Dual domestications and origin of traits in grapevine evolution

Yang Dong[1,2]†, Shengchang Duan[1,2]†, Qiuju Xia[3]†, Zhenchang Liang[4]†, Xiao Dong[1,2]†§, Kristine Margaryan[5,6]‡, Mirza Musayev[7]‡, Svitlana Goryslavets[8]‡, Goran Zdunić[9]‡, Pierre-François Bert[10]‡, Thierry Lacombe[11]‡, Erika Maul[12]‡, Peter Nick[13]‡, Kakha Bitskinashvili[14]‡, György Dénes Bisztray[15]‡, Elyashiv Drori[16,17]‡, Gabriella De Lorenzis[18]‡, Jorge Cunha[19,20]‡, Carmen Florentina Popescu[21]‡, Rosa Arroyo-Garcia[22]‡, Claire Arnold[23]‡, Ali Ergül[24]‡, Yifan Zhu[1]‡, Chao Ma[25]‡, Shufen Wang[1,2], Siqi Liu[1,2], Liu Tang[1,2], Chunping Wang[1,2], Dawei Li[1,2], Yunbing Pan[1,2], Jingxian Li[1,2], Ling Yang[1,2], Xuzhen Li[1,2], Guisheng Xiang[1,2], Zijiang Yang[1,2], Baozheng Chen[1,2], Zhanwu Dai[4], Yi Wang[4], Arsen Arakelyan[5,26,27], Varis Kuliyev[28], Gennady Spotar[8], Nabil Girollet[10], Serge Delrot[10], Nathalie Ollat[10], Patrice This[11], Cécile Marchal[29], Gautier Sarah[11], Valérie Laucou[11], Roberto Bacilieri[11], Franco Röckel[12], Pingyin Guan[13], Andreas Jung[30], Michael Riemann[13], Levan Ujmajuridze[14], Tekle Zakalashvili[14], David Maghradze[14], Maria Höhn[15], Gizella Jahnke[15], Erzsébet Kiss[15], Tamás Deák[15], Oshrit Rahimi[16], Sariel Hübner[31], Fabrizio Grassi[32], Francesco Mercati[33], Francesco Sunseri[34], José Eiras-Dias[19,20], Anamaria Mirabela Dumitru[21], David Carrasco[22], Alberto Rodriguez-Izquierdo[22], Gregorio Muñoz[35], Tamer Uysal[36], Cengiz Özer[36], Kemal Kazan[37], Meilong Xu[38], Yunyue Wang[1], Shusheng Zhu[1], Jiang Lu[39], Maoxiang Zhao[25], Lei Wang[25], Songtao Jiu[25], Ying Zhang[40], Lei Sun[40], Huanming Yang[41], Ehud Weiss[42], Shiping Wang[25], Youyong Zhu[1], Shaohua Li[4]*, Jun Sheng[1,2]*, Wei Chen[1,2]*

Affiliations:

[1]State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan Agricultural University; Kunming, 650201, China.

[2]Yunnan Research Institute for Local Plateau Agriculture and Industry; Kunming, 650201, China.

[3]State Key Laboratory of Agricultural Genomics, BGI-Shenzhen; Shenzhen, 518083, China.

[4]Beijing Key Laboratory of Grape Science and Oenology and Key Laboratory of Plant Resources, Institute of Botany, the Chinese Academy of Sciences; Beijing, 100093, China.

[5]Institute of Molecular Biology, NAS RA; Yerevan, 0014, Armenia.

[6]Yerevan State University; Yerevan, 0014, Armenia.

[7]Genetic Resources Institute, Azerbaijan National Academy of Sciences; Baku, AZ1106, Azerbaijan.

[8]National Institute of Viticulture and Winemaking ‘Magarach’; Yalta, 298600, Crimea.

[9]Institute for Adriatic Crops and Karst Reclamation; Split, 21000, Croatia.

[10]Bordeaux University, Bordeaux Sciences Agro, INRAE, UMR EGFV, ISVV; Villenave d'ornon, 33882, France.

[11]AGAP Institut, University of Montpellier, CIRAD, INRAE, Institut Agro Montpellier; Montpellier, 34398, France.

[12]Julius Kühn Institute (JKI) – Federal Research Center for Cultivated Plants, Institute for Grapevine Breeding Geilweilerhof; Siebeldingen, 76833, Germany.

[13]Botanical Institute, Karlsruhe Institute of Technology; Karlsruhe, 76131, Germany.

[14]LEPL Scientific Research Center of Agriculture; Tbilisi, 0159, Georgia.

[15]Hungarian University of Agriculture and Life Sciences (MATE); Budapest, 1118, Hungary.

[16]Department of Chemical Engineering, Ariel University; Ariel, 40700, Israel.

[17]Eastern Regional R&D Center; Ariel, 40700, Israel.

[18]Department of Agricultural and Environmental Sciences, University of Milano; Milano, 20133, Italy.

[19]Instituto Nacional de Investigação Agrária e Veterinária, I.P./INIAV-Dois Portos; Torres Vedras, 2565-191, Portugal.

[20]Green-it Unit, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa; Oeiras, 2780-157, Portugal.

[21]National Research and Development Institute for Biotechnology in Horticulture; Stefanesti, Arges, 117715, Romania.

[22]Center for Plant Biotechnology and Genomics, UPM-INIA/CSIC; Pozuelo de Alarcon, Madrid, 28223, Spain.

[23]University of Lausanne; Lausanne, 1015, Switzerland.

[24]Biotechnology Institute, Ankara University; Ankara, 06135, Türkiye.

[25]Department of Plant Science, School of Agriculture and Biology, Shanghai JiaoTong University; Shanghai, 200240, China.

[26]Armenian Bioinformatics Institute; Yerevan, 0014, Armenia.

[27]Biomedicine and Pharmacy, RAU; Yerevan, 0051, Armenia.

[28]Institute of Bioresources, Nakhchivan Branch of the Azerbaijan National Academy of Sciences; Nakhchivan, AZ7000, Azerbaijan.

[29]Vassal-Montpellier Grapevine Biological Resources Center, INRAE; Marseillan-plage, 34340, France.

[30]Historische Rebsorten-Sammlung, Rebschule (K39); Gundheim, 67599, Germany.

[31]Galilee Research Institute (Migal), Tel-Hai Academic College; Upper Galilee, 12210, Israel.

[32]Department of Biotechnology and Biosciences, University of Milano-Bicocca; Milano, 20126, Italy.

[33]Institute of Biosciences and Bioresources, National Research Council; Palermo, 90129, Italy.

[34]Department AGRARIA, University Mediterranea of Reggio Calabria; Reggio Calabria, 89122, Italy.

[35]IMIDRA, Alcalá de Henares; Madrid, 28805, Spain.

[36]Viticulture Research Institute, Ministry of Agriculture and Forestry; Tekirdağ, 59200, Türkiye.

[37]Commonwealth Scientific and Industrial Research Organization (CSIRO) Agriculture and Food; Queensland Bioscience Precinct, St. Lucia, Queensland, 4067, Australia.

[38]Institute of Horticulture, Ningxia Academy of Agricultural and Forestry Sciences; Yinchuan, 750002, China.

[39]Center for Viticulture and Oenology, School of Agriculture and Biology, Shanghai JiaoTong University; Shanghai, 200240, China.

[40]Zhengzhou Fruit Research Institutes, CAAS; Zhengzhou, 450009, China.

[41]BGI-Shenzhen; Shenzhen, 518083, China.

