Plant Molecular Biology Reporter (2025) 43:1588–1604
https://doi.org/10.1007/s11105-025-01573-7

RESEARCH

Genomic Variation in *Vitis Vinifera*, from Extant to Contemporary Varieties

Zeki Kaya[1,2] · Funda Özdemir Değirmenci[3] · Sevgi Yorulmaz[2] · Kıvılcım B. Vural[2] · Christina Luke[4] · Christopher H. Roosevelt[4] · Mehmet Somel[2] · David B. Neale[1]

Received: 5 December 2023 / Accepted: 28 April 2025 / Published online: 9 May 2025

Abstract

Ancient DNA (*a*DNA) of 3500-year-old charred grapevine (*Vitis vinifera*) seed material excavated from the archaeological site of Kaymakçı, Gölmarmara, Türkiye, was subjected to whole genome shotgun sequencing using next-generation sequencing (NGS) technology. After grapevine DNA sequence data was filtered by quality scores and the grapevine DNA file was aligned with the two reference grapevine genomes (Pinot Noir-clone ENTAV115 and PN40024), 44 consensus mapped (CMap) ancient DNA sequences common to both reference genomes were obtained. Majority of these sequences (63.6%) were specific to the nuclear genome while 20.5% and 15.9% were specific to chloroplast and mitochondria genomes, respectively. Thus, 57.1%, 42.8%, and 100% of the CMap DNA sequences were related to previously identified genes in the grapevine nuclear, mitochondria, and chloroplast DNA genomes, respectively. Despite the vegetative propagation of grapevine, substantial base differences between ancient and modern varieties were found; that is, mean SNP (single nucleotide polymorphic) sites per the CMap DNA sequences varied between 1.71 in mitochondria and 2.89 in chloroplast sequences. In general, mean SNPs were low in the CMap DNA sequences annotated to known genes compared to those without any gene assignments. Further analysis of the association of the CMap DNA sequences of nuclear and organelle genomes with the SNP sites pointed out that base changes occurred in modern grapevine varieties compared to ancient grape variety involving genes controlling colour and aroma of grape fruits. To gain further insights about the role of domestication and clonal propagation in grape, future studies dealing with ancient grape DNA need to focus on certain gene regions using whole genome *a*DNA enrichment as the source of template DNA. The results of this study could provide invaluable reference information for identifying the genes or gene regions in such future studies. However, given the scarcity of ancient grape seeds, the degraded nature of ancient DNA, and measures to prevent contamination and bias, the limited CMap sequences likely represent only a small fraction of the ancient grapevine genome. Accordingly, the findings should be interpreted with these limitations in mind.

Keywords Ancient grape · Ancient DNA · *Vitis Vinifera* · Genetic differences · Kaymakçı

Introduction

Eurasian wild grape (*Vitis vinifera* subsp. *sylvestris*) is a dioecious, perennial, woody plant which was naturally and widely distributed in Anatolia before its domestication. Archaeological studies suggested that the domestication of the species occurred in Southern Caucasus about 6000–8000 years ago (Zohary and Spiegel-Roy 1975; McGovern et al. 1996; Myles et al. 2011; Dong et al. 2023). Following domestication of the species, cultivars were dispersed to Southern Anatolia, the Fertile.

Crescent, to Egypt about 5000 years ago, and to Europe about 2800 years ago. Today, there are more than 10,000

✉ Zeki Kaya
kayaz@metu.edu.tr

1 Department of Plant Sciences, University of California, Davis, CA, USA

2 Department of Biological Sciences, Middle East Technical University, Ankara, Turkey

3 Department of Crop Science, Kırşehir Ahi Evran University, Kırşehir, Turkey

4 Department of Archaeology, Koç University, Istanbul, Turkey

domesticated grape cultivars in the world (Focus OIV 2017). Turkey is an important genetic diversity center both for cultivated *Vitis vinifera* and wild *Vitis vinifera* ssp. *sylvestris* (Arroyo-Garcia et al. 2006). Over 1200 cultivars could be found in Turkey. A genetic diversity exploration study dealing with wild and cultivated *Vitis vinifera* from Eastern and Southeastern Turkey revealed that accessions from both maintained high genetic diversity (Karataş et al. 2014). The oldest seeds of domesticated grapes dated around 8,000 before present (BP) were reported in Georgia and Turkey (This et al. 2006) and the oldest wild grape (*Vitis vinifera* ssp. *sylvestris*) seeds dated 8400 years BP were excavated in the Euphrates Valley in Southeastern Turkey (Gökbayrak and Söylemezoğlu 2010).

