b; Supplementary Data 21), and K = 11 at the sub-populational level (Supplementary Fig. 13b; Supplementary Data 22). Ancestry standard errors were estimated from $N = 50$ bootstrap pseudo-replicates, using the --haploid = "*" flag. Supervised analyses were carried out to describe ancient samples at the populational (Fig. 3b; Supplementary Fig. 14; Supplementary Data 23) and sub-populational levels (Fig. 3c; Supplementary Fig. 15; Supplementary Data 24).

The genetic makeup of ancient individuals was modelled using qpADM (v1520)[30], considering a single, two, three, and four possible sources. The modelling procedure follows the "rotating" scheme from Harney et al.[30], with sources representing either the two *V. vinifera* subsp., the 10 populations from Dong et al.[20], or those newly inferred 21 sub-populations. Non-tested source populations, as well as non-*V. vinifera* subspecies were included as right populations (i.e., outgroups). Models showing $p \geq 0.01$ were considered to show statistically significant data fit, although the most parsimonious models (i.e., minimizing the number of contributing sources) associated with the largest *P* value at the populational and/or sub-populational levels were considered as the best (Fig. 4d–g; Supplementary Fig. 16; Supplementary Data 25). All significant models were also re-inferred with the SNPs from Ramos-Madrigal et al.[16] present in the extended data set ($N = 4363$; Supplementary Fig. 20; Supplementary Data 25), to compare *p* and standard deviations.

Genetic relatedness

Genetic relatedness between pairs of ancient samples, and between ancient samples and modern accessions, were assessed using READv2 (v2.00)[33], a method specifically designed to estimate relatedness from pseudo-haploid genomic data. READv2 divides the genome into windows of 1 Mb to calculate the distribution of genetic distances (P0) between all pairs of individuals. By default, a normalization step is done considering the median P0 value observed across the full data set. The normalized P0 is then used to infer genetic relatedness between individuals up to the 3rd degree, leaving all other pairs of individuals as unrelated. This procedure was, however, originally developed for human data and may, thus, not immediately apply to species with high heterozygosity and mixed sexual and clonal reproduction systems, such as grapevine.

To define an appropriate normalization value for *V. vinifera* accessions, we used $n = 11{,}758$ known parent-offspring relationships from the *Vitis* International Varieties Catalogue (VIVC)[65] to reconstruct relatedness up to the 3rd degree among cultivars from the extended data set. As no identical relationships were identified, we generated a second extended data set using the pseudo-haploidization procedure described above, which was merged with the original. This resulted in each accession being present twice, with minor differences between duplicates due to the stochastic nature of the pseudo-haploidization process. We then assessed misassignment rates for each relatedness level using our known-relatedness relationship across a range of normalization thresholds, comparing them to the default median P0 value used by READv2[33]. The calibrated normalization value that minimized the false assignment rate across all relatedness levels was selected (P0 = 0.238; Supplementary Fig. 17; Supplementary Data 26), and applied to samples with fully domesticated ancestries (ACG) (Supplementary Data 27). We also measured how using the restrained SNPs data set from Ramos-Madrigal et al.[16] impacted the misassignment rate for each relatedness level (Supplementary Fig. 21; Supplementary Data 28).

Due to differences in genetic diversity and reproduction systems within *V. vinifera* subsp. (Supplementary Fig. 18a), a second normalization threshold was defined for *V. sylvestris* individuals. In the absence of large, well-characterized wild pedigrees, we used the relative position of the selected normalization value for *V. vinifera* along the misassignment curve of artificially duplicated identical *V. vinifera* accessions misidentified as unrelated (Supplementary Fig. 18c; Supplementary Data 29). This relative position was then projected onto the misassignment curve of artificially duplicated Western European *V. sylvestris* individuals (Supplementary Fig. 18d). The final normalization threshold for ancient samples with purely wild (ASyl), or significant wild (AHyb) ancestry, was defined as the mode of the resulting distribution of projected normalization values, with P0 = 0.138 (Supplementary Fig. 18e). Eastern wild individuals were excluded from this process due to their higher genetic diversity (Supplementary Fig. 18b), which would have artificially inflated the normalization threshold and increased the false positive rate in Western wild-like samples, which represent the dominant type present in our ancient data set.