[42]The Martin (Szusz) Department of Land of Israel Studies and Archaeology, Bar-Ilan University, Ramat-Gan, 5290002, Israel.

†These authors contributed equally to this work.

‡Institution contacts for biological samples. Ordered by country names.

§Current Address: Department of Chromosome Biology, Max Planck Institute for Plant Breeding Research; Cologne, 50829, Germany.

*Corresponding author. Email: wchenntr@gmail.com (W.C.); shengjun@dongyang-lab.org (J.S.); shhli@ibcas.ac.cn (S.L.).

Abstract: We elucidate grapevine evolution and domestication histories with 3,525 cultivated and wild accessions worldwide. In the Pleistocene, harsh climate drove the separation of wild ecotypes due to continuous habitat fragmentation. Then, domestication occurred concurrently about 11,000 years ago in the Near East and the Caucasus to yield table and wine grapevines. The Near East domesticates dispersed into Europe with early farmers, introgressed with ancient wild western ecotypes, and subsequently diversified along human migration trails into muscat and unique western wine grape ancestries by the late Neolithic. Analyses of domestication traits also reveal novel insights into selection for berry palatability, hermaphroditism, muscat flavor, and berry skin color. These data demonstrate the role of grapevine in the early inception of agriculture across Eurasia.

One-Sentence Summary: The dual origin and diversification of grapevine is a testament to early human migration and the development of various Eurasian civilizations.

The cultivated grapevine (*V. vinifera* ssp. *vinifera*, hereafter *V. vinifera*) shares a close relationship with humans (*1*). With unmatched cultivar diversity, this food source (table and raisin grapes) and winemaking ingredient (wine grapes) became an emblem of cultural identity in major Eurasian civilizations (*1–3*), leading to intensive research in ampelography, archaeobotany, and historical records to reveal its history (*4*). Early work asserted that *V. vinifera* originated from its wild progenitor *V. vinifera* ssp. *sylvestris* (hereafter *V. sylvestris*) about 8,000 years ago (ya) during the Neolithic agricultural revolution in the Near East (*5*, *6*). In recent years, various genetic studies explored this proposition (*6–13*), but the critical details of grapevine domestication were often inconsistent. Studies argued for the existence of domestication centers in the western Mediterranean (*13*), Caucasus (*12*, *14*), and Central Asia (*12*), which in turn cast doubt on the popular notion of a single past domestication event (*10*, *11*). Three demographic inferences yielded population split times between *V. vinifera* and *V. sylvestris* to dates between 15-400 thousand years ago (Kya), predating the historical consensus on domestication time (*7–9*). As early domesticates spread to other parts of Eurasia via poorly defined migration routes in the ensuing millennia (*5*), the single-origin theory also confounds the origin order between table and wine grapevines. One view proposes a wine grapevine-first model with the two types diverging about 2,500 ya (*7*, *10*, *11*). Hybridization with local *V. sylvestris* was common in creating extant European wine grapes (*10*, *11*), but when these introgression events occurred is unknown. Several studies suggest that the earliest cultivation of European wine grapes in France and Iberia postdates 3000 ya (*10*, *15*). These discrepancies primarily result from the inadequate sampling of grapevine accessions and the limited resolution of genetic data in previous analyses. Therefore, we report the genomic variation dataset from a global cohort to systematically

delineate the structure of grapevine genetic diversity, explore the origin of *V. vinifera*, deduce a putative dispersal history, and investigate key domestication traits and diversification signatures.

Results

We constructed a chromosomal-level reference *V. sylvestris* genome assembly (VS-1 from Tunisia) to attain genomic variations, which shows a higher percentage of anchored chromosomal lengths than PN40024 (fig. S1, table S1-S9) (*16*). From the 3,304 assembled accessions from a dozen Eurasian germplasm and private collections, we obtained good-quality Illumina paired-end sequencing data to an average 20× coverage for 3,186 grapevine accessions (2,237 *V. vinifera* and 949 *V. sylvestris*; table S10-S13). The sample selection preferentially includes old, autochthonous, and economically important varieties to maximize the spectrum of genetic diversity. We also included genomic data for 339 previously sequenced accessions (266 *V. vinifera* and 73 *V. sylvestris*; table S14) in the analyses (*7*, *8*, *17*), producing the final cohort of 3,525 grapevine accessions (2,503 *V. vinifera* and 1,022 *V. sylvestris*). The alignment of the Illumina reads to the VS-1 reference genome identifies 45,624,306 biallelic SNPs and 7,314,397 biallelic short Indels (≤40 bp; 73.2% shorter than five bps) (*16*), among which rare alleles (minor allele frequency ≤ 1%) account for the majority (fig. S2, table S15-S22).

Core accessions differentiate by eight distinct genetic ancestries

Clones, mutants, synonyms, and homonyms are common phenomena in grapevine germplasm and collections (*18*). Using the identity-by-state sharing pattern estimators, we found 1,534 accessions sharing the genetic profile with at least one other in the cohort, totaling 498 distinct genotypes (fig. S3, table S23) (*16*). We kept one accession for each distinct genotype, corrected

misidentified accessions, and excluded interspecific hybrids for a core cohort of 2,448 grapevines (1,604 *V. vinifera* and 844 *V. sylvestris*; fig. S3), which remain representative of the major viticultural regions (*19*) in the world (Fig. 1A, fig. S3).

The principal component analysis (PCA) shows that *V. sylvestris* and *V. vinifera* separately spread out along the first two axes (total variance explained, PC1 7.56%, PC2 1.71%), with both displaying a crude Near East to Western Europe gradient (Fig. 1B, fig. S4-S5). The PC3 axis (1.26% variance) separates *V. vinifera* individuals according to their utilization, agreeing with the main table and wine grapevine clades in the maximum likelihood phylogenetic tree and reticulate phylogenetic network (fig. S6-S7). Notably, the *V. vinifera* accessions show a weak isolation-by-distance correlation (Fig. 1C), suggesting a disconnection between the viticultural geographic pattern and the genetic structures in the grapevine (*20*). This observation could be due to the extensive exchange of superior cultivars across regions and the subsequent interbreeding throughout history.

Given the poor resolution of viticultural regions in defining grapevine diversity, we leveraged genetic ancestry information from an unsupervised ADMIXTURE analysis to categorize core accessions (Fig. 1D, fig. S8) (*16*). At K=2, all *V. vinifera* accessions contain a majority east (red) ancestry that matches the ancestry of the *V. sylvestris* accessions in the East Mediterranean region. At K=8, hierarchical clustering of ancestry components identifies four *V. sylvestris* groups from distinct geographic regions: the Near East (Syl-E1, 84.3% K2), the Caucasus (Syl-E2, 72.7% K6), Central Europe (Syl-W1, 94.7% K1), and the Iberian Peninsula (Syl-W2, 69.8% K8; Fig. 1D-1F). *V. sylvestris* accessions collected from other regions show admixed genetic

structures (*16*). For cultivated grapevines, six genetic ancestries could designate six distinctive groups (CG1 to CG6), all covering a broad range of viticultural regions (Fig. 1D to 1F) (*16*). Accessions with pure or close to pure ancestries (Fig. S9) (*16*) help ascribe names to these groups as Near East table grapevines (CG1, 73.9% *K*2), Caucasian wine grapevines (CG2, 66.4% *K*6), muscat grapevines (CG3, 87.7% *K*5), Balkan wine grapevines (CG4, 69.9% *K*4), Iberian wine grapevines (CG5, 68.8% *K*7), and Western European wine grapevines (CG6, 68.4% *K*3). The admixed *V. vinifera* accessions showed different combinations of genetic ancestries (fig. S9). The four *V. sylvestris* and six *V. vinifera* groups, supported by archetypal analysis at *K*=8 (fig. S10), form identifiable clusters in the PCA plots (Fig. 1G, fig. S4), thus suitable for population genomic investigations.