Recent studies suggest that the Fertile Crescent is not the only region of domestication in the Eastern Mediterranean. Studies of ancestral connections reveal that populations in Western Anatolia show closer relations with those in Europe, rather than communities along the Eastern Mediterranean (Kılınç et al. 2016). The Late Bronze Age in the Western Anatolia, around 1700/1650–1200 before the common era (B.C.E.) plays an important role in Anatolian cultural history. Within this context, the Gediz Valley served as a hub of exchange between the central Anatolian plateau and the Aegean coast. The Marmara Lake basin of the middle Gediz Valley hosted a number of small settlements from the beginning of the Middle Bronze Age (2000–1700/1650 B.C.E.) through the end of the Late Bronze Age (Luke et al. 2015). Among these settlements, the archaeological site of Kaymakçı was the largest and most complex (Roosevelt et al. 2018; Luke and Roosevelt 2017; Roosevelt and Luke 2017).

Excavations at Kaymakçı in the Marmara Lake basin since 2014 have resulted in the discovery of structures likely functioning as grain silos, as demonstrated at many contemporary sites in central and western Anatolia. Analysis of unearthed botanical samples from Kaymakçı point to cultivation of crops such as barley (*Hordeum vulgare* L.), free-threshing wheat (*Triticum aestivum/durum*), emmer wheat (*Triticum turgidum spp. dicoccum* (Schrank) Thell.), and einkorn wheat (*Triticum monococcum L.*) for human consumption and to some extent for animal husbandry (Roosevelt et al. 2018). Abundant evidence of other crops was reported, also, including bitter vetch (*Vicia ervilia*), chickpea (*Cicer arietinum*), lentil (*Lens culinaris* Medik.), and grass pea (*Lathyrus sativus* L.) (Shin et al. 2021; Marston et al. 2021). The only evidence of fruit cultivation is that of grape (*Vitis vinifera*).

The field of ancient DNA (*a*DNA) concerning the molecular study of archaeological and paleontological specimens where DNA can be extracted from a wide range of well-preserved tissues is rapidly advancing with new DNA sequencing techniques (Parducci et al. 2019). To date, studies have been carried out on sample types ranging from seeds (Manen et al. 2003; Cappellini et al. 2010; Çiftci et al. 2019; Özgen et al. 2012; Cohen et al. 2023), sediments (Willerslev et al. 2004; Anderson-Carpenter et al. 2011), and woods (Rogers and Kaya 2006). Ancient specimens generally contain low amounts of endogenous DNA, which consists of short molecules that are highly fragmented and exhibit characteristic DNA damage patterns. However, plant seeds excavated either from archaeological settlements or lake sediments can provide *a*DNA which could be used to explore the past of plant species. *a*DNA will be particularly useful to understand the evolution and domestication of agricultural plants.

In a previous study (Çiftci et al. 2019), free-threshing wheat (*Triticum aestivum*), bitter vetch (*Vicia ervilia*), chickpea (*Cicer arietinum*), and grape (*Vitis vinifera*) were studied using *a*DNA approaches on archaeological materials to characterize the seeds according to true type and how much change had occurred in them compared to contemporary relatives. Although very limited regions of genomes in the *a*DNA of these crops species were studied, such as ribosomal gene 26S, the results indicated that the specific base locations were either lost or substituted with different DNA bases in the contemporary DNAs. Compared to studied crops species, the fewest variations between ancient and modern species were found in *Vitis vinifera* samples, although the changes were base deletions occurring in contemporary grape samples. The effects of continued domestication and breeding activities on these changes need to be studied further.

The archaeological charred grape seeds previously studied (Çiftci et al. 2019) were the first from western Anatolia to be analysed for *a*DNA. *Vitis vinifera* has a small genome size, about 500 Mb (Velasco et al. 2007; Jaillon et al. 2007), which makes genomic study much easier compared to other crops species reported by Çiftci et al. (2019). Furthermore, grape has been propagated clonally in cultivation for a long time, although reproduction by seeds still exists in the wild. This means that the number of generations between extant and current varieties will be much less than those reproduced mainly by seeds (Brown et al. 2015). In fact, the recent study by Cohen et al. (2023) points out long and continuous grape cultivation in the eastern Mediterranean that reveals high similarities of ancient grape cultivars to the modern vine cultivars. The large amount and good quality of genomic DNA obtained from charred ancient seeds in the previous study (Çiftci et al. 2019) could be used for high-throughput sequencing of ancient DNA that can generate abundant genetic data for answering important archaeobotanical and domestication questions. Concerning the selection of the reference genome ENTAV published by Velasco et al. 2007, instead the PN40024 released by Jaillon et al. 2007, the former is preferred due to its accuracy and systematic curation, Future work using a whole genome sequencing approach and tackling multiple, well-characterized gene

regions responsible for agronomic traits in grapes will provide important information about indigenous cultivation in the past and the origins and distribution of grape species.