Reporting summary

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

Data availability

Demultiplexed sequencing data and alignments against the 12x.2 grape reference genome generated in this study are deposited in the European Nucleotide Archive (ENA) under the following accession code PRJEB94459. Previously published sequencing datasets reanalyzed in this study are available under the following accession codes: PRJNA489970[16], PRJNA999173[19], PRJCA009314[20], PRJNA393611[21], PRJNA388292[22], PRJNA485199[23], PRJNA490319[24], PRJNA868429[25], PRJNA731597[26].

All other data supporting the findings of this study are provided in the Supplementary Data files. As the original archaeological grape pips were consumed during destructive sampling, only ancient DNA extracts and sequencing libraries are retained. These materials are curated at the Centre for Anthropobiology and Genomics of Toulouse (CAGT, Toulouse, France). Access to the material can be requested from Laurent Bouby (laurent.bouby@umontpellier.fr) and Ludovic Orlando (ludovic.orlando@utoulouse.fr).

Code availability

No original code has been produced; all other software packages used in this study were previously published, as indicated both in the Methods section and the Reporting Summary.

References

1. Grassi, F. & De Lorenzis, G. Back to the origins: Background and perspectives of grapevine domestication. *Int. J. Mol. Sci.* **22**, 4518 (2021).
2. McGovern, P. et al. Early Neolithic wine of Georgia in the South Caucasus. *Proc. Natl. Acad. Sci.* **114**, 10309–10318 (2017).
3. Brun, JP. Viticulture et oléiculture en Gaule. In: Ouzoulias, P. & Tranoy, L. (eds), *Comment les Gaules devinrent romaines, La Découverte*, Chap. **15**, 231-253 (2010).
4. Brun JP. Archéologie du vin et de l'huile. De la préhistoire à l'époque hellénistique. Paris: Errance. 226 (2004).
5. Dietler, M. Alcohol: Anthropological/archaeological perspectives. *Annu. Rev. Anthropol.* **35**, 229–249 (2006).
6. Estreicher, S. K. Wine and France: A brief history. *Eur. Rev.* **31**, 91–179 (2023).
7. Alonso Ugaglia, A., Cardebat, J.-M. & Jiao, L. The French Wine Industry. In *The Palgrave Handbook of Wine Industry Economics*