Separation of *V. sylvestris* ecotypes in Pleistocene

According to the genetic ancestries and the occupied ecological niches in the western Eurasia continent, we designate *V. sylvestris* accessions in the Near East and the Caucasus as the eastern ecotype (Syl-E) and accessions in Central Europe and the Iberian Peninsula as the western ecotype (Syl-W; Fig. 2A). The large between-ecotype fixation index values (e.g., Syl-E1 vs. Syl-W1, F_{ST}=0.340) and the small within-ecotype fixation index values (Syl-E1 vs. Syl-E2, F_{ST}=0.101; Syl-W1 vs. Syl-W2, F_{ST}=0.072; fig. S11, table S26) support this designation. Both nucleotide diversity (π) and individual heterozygosity show that the western ecotype (especially Syl-W1) has significantly reduced variation compared to its eastern counterpart (fig. S11). Furthermore, the linkage disequilibrium decay (LD, r^2) was much slower in Syl-W (1.0-1.6Kb at

half of maximum r^2) than in Syl-E (400-600bp at half of maximum r^2; fig. S12). These data demonstrate that the eastern ecotype retains more genetic diversity.

Demographic inference with folded SNP frequency spectra reveals an ancient population bottleneck in Syl-E around 400-800 Kya and in Syl-W around 150-400 Kya (Fig. 2B, fig. S13). This Pleistocene period, characterized by changing climate cycles (*21*, *22*), also witnessed the deduced population split (median time ~200-400 Kya) between the two ecotypes (Fig. 2C). The slow descent of the split line suggests that the geographic isolation process was gradual (fig. S13). At ~56 Kya, the population split between Syl-E1 and Syl-E2 occurred during the last glacial cycle (11.7-115 Kya) when the global climate trended toward dryer and colder conditions (*23*). Close to the time of the Last Glacial Maximum (LGM at ~21 Kya), *V. sylvestris* subgroups experienced a second population bottleneck (~40 Kya), with effective population sizes (N_e) reaching a minimum of 10,000 to 40,000 (Fig. 2B, fig. S13). Following this result, ecological niche modeling predicts that the areas with suitable environmental conditions for Syl-E and Syl-W (suitability>0.75) remained connected at the Pleistocene Last Interglacial (~130 Kya, fig. S14) but became entirely separated at the LGM (Fig. 2D). The post-bottleneck N_e rebound was steeper in the Syl-W accessions, but the numbers decreased to lower levels in recent times (Fig. 2B, fig. S13). This result agrees with the reduced genetic diversity in Syl-W and the abrupt population split between Syl-W1 and Syl-W2 at ~2.5 Kya.

Dual origin of *V. vinifera* at the advent of agriculture

The wet climate in the Early Holocene (~11.7-8.3 Kya) (*24*) facilitated the expansion of suitable habitats for Syl-E, resulting in a large geographic span from Central Asia to the Iberian Peninsula

(Fig. 2D). This expansion supports the eastern origin and subsequent continental dispersal of *V. vinifera*. Since CG1 shares the main ancestral component with Syl-E1 and CG2 with Syl-E2 (Fig. 1D and 1F), the possibility of two domestication events becomes evident. Indeed, both CG1 and CG2 maintain the highest genetic diversity and manifest the quickest LD decay among all CG groups (fig. S11 and S12). Furthermore, they are less differentiated from their corresponding wild ecotypes (Fig. 3A, fig. S11). The AIC-based phylogenetic selection also prefers a dual origin tree model (fig. S15), which agrees with the outgroup f_3 statistics bi-plots that CG1 and CG2 are genetically closer to Syl-E1 and Syl-E2, respectively (Fig. 3B, fig. S15, table S27). Notably, the population split lines of CG1/Syl-E2 and CG2/Syl-E1 pairs resemble that of Syl-E1/Syl-E2 and differ from those of CG1/Syl-E1 and CG2/Syl-E2 pairs (Fig. 3C, fig. S16). These data collectively support a dual origin of *V. vinifera* and reject the popular theory of a single primary domestication center (*10*, *11*). Both CG1/Syl-E1 and CG2/Syl-E2 population pairs separated quickly (Fig. 3C), which is compatible with a clean split scenario. We estimate the median population split time to be ~11 Kya (95% confidence interval: ~10.5-12.5 Kya) for both pairs, suggesting that the independent domestication events took place concurrently around the advent of agriculture. As CG1 and CG2 separately represent table and wine grapevine ancient genetic backgrounds (*K*2 and *K*6; fig. S9), the dual origin rejects the assumption that wine grapevine predates table grapevine (*7*, *10*, *11*).

The dispersal of grapevine domesticates along human migration routes

The geographic distributions of CG1 and CG2 cultivars across Eurasia and North Africa correspond to vastly different human migration routes for the two grapevine groups (Fig. 3D). The CG2 cultivars were mainly confined to both sides of the Caucasus Mountains, with a limited

dispersal into the Carpathian Basin by the northern Black Sea. This result contrasts with previous models implying that CG2 played a central role in the formation of wine grapevines in Europe (*3*). Instead, CG2 represents a local domestication effort that had a minor impact on grapevine diversification. In comparison, the dispersal of CG1 in four directions spanned Eurasia and North Africa. First, the eastward expansion through Central Asia into India and China follows the Inner Asia Mountain Corridor, a path that also witnessed the exchange of other crops (i.e., wheat, barley, and millet) between the West and the East (*25*). Second, the northbound expansion could mirror the early cultural contact of the Near East over the Zagros mountains with the Caucasus (*26*, *27*). Third, the northwest expansion via Anatolia into the Balkans bespeaks the spread of farming into Europe (*28*, *29*). Finally, a westward expansion moved across the North African coastline to reach Morocco (*30*). Even though grapevine domesticates followed the trails of past human migration, the timing and dispersal details require paleogenomic data for delineation.

Shared and unique domestication signatures in CG1 and CG2 grapevines

Given the dual origin scenario, we investigated domestication signatures in both Syl-E1/CG1 and Syl-E2/CG2 group pairs by selecting genomic regions that display increased nucleotide diversity differences and population differentiation (both top 5%; Fig. 3D). This method yields 1,140 domestication selective sweep genes in 132 regions for CG1 and 887 genes in 137 regions for CG2 (table S28), among which only 189 genes in 31 regions exist in both groups (table S29). Most shared signals are on chromosomes 2 and 17, confirming previous investigations that the selection on flower sexual morphs (sex determination region, SDR), berry skin color (*VvMybA* gene cluster), and berry development (*SDH* gene cluster) are of great importance during grapevine domestication (*8*, *11*). In addition, our analysis also identifies shared domestication

genes that possibly underlie grapevine growth (e.g., *NPF*; see (*16*) for gene descriptions), physiology (e.g., *FER4*), fruit set (e.g., *GA2OX* gene cluster), and resistance to biotic/abiotic stress (e.g., *FER4*; *PPR* gene cluster; *RNF181* gene cluster).