(eds. Alonso Ugaglia, A., Cardebat, J.-M. & Corsi, A.) 17–46 (Springer International Publishing, Cham, 2019).
8. Bouby, L. et al. Bioarchaeological Insights into the Process of Domestication of Grapevine (*Vitis vinifera L.*) during Roman Times in Southern France. *PLoS ONE* **8**, e63195 (2013).
9. Rice, C. M. Comparative advantage, specialized viticulture, and the economic development of Gallia Narbonensis. *J. Rom. Archaeol.* **36**, 261–299 (2023).
10. Bouby, L. et al. The Holocene history of grapevine (*Vitis vinifera*) and viticulture in France retraced from a large-scale archaeobotanical dataset. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **625**, 111655 (2023).
11. Bouby, L. et al. Seed morphometrics unravels the evolutionary history of grapevine in France. *Sci. Rep.* **14**, 22207 (2024).
12. Van Limbergen, et al. (2025). Vine-growing and Winemaking in the Roman World: New data and original perspectives. Peeters Publishers (2025).
13. Dodd, E. The Archaeology of Wine Production in Roman and Pre-Roman Italy. *Am. J. Archaeol.* **126**, 443–480 (2022).
14. Boissinot, P. Les vignobles des environs de Mégara Hyblaea et les traces de la viticulture italienne durant l'Antiquité. *Mélanges LÉcole Fr. Rome Antiq.* **121**, 83–132 (2009).
15. Laucou, V. et al. Extended diversity analysis of cultivated grapevine *Vitis vinifera* with 10K genome-wide SNPs. *PLOS ONE* **13**, e0192540 (2018).
16. Ramos-Madrigal, J. et al. Palaeogenomic insights into the origins of French grapevine diversity. *Nat. Plants* **5**, 595–603 (2019).
17. Mayer-Olivé, M. "THE MEDITERRANEAN NAVIGATION AND THE BALEARIC ISLANDS: 432. Shipping and the Movement of Materials and Products in the Roman Mediterranean, with Particular Reference to the Reflection in the Balearic Islands." In J. Velaza, Ed., Insularity, Identity and Epigraphy in the Roman World, Cambridge (2017).
18. Briggs, A. W. et al. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci.* **104**, 14616–14621 (2007).
19. Breglia, F., Bouby, L., Wales, N., Ivorra, S. & Fiorentino, G. Disentangling the origins of viticulture in the western Mediterranean. *Sci. Rep.* **13**, 17284 (2023).
20. Dong, Y. et al. Dual domestications and origin of traits in grapevine evolution. *Science* **379**, 892–901 (2023).
21. Liang, Z. et al. Whole-genome resequencing of 472 Vitis accessions for grapevine diversity and demographic history analyses. *Nat. Commun.* **10**, 1190 (2019).
22. Zhou, Y. et al. The population genetics of structural variants in grapevine domestication. *Nat. Plants* **5**, 965–979 (2019).
23. Kim, M.-S., Hur, Y. Y., Kim, J. H. & Jeong, S.-C. Genome resequencing, improvement of variant calling, and population genomic analyses provide insights into the seedlessness in the genus *Vitis*. *G3 GenesGenomesGenetics* **10**, 3365–3377 (2020).
24. Liu, B.-B. et al. Capturing single-copy nuclear genes, organellar genomes, and nuclear ribosomal DNA from deep genome skimming data for plant phylogenetics: A case study in Vitaceae. https://doi.org/10.1111/jse.12806 (2021).
25. Ma, Z.-Y. et al. Phylogenomic relationships and character evolution of the grape family (Vitaceae). *Mol. Phylogenet. Evol.* **154**, 106948 (2021).
26. Morales-Cruz, A. et al. Introgression among North American wild grapes (Vitis) fuels biotic and abiotic adaptation. *Genome Biol.* **22**, 254 (2021).
27. Ma, Z. et al. Phylogenetic relationships, hybridization events, and drivers of diversification of East Asian wild grapes as revealed by phylogenomic analyses. *J. Syst. Evol.* **61**, 273–283 (2023).
28. Green, R. E. et al. A Draft Sequence of the Neandertal Genome. *Science* **328**, 710–722 (2010).
29. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**, 1655–1664 (2009).
30. Harney, É, Patterson, N., Reich, D. & Wakeley, J. Assessing the performance of qpAdm: a statistical tool for studying population admixture. *Genetics* **217**, iyaa045 (2021).
31. Di Vecchi-Staraz, M. et al. Low level of pollen-mediated gene flow from cultivated to wild grapevine: Consequences for the evolution of the endangered subspecies *Vitis vinifera L.* subsp. *silvestris*. *J. Hered.* **100**, 66–75 (2009).
32. Xiao, H. et al. Adaptive and maladaptive introgression in grapevine domestication. *Proc. Natl. Acad. Sci.* **120**, e2222041120 (2023).
33. Alaçamlı, E. et al. READv2: Advanced and user-friendly detection of biological relatedness in archaeogenomics. *Genome Biol.* **25**, 216 (2024).
34. Brun, JP. Vin et Huile Dans La Mediterranee Antique: Viticulture, oléiculture et procédés de fabrication. Paris errance (2003).
35. Jung, C et al. Lattes, Saint-Pierre Sud. Un ensemble funéraire protohistorique au sein de la campagne lattoise. Occitanie, Hérault (34). Ligne Grande Vitesse, Contournement de Nîmes et Montpellier. [Rapport de recherche] Institut national de recherches archéologiques préventives, INRAP. (2017).
36. Wagner, S. et al. High-Throughput DNA sequencing of ancient wood. *Mol. Ecol.* **27**, 1138–1154 (2018).
37. Wagner, S. et al. Tracking population structure and phenology through time using ancient genomes from waterlogged white oak wood. *Mol. Ecol.* **33**, e16859 (2024).
38. McGovern, P. E. et al. Beginning of viniculture in France. *Proc. Natl. Acad. Sci.* **110**, 10147–10152 (2013).
39. Chabal, L. & Heinz, C. Reconstructing the heterogeneity of past woodlands in anthracology using the spatial distribution of charcoals in archaeological layers: Applied to the postglacial occupation of the Abeurador cave (Hérault) in the South of France. *Quat. Int.* **593–594**, 19–35 (2021).
40. Pérez-Jordà, G., Peña-Chocarro, L., García Fernández, M. & Vera Rodríguez, J. C. The beginnings of fruit tree cultivation in the Iberian Peninsula: plant remains from the city of Huelva (southern Spain). *Veg. Hist. Archaeobotany* **26**, 527–538 (2017).
41. Bouby, L. L'agriculture dans le bassin du Rhône du bronze final à l'Antiquité: agrobiodiversité, économie, cultures. (AED, Archives d'écologie préhistorique, Toulouse, 2014).
42. Dion, R. *Histoire de La Vigne et Du Vin En France Des Origines Au XIXe Siècle*. CNRS (2011).
43. André, J. & Levadoux, L. La vigne et le vin des Allobroges. *J. Savants* **3**, 169–181 (1964).
44. Bonhomme, V. et al. Seed morphology uncovers 1500 years of vine agrobiodiversity before the advent of the Champagne wine. *Scientific Reports* **11**, 2305 (2021).
45. Bowers, J. et al. Historical Genetics: The Parentage of Chardonnay, Gamay, and Other Wine Grapes of Northeastern France. *Science* **285**, 1562–1565 (1999).
46. Lacombe, T. et al. Large-scale parentage analysis in an extended set of grapevine cultivars (*Vitis vinifera L.*). *Theor. Appl. Genet.* **126**, 401–414 (2013).
47. Anderson, K. & Nelgen, S. Internationalization, premiumization and diversity of the world's winegrape varieties. *J. Wine Res.* **32**, 247–261 (2021).
48. Grillon, G., Medigue, C., Garcia, J.-P. & Labbé, T. Le «très loyal pinot»: itinéraire d'un cépage mythique de la Bourgogne. *Crescentis* https://doi.org/10.58335/crescentis.1003 (2019).
49. Sousa Da Mota, B. et al. Imputation of ancient human genomes. *Nat. Commun.* **14**, 3660 (2023).
50. Cohen, P. et al. Ancient DNA from a lost Negev Highlands desert grape reveals a Late Antiquity wine lineage. *Proc. Natl. Acad. Sci.* **120**, e2213563120 (2023).