As expected for dual domestications, most selective sweep signatures in CG1 and CG2 are unique and target distinctive chromosomal regions (Fig. 3E). Even though CG1 and CG2 correspondingly represent table and wine grapevines, many unique signatures seem to suggest a convergent selection mechanism targeting different aspects of common domestication traits. An obvious example is the improvement of berry palatability through the reduction of alkaloid biosynthesis (*MecgoR* gene cluster in CG1; *TR2* and *SSL* gene clusters in CG2) and the enhancement of carbohydrate metabolism (*SWEET17* in CG1; *PFKFB1* in CG2). Other examples include perceived berry desirability (*BEAT* gene cluster for floral scent in CG1; *UFGT* gene cluster for berry color in CG2) and response to environmental stresses (*UPL6* in CG1; *WAK* in CG2). These findings suggest that the initial cultivation of CG1 and CG2 may have been to serve early humans' caloric and micronutrient needs. The selection of genetic features suitable for winemaking in CG2 could have been serendipitous, and the practice of winemaking with CG2 (e.g., 8000 ya) (*14*) possibly postdates grapevine domestication. Since gene annotation depends on homology-based inference, it should be noted that many genes mentioned here need further verification in grapevine.

Wine grapevine diversification in Europe

As the CG1 early domesticates dispersed into Europe via Anatolia, a crucial question concerns the diversification history of European wine grapevines in the ensuing millennia. In particular,

the shared areas of suitable habitats for Syl-E and Syl-W in the early Holocene (black area in Fig. 2D) formed an ecological foundation for the genetic exchange between CG1 and local refugia Syl-W accessions in the coastal regions of northern Mediterranean and southern Black Sea, the Iberian Peninsula, and an area corresponding to present western France. It is therefore important to examine where and how distinct grapevine genetic ancestries (CG3-CG6) formed with relevance to Syl-W introgression (*10*, *11*). We have chosen cultivars in each group with at least 75% major ancestry (also average Syl-W ancestry in each *V. vinifera* group <3%) to perform population analyses. This selection rules out many old varieties [i.e., 'Lambrusco' cultivars deriving about half of their ancestries from Syl-W (fig. S9)], which likely showcase secondary diversification efforts after the distinct ancestries had been established. Interestingly, the TreeMix analysis finds one migration edge that points from Syl-W to a population ancestral to CG3-CG6 (estimated weight 0.114; Fig. 4A, fig. S17), suggesting an ancient introgression event occurred before the diversification of all European grapevines. An additional migration edge also points from Syl-W to CG6 (estimated weight, 0.292), which implies an independent introgression event unique to Western European wine grapevines in the past. Various combinations of D-statistics testing the gene flow from Syl-W into CG groups (Z-score>3.0, adjusted $P<4.17\times10^{-5}$; Fig. 4B, table S31) support this introgression history. Additionally, gene flow from Syl-W into CG3-CG6 inferred from Momi2 align with their corresponding divergence from CG1, further supporting the introgression history (Fig. 4C). Notably, the estimated median divergence times date the creation of muscat grapes (CG3) to 10,500 ya, Balkan wine grapes (CG4) to 8,070 ya, Iberian wine grapevines (CG5) to 7,740 ya, and Western European wine grapevines to 6,910 ya (Fig. 4D). These stepwise diversification times accord with the historical

migration of Anatolian farmers into Europe (*26*, *29*, *31*, *32*), substantiating the role of viticulture in forming Neolithic agricultural societies.

The migration edge weights, f_4-ratio, and Momi2 estimates collectively show that ancient introgression from Syl-W accounts for about 11.4-18.0% of the CG3-CG6 genomes (Fig. 4, table S30). On top of this, at least one other independent introgression event contributed about 25.0-30.0% additional Syl-W to the CG6 ancestry. We have screened the introgression tracts in CG3-CG6 by choosing the genomic windows with the top 1% d_f and f_{dM} values (fig. S18). A total of ten shared regions among CG3-CG6 groups contain genes that are putatively involved in plant immunity (e.g., *CYSK*), abiotic stress response (e.g., *GBA*), and carbohydrate metabolism (e.g., *TPS/TPP*; table S31). This result agrees with the proposal that introgression helps grapevines adapt to new environments and become more suitable for wine making (*10*, *11*).

Genetic analyses of domestication and diversification traits

Hermaphroditism: origin of H2 haplotype

The transition from dioecy in *V. sylvestris* (male, M/f; female, f/f) to hermaphroditism in *V. vinifera* is the most prominent phenotypic change during domestication (*33*). It involves recombination events between M and f around a selective sweep region on chromosome 2 known as the sex determination region (SDR; Fig. 5A). Previous studies have identified two major hermaphroditic haplotypes (H1 and H2) and four hermaphroditic genotypes (H1/f, H2/f, H1/H1, and H1/H2) from select cultivars (*33*), but the recombination history remains unclear. The analysis of our grapevine cohort reveals five recombination sites in the SDR region (Fig. 5B), which not only confirms known genotypes but also identifies novel minor haplotypes (male

variant Mv, female variant fv, H3, H4, and H5) and genotypes (Mv/f, M/H1, M/H5, H1/fv, H5/f, H4/f, H2/H2, and H2/H3) in both wild and cultivated grapevines (Fig. 5B, table S32). Among all SDR haplotypes, both M and H1 manifest the highest subtype diversity (fig. S19-S22). Furthermore, the SDR genotype statistics reveal a distribution bias of the H2-containing SDRs in the Iberian (CG5) and Western European grapevines (CG6; Fig. 5C, fig. S23). To investigate this observation, we constructed a putative recombination history for all known SDR haplotypes (Fig. 5D). It shows that a first recombination event between the parental M and f haplotypes created Mv (site 4), fv (site 3), H1 (site 2), and H4 (site 1). On this basis, H1 experienced a second recombination event with f to produce H3 (site 5) and H5 (site 4), whereas H4 recombined again with f at site 5 to bring about H2. Since three Syl-E *V. sylvestris* (IS164, IS167, and IS180) and 11 *V. vinifera* accessions in the cohort contain H4 (Fig. 4G), a likely scenario supports a westward dispersal of H4 after human selection to reach the Iberian Peninsula [e.g., in extant old Iberian cultivar 'Malvasia Fina' (PO153)], where H2 originated from H4 via secondary recombination and later became dominant during the diversification of Iberian and Western European cultivars.

Muscat flavor: trait selection may reduce grapevine fitness

Muscat grapevine is unique for its floral aromas, which result from a hard-to-define concoction of monoterpenoids in the fruit (*34*). Given the broad geographic distribution (fig. S24) and ancient history of muscat grapevines, it is difficult to pinpoint the center of origin. Momi2 estimate predicts a population split from CG1 at around 10,564 ya (Fig. 4C), which would suggest an origination site close to the Near East. This scenario agrees with the relatively low F_{ST} values and sizeable gene flow with CG1 (Fig. 4 and fig. S11). The CG3 group also show low

genetic diversity and high LD extent compared to others (fig. S11 and fig. S12). One possible reason is the gradual loss of ancient CG3 cultivars in Anatolia and the surrounding regions throughout history (fig. S24). Even though the muscat aroma is a complex trait, genome-wide association (GWA) analysis based on a binary differentiation reveals 18 SNP signatures on chromosomes 5 and 18 (fig. S24, table S33). This set includes a nonsynonymous SNP Chr5:19419686 in the *VvDXS* gene linked to the trait (*34*). Examination of the genotype at this locus shows that 108 out of the 134 muscat grapevines (including 'Muscat Hamburg', 'Königin der Weingärten', and 'Muscat of Alexandria' commonly used as parental cultivars) are heterozygous (G/T) and only eight individuals are homozygous (T/T) for the alternative SNP (exact test for Hardy-Weinberg Equilibrium, D=20.68, P=2.01×10^{-13}). Additionally, the majority of grapevines without muscat aroma are homozygous for the reference SNP (G/G; 1,451 out of 1,468; exact test for Hardy-Weinberg Equilibrium, D=0.049, P=1.00). This result suggests that selection on this allele might have put constraint on grapevine fecundity, thereby preventing the alternative SNP from reaching fixation.

Berry skin color: novel genes associated with white grapes

The emergence of white grapes from their red-berried congeners is an essential domestication episode in viticulture history. The color change results from a reduction of anthocyanin synthesis in berry skin cells, where the expression of proposed master regulators, such as *VvMybA*, decreased significantly in select cultivars due to either a *Gret1* retrotransposon (*35*), non-conservative exonic mutations (*36*), or large deletions in the locus (*37*). We performed GWA analysis on this large grapevine cohort (fig. S25A and B) and identified multiple significant SNPs across the genome (fig. S25C). The most prominent peak spans a broad genomic region

from 3.51Mb to 16.05 Mb on chromosome 2, overlapping the *VvMybA* locus. Among all significant exonic SNPs in this region (table S34), nonsynonymous SNPs with the smallest *P*-values localize to two uncharacterized genes outside the *VvMybA* locus (fig. S25D), whose putative protein functions are acylaminoacyl-peptidase (*Vvsyl02G000229*) and lysine-specific demethylase (*Vvsyl02G001064*), respectively. These SNPs are overwhelmingly homozygous for the reference allele in white grapes and heterozygous in red grapes (fig. S25E). We validated the SNPs in red-berried *V. sylvestris* accessions to account for possible false positives and confirmed their genotypes as predominant heterozygous (fig. S25E, table S34). In comparison, significant exonic SNPs in *VvMybA* genes [including Chr2:5116947 G/T reported previously in (*36*)] show shared genotypes between white grapes and the *V. sylvestris* accessions (fig. S25E). It is not clear how *Vvsyl02G000229* and *Vvsyl02G001064* might regulate anthocyanin synthesis, but these results demonstrate that exonic mutations in the two genes are better predictors of berry skin colors. Furthermore, the heterozygous SNP states in *V. sylvestris* accessions suggest that the white berry alleles existed in natural wild populations before grapevine domestication.

Discussion

Our systematic genomic survey of *V. sylvestris* and *V. vinifera* accessions paints a defined picture of grapevine evolutionary history, which echoes key events in the history of world climate change and human migration (Fig. 6). The Pleistocene era witnessed the continuous fragmentation of habitats, the decline of effective population size, and the separation of ecotypes for *V. sylvestris*. It is highly likely that modern humans extensively utilized grapevines as an energy source from the late Pleistocene, but the harsh climate was not suited for agriculture (*38*). As the climatic conditions ameliorated at the Pleistocene-Holocene transition, the grapevine with

its relatively stable perennial yield unsurprisingly became one of the earliest candidates for domestication. The dual events underpin the model that plant domestication occurs in large culturally connected areas over a long time (*39*), but the domestication time gap remains between genomic inference and archaeological evidence (table S35, fig. S26-27) (*16*). The diverse SDR haplotypes suggest that an early goal could be the conscious selection (*40*) and propagation of rare naturally-occurring hermaphroditic individuals from the *V. sylvestris* population, because they allow mass plantation without male plants. The selection on phenotype, but not on genotype, also implies that the different hermaphroditic haplotypes were subject to strong genetic drift, supported by the high frequency of H1 and almost extinct H4 in extant cultivars.

The Mesolithic and Neolithic periods also saw the early dispersal and diversification of grapevines where unique ancestries emerged in the Balkans, Iberia, and Western Europe with the help of *V. sylvestris* introgression into CG1. This event mirrors early farmer migration in Europe, consolidating the role of viticulture in forming sedentary societies. A higher level of cultural exchange characterizes the last stage since the Bronze Age, thus the trading of superior grapevine cultivars along trade routes. It is especially evident in the plethora of Italian cultivars with three or more genetic ancestries, but unfortunately poses a challenge to disentangle the genealogical history of each grapevine cultivar (*20*). Lastly, genetic reliable wild grapevines from Central Asia, a region battered by climate change and social instability for the past few millennia, are no longer available to test Vavilov's theory for a diversity center or a hypothetical turnover of grapevine types due to Islam conversion in the region. Paleogenomic data may help to resolve these questions in the future.

References and Notes

1. P. E. McGovern, U. Hartung, V. R. Badler, D. L. Glusker, L. J. Exner, The beginnings of winemaking and viniculture in the ancient Near East and Egypt. *Expedition*. 39, 3–21 (1997).

2. P. This, T. Lacombe, M. R. Thomas, Historical origins and genetic diversity of wine grapes. *Trends Genet*. 22, 511–519 (2006).

3. F. Grassi, G. D. Lorenzis, Back to the Origins: Background and Perspectives of Grapevine Domestication. *Int J Mol Sci*. 22, 4518 (2021).

4. D. Cantu, M. A. Walker, *The Grape Genome* (Springer Nature Switzerland AG, 2019; https://doi.org/10.1007/978-3-030-18601-2).

5. D. Zohary, M. Hopf, E. Weiss, *Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin* (Oxford University Press, Oxford, UK, 2012).

6. S. Myles, A. R. Boyko, C. L. Owens, P. J. Brown, F. Grassi, M. K. Aradhya, B. Prins, A. Reynolds, J.-M. Chia, D. Ware, C. D. Bustamante, E. S. Buckler, Genetic structure and domestication history of the grape. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3530–3535 (2011).

7. Y. Zhou, M. Massonnet, J. S. Sanjak, D. Cantu, B. S. Gaut, Evolutionary genomics of grape (Vitis vinifera ssp. vinifera) domestication. *Proc. Natl. Acad. Sci. U.S.A.* 114, 11715–11720 (2017).

8. Z. Liang, S. Duan, J. Sheng, S. Zhu, X. Ni, J. Shao, C. Liu, P. Nick, F. Du, P. Fan, R. Mao, Y. Zhu, W. Deng, M. Yang, H. Huang, Y. Liu, Y. Ding, X. Liu, J. Jiang, Y. Zhu, S. Li, X. He, W. Chen, Y. Dong, Whole-genome resequencing of 472 Vitis accessions for grapevine diversity and demographic history analyses. *Nat Commun*. 10, 1190 (2019).

9. A. Sivan, O. Rahimi, B. Lavi, M. Salmon-Divon, E. Weiss, E. Drori, S. Hübner, Genomic evidence supports an independent history of Levantine and Eurasian grapevines. *Plants People Planet*. 3, 414–427 (2021).

10. S. Freitas, M. A. Gazda, M. Â. Rebelo, A. J. Muñoz-Pajares, C. Vila-Viçosa, A. Muñoz-Mérida, L. M. Gonçalves, D. Azevedo-Silva, S. Afonso, I. Castro, P. H. Castro, M. Sottomayor, A. Beja-Pereira, J. Tereso, N. Ferrand, E. Gonçalves, A. Martins, M. Carneiro, H. Azevedo, Pervasive hybridization with local wild relatives in Western European grapevine varieties. *Sci Adv*. 7, eabi8584 (2021).

11. G. Magris, I. Jurman, A. Fornasiero, E. Paparelli, R. Schwope, F. Marroni, G. D. Gaspero, M. Morgante, The genomes of 204 Vitis vinifera accessions reveal the origin of European wine grapes. *Nat Commun*. 12, 7240 (2021).