51. Liu, Z. et al. Grapevine pangenome facilitates trait genetics and genomic breeding. *Nat. Genet*. https://doi.org/10.1038/s41588-024-01967-5 (2024).
52. Wales, N. et al. The limits and potential of paleogenomic techniques for reconstructing grapevine domestication. *J. Archaeol. Sci.* **72**, 57–70 (2016).
53. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20130624 (2015).
54. Schubert, M. et al. Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nat. Protoc.* **9**, 1056–1082 (2014).
55. Canaguier, A. et al. A new version of the grapevine reference genome assembly (12X.v2) and of its annotation (VCost.v3). *Genomics Data* **14**, 56–62 (2017).
56. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
57. Poullet, M. & Orlando, L. Assessing DNA sequence alignment methods for characterizing ancient genomes and methylomes. *Front. Ecol. Evol.* **8**, 105 (2020).
58. 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).
59. Bouby, L. et al. Tracking the history of grapevine cultivation in Georgia by combining geometric morphometrics and ancient DNA. *Veg. Hist. Archaeobotany* **30**, 63–76 (2021).
60. Chang, C. C. et al. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **4**, s13742–015-0047–8 (2015).
61. Charrad, M., Ghazzali, N., Boiteau, V. & Niknafs, A. NbClust: An *R* package for determining the relevant number of clusters in a data set. *J. Stat. Softw.* **61**, 1–36 (2014).
62. Lefort, V., Desper, R. & Gascuel, O. FastME 2.0: A Comprehensive, accurate, and fast distance-based phylogeny inference program: Table 1. *Mol. Biol. Evol.* **32**, 2798–2800 (2015).
63. Patterson, N. et al. Ancient Admixture in Human History. *Genetics* **192**, 1065–1093 (2012).
64. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet* **2**, e190 (2006).
65. Röckel, F. et al. Vitis International Variety Catalogue. Available at: http://www.vivc.de (accessed June 2025).

Acknowledgements

We thank all staff members at the Centre for Anthropobiology and Genomics of Toulouse for fruitful discussion, as well as Stéphanie Schiavinato and Laure Tonasso-Calvière for technical support and maintenance of ancient DNA facilities, and Antoine Lacombe for hardware maintenance of computational clusters. This project received funding from the French national research agency (Agence Nationale pour la Recherche), through the Viniculture (ANR-16-CE27- 0013; LB, LO), MICA (ANR-22-CE27-0026; LB, LO, RN), ISEMA (ANR-23-CE27-0003; J.R. Cine Serra/Ses Feixes project), and ArkaeoAG (ANR-20-CE27-0013; LO) collaborative research projects. Over the course of this study, Ludovic Orlando was supported by the European Research Council (PEGASUS-681605 and HorsePower-101071707), and Roberto Bacillieri work was financially supported by INRAE France.