12. S. Riaz, G. D. Lorenzis, D. Velasco, A. Koehmstedt, D. Maghradze, Z. Bobokashvili, M. Musayev, G. Zdunic, V. Laucou, M. A. Walker, O. Failla, J. E. Preece, M. Aradhya, R. Arroyo-Garcia, Genetic diversity analysis of cultivated and wild grapevine (Vitis vinifera L.) accessions around the Mediterranean basin and Central Asia. *Bmc Plant Biol*. 18, 137 (2018).

13. R. Arroyo-García, L. Ruiz-García, L. Bolling, R. Ocete, M. A. López, C. Arnold, A. Ergul, G. Söylemezo˝lu, H. I. Uzun, F. Cabello, J. Ibáñez, M. K. Aradhya, A. Atanassov, I. Atanassov, S. Balint, J. L. Cenis, L. Costantini, S. Gorislavets, M. S. Grando, B. Y. Klein, P. E. McGovern, D. Merdinoglu, I. Pejic, F. Pelsy, N. Primikirios, V. Risovannaya, K. A. Roubelakis-Angelakis, H. Snoussi, P. Sotiri, S. Tamhankar, P. This, L. Troshin, J. M. Malpica, F. Lefort, J. M. Martinez-Zapater, Multiple origins of cultivated grapevine (Vitis vinifera L. ssp. sativa) based on chloroplast DNA polymorphisms. *Mol Ecol*. 15, 3707–3714 (2006).

14. P. McGovern, M. Jalabadze, S. Batiuk, M. P. Callahan, K. E. Smith, G. R. Hall, E. Kvavadze, D. Maghradze, N. Rusishvili, L. Bouby, O. Failla, G. Cola, L. Mariani, E. Boaretto, R. Bacilieri, P. This, N. Wales, D. Lordkipanidze, Early Neolithic wine of Georgia in the South Caucasus. *Proc. Natl. Acad. Sci U.S.A.* 114, E10309–E10318 (2017).

15. J. Ramos-Madrigal, A. K. W. Runge, L. Bouby, T. Lacombe, J. A. S. Castruita, A.-F. Adam-Blondon, I. Figueiral, C. Hallavant, J. M. Martínez-Zapater, C. Schaal, R. Töpfer, B. Petersen, T. Sicheritz-Pontén, P. This, R. Bacilieri, M. T. P. Gilbert, N. Wales, Palaeogenomic insights into the origins of French grapevine diversity. *Nature Plants*. 5, 595–603 (2019).

16. See the supplementary materials.

17. M. J. Roach, D. L. Johnson, J. Bohlmann, H. J. J. van Vuuren, S. J. M. Jones, I. S. Pretorius, S. A. Schmidt, A. R. Borneman, Population sequencing reveals clonal diversity and ancestral inbreeding in the grapevine cultivar Chardonnay. *PLoS Genet*. 14, e1007807 (2018).

18. T. Lacombe, J.-M. Boursiquot, V. Laucou, M. D. Vecchi-Staraz, J.-P. Péros, P. This, Large-scale parentage analysis in an extended set of grapevine cultivars (Vitis vinifera L.). *Theor Appl Genet*. 126, 401–414 (2013).

19. R. Bacilieri, T. Lacombe, L. L. Cunff, M. D. Vecchi-Staraz, V. Laucou, B. Genna, J.-P. Péros, P. This, J.-M. Boursiquot, Genetic structure in cultivated grapevines is linked to geography and human selection. *Bmc Plant Biol*. 13, 25–25 (2013).

20. F. Mercati, G. D. Lorenzis, A. Mauceri, M. Zerbo, L. Brancadoro, C. D'Onofrio, C. Morcia, M. G. Barbagallo, C. Bignami, M. Gardiman, L. de Palma, P. Ruffa, V. Novello, M. Crespan, F. Sunseri, Integrated Bayesian Approaches Shed Light on the Dissemination Routes of the Eurasian Grapevine Germplasm. *Front Plant Sci*. 12, 692661 (2021).

21. R. Hosfield, J. Cole, Early hominins in north-west Europe: A punctuated long chronology? *Quaternary Sci Rev*. 190, 148–160 (2018).

22. A. Timmermann, K.-S. Yun, P. Raia, J. Ruan, A. Mondanaro, E. Zeller, C. Zollikofer, M. P. de León, D. Lemmon, M. Willeit, A. Ganopolski, Climate effects on archaic human habitats and species successions. *Nature*, 1–7 (2022).

23. E. C. Corrick, R. N. Drysdale, J. C. Hellstrom, E. Capron, S. O. Rasmussen, X. Zhang, D. Fleitmann, I. Couchoud, E. Wolff, Synchronous timing of abrupt climate changes during the last glacial period. *Science*. 369, 963–969 (2020).

24. M. Engel, H. Brückner, A. Pint, K. Wellbrock, A. Ginau, P. Voss, M. Grottker, N. Klasen, P. Frenzel, The early Holocene humid period in NW Saudi Arabia – Sediments, microfossils and palaeo-hydrological modelling. *Quatern Int*. 266, 131–141 (2012).

25. C. J. Stevens, C. Murphy, R. Roberts, L. Lucas, F. Silva, D. Q. Fuller, Between China and South Asia: A Middle Asian corridor of crop dispersal and agricultural innovation in the Bronze Age. *Holocene*. 26, 1541–1555 (2016).

26. I. Lazaridis, D. Nadel, G. Rollefson, D. C. Merrett, N. Rohland, S. Mallick, D. Fernandes, M. Novak, B. Gamarra, K. Sirak, S. Connell, K. Stewardson, E. Harney, Q. Fu, G. Gonzalez-Fortes, E. R. Jones, S. A. Roodenberg, G. Lengyel, F. Bocquentin, B. Gasparian, J. M. Monge, M. Gregg, V. Eshed, A.-S. Mizrahi, C. Meiklejohn, F. Gerritsen, L. Bejenaru, M. Blüher, A. Campbell, G. Cavalleri, D. Comas, P. Froguel, E. Gilbert, S. M. Kerr, P. Kovacs, J. Krause, D. McGettigan, M. Merrigan, D. A. Merriwether, S. O'Reilly, M. B. Richards, O. Semino, M. Shamoon-Pour, G. Stefanescu, M. Stumvoll, A. Tönjes, A. Torroni, J. F. Wilson, L. Yengo, N. A. Hovhannisyan, N. Patterson, R. Pinhasi, D. Reich, Genomic insights into the origin of farming in the ancient Near East. *Nature*. 536, 419–424 (2016).

27. C.-C. Wang, S. Reinhold, A. Kalmykov, A. Wissgott, G. Brandt, C. Jeong, O. Cheronet, M. Ferry, E. Harney, D. Keating, S. Mallick, N. Rohland, K. Stewardson, A. R. Kantorovich, V. E. Maslov, V. G. Petrenko, V. R. Erlikh, B. Ch. Atabiev, R. G. Magomedov, P. L. Kohl, K. W. Alt, S. L. Pichler, C. Gerling, H. Meller, B. Vardanyan, L. Yeganyan, A. D. Rezepkin, D. Mariaschk, N. Berezina, J. Gresky, K. Fuchs, C. Knipper, S. Schiffels, E. Balanovska, O. Balanovsky, I. Mathieson, T. Higham, Y. B. Berezin, A. Buzhilova, V. Trifonov, R. Pinhasi, A. B. Belinskij, D. Reich, S. Hansen, J. Krause, W. Haak, Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. *Nat Commun*. 10, 590 (2019).

28. R. Pinhasi, J. Fort, A. J. Ammerman, Tracing the Origin and Spread of Agriculture in Europe. *Plos Biol*. 3, e410 (2005).

29. I. Mathieson, S. A. Roodenberg, C. Posth, A. Szécsényi-Nagy, N. Rohland, S. Mallick, I. Olalde, N. Broomandkhoshbacht, F. Candilio, O. Cheronet, D. Fernandes, M. Ferry, B. Gamarra, G. G. Fortes, W. Haak, E. Harney, E. Jones, D. Keating, B. Krause-Kyora, I. Kucukkalipci, M. Michel, A. Mittnik, K. Nägele, M. Novak6, J. Oppenheimer, N. Patterson, S. Pfrengle, K. Sirak6, K. Stewardson, S. Vai, S. Alexandrov, K. W. Alt1, R. Andreescu, D. Antonović, A. Ash, N. Atanassova, K. Bacvarov, M. B. Gusztáv, H. Bocherens, M. Bolus, A. Boroneanţ, Y.

Boyadzhiev, A. Budnik, J. Burmaz, S. Chohadzhiev, N. J. Conard, R. Cottiaux, M. Čuka, C. Cupillard, D. G. Drucker, N. Elenski, M. Francken, B. Galabova, G. Ganetsovski, B. Gély, T. Hajdu, V. Handzhyiska, K. Harvati, T. Higham, S. Iliev, I. Janković, I. Karavanić, D. J. Kennett, D. Komšo, A. Kozak, D. Labuda, M. Lari, C. Lazar, M. Leppek, K. Leshtakov, D. L. Vetro, D. Los, I. Lozanov, M. Malina, F. Martini, K. McSweeney, H. Meller, M. Menđušić, P. Mirea, V. Moiseyev, V. Petrova, T. D. Price, A. Simalcsik, L. Sineo, M. Šlaus, V. Slavchev, P. Stanev, A. Starović, T. Szeniczey, S. Talamo, M. Teschler-Nicola, C. Thevenet, I. Valchev, F. Valentin, S. Vasilyev, F. Veljanovska, S. Venelinova, E. Veselovskaya, B. Viola, C. Virag, J. Zaninović, S. Zäuner, P. W. Stockhammer, G. Catalano, R. Krauß, D. Caramelli, G. Zariņa, B. Gaydarska, M. Lillie, A. G. Nikitin, I. Potekhina, A. Papathanasiou, D. Borić, C. Bonsall, J. Krause, R. Pinhasi, D. Reich, The Genomic History of Southeastern Europe. *Nature*. 555, 197–203 (2018).

30. R. Fregel, F. L. Méndez, Y. Bokbot, D. Martín-Socas, M. D. Camalich-Massieu, J. Santana, J. Morales, M. C. Ávila-Arcos, P. A. Underhill, B. Shapiro, G. Wojcik, M. Rasmussen, A. E. R. Soares, J. Kapp, A. Sockell, F. J. Rodríguez-Santos, A. Mikdad, A. Trujillo-Mederos, C. D. Bustamante, Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6774–6779 (2018).

31. I. Olalde, S. Mallick, N. Patterson, N. Rohland, V. Villalba-Mouco, M. Silva, K. Dulias, C. J. Edwards, F. Gandini, M. Pala, P. Soares, M. Ferrando-Bernal, N. Adamski, N. Broomandkhoshbacht, O. Cheronet, B. J. Culleton, D. Fernandes, A. M. Lawson, M. Mah, J. Oppenheimer, K. Stewardson, Z. Zhang, J. M. J. Arenas, I. J. T. Moyano, D. C. Salazar-García, P. Castanyer, M. Santos, J. Tremoleda, M. Lozano, P. G. Borja, J. Fernández-Eraso, J. A. Mujika-Alustiza, C. Barroso, F. J. Bermúdez, E. V. Mínguez, J. Burch, N. Coromina, D. Vivó, A. Cebrià, J. M. Fullola, O. García-Puchol, J. I. Morales, F. X. Oms, T. Majó, J. M. Vergès, A. Díaz-Carvajal, I. Ollich-Castanyer, F. J. López-Cachero, A. M. Silva, C. Alonso-Fernández, G. D. de Castro, J. J. Echevarría, A. Moreno-Márquez, G. P. Berlanga, P. Ramos-García, J. Ramos-Muñoz, E. V. Vila, G. A. Arzo, Á. E. Arroyo, K. T. Lillios, J. Mack, J. Velasco-Vázquez, A. Waterman, L. B. de L. Enrich, M. B. Sánchez, B. Agustí, F. Codina, G. de Prado, A. Estalrrich, Á. F. Flores, C. Finlayson, G. Finlayson, S. Finlayson, F. Giles-Guzmán, A. Rosas, V. B. González, G. G. Atiénzar, M. S. H. Pérez, A. Llanos, Y. C. Marco, I. C. Beneyto, D. López-Serrano, M. S. Tormo, A. C. Valera, C. Blasco, C. Liesau, P. Ríos, J. Daura, M. J. de P. Michó, A. A. Diez-Castillo, R. F. Fernández, J. F. Farré, R. Garrido-Pena, V. S. Gonçalves, E. Guerra-Doce, A. M. Herrero-Corral, J. Juan-Cabanilles, D. López-Reyes, S. B. McClure, M. M. Pérez, A. O. Foix, M. S. Borràs, A. C. Sousa, J. M. V. Encinas, D. J. Kennett, M. B. Richards, K. W. Alt, W. Haak, R. Pinhasi, C. Lalueza-Fox, D. Reich, The genomic history of the Iberian Peninsula over the past 8000 years. *Science*. 363, 1230–1234 (2019).

32. S. Brunel, E. A. Bennett, L. Cardin, D. Garraud, H. B. Emam, A. Beylier, B. Boulestin, F. Chenal, E. Ciesielski, F. Convertini, B. Dedet, S. Desbrosse-Degobertiere, S. Desenne, J. Dubouloz, H. Duday, G. Escalon, V. Fabre, E. Gailledrat, M. Gandelin, Y. Gleize, S. Goepfert, J. Guilaine, L. Hachem, M. Ilett, F. Lambach, F. Maziere, B. Perrin, S. Plouin, E. Pinard, I. Praud, I. Richard, V. Riquier, R. Roure, B. Sendra, C. Thevenet, S. Thiol, E. Vauquelin, L. Vergnaud, T. Grange, E.-M. Geigl, M. Pruvost, Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proc. Natl. Acad. Sci. U.S.A.* 117, 12791–12798 (2020).

33. C. Zou, M. Massonnet, A. Minio, S. Patel, V. Llaca, A. Karn, F. Gouker, L. Cadle-Davidson, B. Reisch, A. Fennell, D. Cantu, Q. Sun, J. P. Londo, Multiple independent recombinations led to hermaphroditism in grapevine. *Proc. Natl. Acad. Sci. U.S.A.* 118 (2021), doi:10.1073/pnas.2023548118.

34. F. Emanuelli, J. Battilana, L. Costantini, L. L. Cunff, J.-M. Boursiquot, P. This, M. S. Grando, A candidate gene association study on muscat flavor in grapevine (Vitis vinifera L.). *Bmc Plant Biol*. 10, 241–241 (2010).

35. S. Kobayashi, N. Goto-Yamamoto, H. Hirochika, Retrotransposon-Induced Mutations in Grape Skin Color. *Science*. 304, 982–982 (2004).