Author contributions

Designed the study: L.B., R.B. and L.O. Performed wet lab work: L.C., N.W., with input from S.W. and L.O. Performed data analysis: R.N. and L.O., with input O.E. Prepared figures: R.N., with input from L.O. Provided material, samples and reagents: E.B., M.C., M.D., I.F., C.H., P.M., V.M., J.R., N.R., M.T., G.P., C.J., I.D., M.R-Z., B.H., R.M., J.R., C.N., O.G., E.G., P.K., E.M., M.P., M.B., J.C.-L., D.D., L.L., H.P., C.R., M.S., M.T.-T., J.T.-C., L.B., and L.O. Interpreted data: R.N., G.B, J.T., R.B, L.B., and L.O. Wrote the manuscript draft: L.O. Edited the manuscript: all co-authors. Wrote the Supplementary Material: R.N., L.B., and L.O.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-70166-z.

Correspondence and requests for materials should be addressed to Laurent Bouby or Ludovic Orlando.

Peer review information *Nature Communications* thanks Emlyn Dodd, Zeki Kaya and the other anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

[1]Centre d'Anthropobiologie et de Génomique de Toulouse, Université de Toulouse, CNRS UMR 5288, 37 Allées Jules Guesde, Toulouse, France. [2]ISEM, Univ Montpellier, CNRS, IRD, EPHE, Inrap, Montpellier, France. [3]Biodiversité Gènes et Communautés,Université de Bordeaux, INRAE UMR 1202, 69 route d'Arcachon, Cestas, France. [4]Archéozoologie, Archéobotanique - Sociétés, Pratiques et Environnement, CNRS-MNHN UMR 7209, 55 rue Buffon, Paris, France. [5]Archéologie Alsace, 11 rue champollion, Sélestat, France. [6]Institut National de Recherches Archéologiques Préventives, rue d'Alésia, Paris, France. [7]GEOLAB - Laboratoire de Géographie Physique et Environnementale, Université Clermont Auvergne, CNRS, Inrap UMR 6042, 39 rue camille-Guérin, 8, Limoges, France. [8]Centre de Recherche Archéologique de la Vallée de l'Oise, CNRS-MNHN UMR 7209, 17 rue James de Rothschild, Compiègne, France. [9]Hadès bureau d'investigation archéologiques, 9 rue d'Ariane, L'Union, France. [10]ASM, Université Paul-Valéry Montpellier, INRAP, Ministère de la culture, Montpellier, France. [11]Ipso Facto, 17 chemin de séverin, Arles, France. [12]Service Art, Histoire et Archéologie, 14 cours Aristide Briand, Martigues, France. [13]Mosaïques Archéologie,

Domaine La Barthe, domaine Barthe, Cournonterral, France. [14]Service archéologique municipal de Valenciennes, Rue des Archers, Valenciennes, France. [15]Ministère de la Culture, DRAC Grand-Est, Metz, France. [16]Antiquarium—Arqueológia y patrimonio, Avda Isidoro Macabich 19, 07820 Sant Antoni de Portmany, Eivissa- Illes Balears, Spain. [17]ARChéologies et Sciences de l'ANtiquité, UMR 7041, Université Paris Nanterre, 200 avenue de la République, Nanterre, France. [18]Culture et Environnement, Préhistoire, Antiquité, Moyen-âge, Université Nice Côte d'Azur, CNRS UMR 7264, 24 avenue des Diables Bleus, Nice, France. [19]ARChéologie et HIstoire ancienne: MEDiterranée-Europe, Université de Strasbourg, CNRS UMR 7044, 5 allée du Général Rouvillois, Strasbourg, France. [20]Arqueología Africana, 971713, Universidad Complutense de Madrid, Madrid, Spain. [21]Centre de Recherche sur la Biodiversité et l'Environnement, Université de Toulouse, CNRS, IRD, UMR 5300, 118 route de Narbonne, Toulouse, France. [22]UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France. ✉e-mail: laurent.bouby@umontpellier.fr; ludovic.orlando@utoulouse.fr