36. A. R. Walker, E. Lee, J. Bogs, D. A. J. McDavid, M. R. Thomas, S. P. Robinson, White grapes arose through the mutation of two similar and adjacent regulatory genes. *Plant J*. 49, 772–785 (2007).

37. A. R. Walker, E. Lee, S. P. Robinson, Two new grape cultivars, bud sports of Cabernet Sauvignon bearing pale-coloured berries, are the result of deletion of two regulatory genes of the berry colour locus. *Plant Mol Biol*. 62, 623–635 (2006).

38. P. J. Richerson, R. Boyd, R. L. Bettinger, Was Agriculture Impossible during the Pleistocene but Mandatory during the Holocene? A Climate Change Hypothesis. *Am Antiquity*. 66, 387–411 (2001).

39. R. G. Allaby, C. J. Stevens, L. Kistler, D. Q. Fuller, Emerging evidence of plant domestication as a landscape-level process. *Trends Ecol Evol*. 37, 268–279 (2021).

40. R. S. Meyer, M. D. Purugganan, Evolution of crop species: genetics of domestication and diversification. *Nat Rev Genet.* 14, 840–852 (2013).

Acknowledgments: We thank Frédérique Pelsy, Laurence Garmendia Auckenthaler, Anne-Françoise Adam-Blondon, Christèle Cornier, Pál Kozma, Olivier Bachmann, François Gillet, Jean-Michel Gobat, Sandrine Dedet, Joachim Daumann, Kerstin Huber, Valentina Risovannaya, Alla Polulyah, Bordenave Louis, Maria Lafargue, Goutouly Jean-Pascal, Gagik Melyan, Dorin Ioan Sumedrea, Naqinezhad Alireza, technical staff from EGFV and UEVB, and the Danube-Auen National Park for their assistance in the sample collection and laboratory work. We thank Peter Kupfer, Elisha D. O. Roberson, and Desislava Petkova for their comments.

Funding: Support for this project was provided by

Natural Science Foundation of China grant 32070599 (WC)

Yunnan Agricultural University Research Fund A2032002519 (WC)

China Agriculture Research System of MOF and MARA CARS-29 (SW)

Science Committee at the Ministry of SCS, RA 20APP-4E007 (KM)

Alliance of International Science Organization ANSO-CR-PP-2020-04-A (KM)

Ministerio de Ciencia, Innovación y Universidades and Agencia Estatal de Investigación of Spain RTI2018-094470-R-C21 (RAG)

Predoctoral Fellowship PRE2019-088446 (ARI)

Israel Ministry of Science and Technology 90-23-020-12 (ED)

Fondation Giacomi and Swiss National Science Foundation SNSF 43307 to (CA)

European Regional Fund KK.05.1.1.02.0010 (GZ)

Georgian state budget (LU, KB, TZ)

TUBITAK and Ministry of Agriculture and Forestry of Republic of Türkiye grant 105G078 (AE)

Israel Science Foundation 551/18 (EW)

Author contributions:

Conceptualization: YD, ZL, SW, JS, WC

Sample collection and validation: ZL, KM, MM, SG, GZ, PFB, TL, FR, PN, KB, GDB, ED, GDL, JC, CFP, RAG, CA, AE, ZD, VK, GS, NG, SD, NO, PT, CM, VL, AJ, LU, TZ,

DM, MH, GJ, EK, TD, FG, FM, FS, JED, AMD, DC, GM, TU, CÖ, KK, MX, JL, MZ, LW, SJ, YZ, LS, and SL

Laboratory work: YZ, CM, SW, SL, LT, CW, DL, YP, JL, LY, XL, GX, ZY, BC, YW, PG, MR, OR, ARI, YW, SZ

Data analysis and visualization: SD, QX, XD

Supervision: YD, HY, YZ, SW, JS, WC

Data Interpretation: All authors participated in the interpretation of the data.

Writing: YD, SD, QX, XD, WC wrote the paper with input from all co-authors.

Competing interests: A.J. is the founder and owner of Historische Rebsorten vineyard. All other authors declare no competing interests.

Data and materials availability: The VS-1 genome assembly is available at the China National Center for Bioinformation under the accession number CRA006898. The raw resequencing data are available at the China National Center for Bioinformation under the accession number CRA006917.

Supplementary Materials

Materials and Methods

Supplementary Text

Figs. S1 to S27

Tables S1 to S35

References and Notes.

Fig. 1. Genetic diversity of global core *V. sylvestris* and *V. vinifera* accessions. (A) Geographical locations of the 2,448 core grapevine accessions. **(B)** Principal component analysis according to major viticultural regions. Large square/circle highlights median position. Star shows VS-1 position. **(C)** Isolation-by-distance test of *V. sylvestris* and *V. vinifera* accessions. Linear regression with 95% confidence interval shown. **(D)** ADMIXTURE clustering of the accessions. **(E)** The geographic locations of the accessions in each group. Gray represents minor locations. **(F)** Average proportion of major genetic ancestries in grapevine groups. **(G)** PC2 vs. PC3 projection according to grapevine groups. Syl-W, *V. sylvestris* western ecotype; Syl-E, *V. sylvestris* eastern ecotype; CG, cultivated grapevine.

Fig. 2. The population history of *V. sylvestris* ecotypes. **(A)** Geographic isolation and population separation of *V. sylvestris* ecotypes. Pie charts show mean ancestry proportion at K=8. Same color scheme in Fig. 1B. **(B)** Demographic histories of *V. sylvestris* populations deduced from Stairway Plot 2. Lines: median with 75% and 95% confidence interval. **(C)** Population split times among ecotypes with MSMC2. Red bars, median with 95% confidence interval. **(D)** Ecological niche modeling of the suitable habitats for *V. sylvestris* ecotypes. The color scale shows suitability score.

Fig. 3. Independent domestications of *V. vinifera* in the Near East and the Caucasus. (A) Pairwise fixation index of major grapevine groups. **(B)** Outgroup f_3 statistics biplot measuring genetic similarity. Rotund, *Muscadinia rotundifolia*. Stars mark the f_3 statistics for CG1/CG2. **(C)** Estimated split times among Syl-E1/2 and CG1/2 with MSMC2 (left). Red bars, median with 95% confidence interval. **(D)** Geographic distribution of CG1 and CG2 in relation to the domestication centers. Human dispersal routes shown. **(E)** Shared (sky blue) and unique domestication selective sweep regions (red and dark teal) in *V. vinifera*.

Fig. 4. Stepwise diversification of *V. vinifera* in Europe. Introgression from Syl-W into European *V. vinifera* groups revealed by TreeMix **(A)** and confirmed by D-statistic **(B)**. **(C)** Four population simulation of split times and genetic introgression using Momi2. Median numbers from 100 bootstrap runs. **(D)** Origination of *V. vinifera* groups (CG3-CG6) by the end of Neolithic. Geographic distribution of CG groups shown by color circles. See fig. S24 for details on CG3.

Fig. 5. Selection and evolution of the sex determination region in the core grapevine accessions. **(A)** The sex determination region (SDR) in VS-1. Red arrows indicate identified recombination sites. **(B)** SDR genotypes from associated SNPs reveal five recombination sites (dashed lines) and genotype diversity (right). Major and minor haplotypes shown on the left. **(C)** Distribution of SDR genotypes in the six major grapevine groups. **(D)** Recombination history of all SDR haplotypes. **(E)** Putative dispersal route of the H4 haplotype and the origination of H2 haplotype.

Fig. 6. Schematic graph of grapevine evolutionary history. Key events in the evolutionary history of grapevines are shown side by side with major events in global climate change and human migration.