The objective of the current study was set to determine the extent of genetic changes occurring in the reference genome compared to the genomic sequences obtained from the ancient grape seed material. Here we report the results from comparison of whole genome shotgun sequences of ancient grape seed to the reference genome sequences of Velasco et al. (2007). Due to the limited number of consensus-mapped (CMap) ancient grapevine DNA sequences obtained from a single seed from Kaymakçı, and in light of the inherent challenges associated with ancient DNA research—including the scarcity of ancient seeds, the degraded and low-yield nature of ancient DNA suitable for next-generation sequencing, and stringent protocols to minimize contamination and analytical bias—the data presented here should be regarded as a partial representation of the ancient grapevine genome. Consequently, interpretations drawn from these results should be made with an understanding of their inherent limitations.

Materials and Methods

In ancient DNA studies, the type of material, age, and conditions of study materials and their availability dictate uniqueness and success of intended studies. Thus, a collaboration with archaeologists is important in every step of obtaining authentic and uncontaminated archaeobotanical material from archaeological sites. Such a collaboration with archaeologists from Koç University was initiated at Kaymakçı in 2017. Archaeological work has been carried out at Kaymakçı since 2014 (https://gygaia.org). A limited number of ancient, charred seeds of wheat (*Triticum aestivum*), bitter vetch (*Vicia ervilia*), chickpea (*Cicer arietinum*), and grape (*Vitis vinifera*) dating from about 3500 years ago were provided to us in 2017. It is expected that continuing excavations at Kaymakçı will produce more ancient crop seeds in the future. Ancient DNA from these seeds were successfully extracted. Since ancient DNA molecules were characterized by varying degrees of degradation, whole genomic enrichments of wheat (*Triticum aestivum*), bitter vetch (*Vicia ervilia*), chickpea (*Cicer arietinum*), and grape (*Vitis vinifera*) were carried out previously. DNA extractions from ancient *Vitis vinifera* seeds were conducted using two distinct methods to evaluate the fidelity of these DNA isolation procedures. The first method employed Kistler's modified CTAB protocol (Kistler and Shapiro, 2011), while the second utilized the commercial High Pure PCR Product Purification Kit (Roche), following the manufacturer's instructions (Çiftci et al. 2019; Lister et al. 2008; Oliveira et al. 2012). For comparative analysis, DNA from one ancient *Vitis vinifera* seed was isolated using the CTAB-based method, and DNA from another seed was extracted using the commercial kit (Supplementary material).

DNA isolation and PCR analyses were conducted in a dedicated laboratory specifically designed and reserved for ancient DNA studies. This facility had no prior history of DNA isolation or Polymerase Chain Reaction (PCR) applications on any materials. To ensure a contamination-free environment, the laboratory was meticulously cleaned using 20% bleach, absolute ethanol, and DNA AWAY™ Surface Decontaminant solutions (Molecular Bioproducts, Inc., San Diego, CA).

Prior to DNA extraction, all isolation equipment and necessary chemicals were sterilized using an autoclave. At the commencement of the isolation process, materials were rigorously cleaned with 70% ethanol and DNA AWAY™ Surface Decontaminant solutions to ensure cleanliness. To minimize the risk of contamination from the laboratory environment, all equipment involved in DNA isolation and PCR was subjected to UV treatment. Strict contamination control measures were implemented, including the use of protective clothing, double-gloving, and face masks during ancient DNA procedures. Additionally, regular decontamination steps, such as bleach treatment, UV irradiation, and frequent glove changes, were performed throughout the experimental process.

The concentrations of DNA isolated using the two methods were compared following quantification by spectrophotometry (Biodrop Lite 7141 V.1.0.4, Biological Sciences Department, METU). The analysis indicated that the commercial kit method produced higher DNA concentrations. To ensure the reliability of the extraction process, extraction blanks were included as negative controls for each seed. In the case of grape, only four charred seeds from Kaymakçı were made available. In the current study, ancient DNA from one of ancient seeds having *a*DNA with high quality and yield (Fig. 1, Supplementary material, Table 1), was used for whole genome shot gun sequencing experiment. Although whole genome DNA amplification treatment of *a*DNA from ancient grape seeds produced high quality and yield, it was not considered in the current study due to scarcity of charred grape seeds, contamination, and bias concerns associated with additional steps involved in whole genome DNA amplification treatments (Sabina and Leamon 2015).

Because four ancient seeds were available from *Vitis vinifera*, DNA extraction procedures were optimized by testing various extraction procedures with the charred seeds from the other crop species mentioned above (Uluğ et al. 2025). The High Pure PCR Product Purification Kit (Roche) yielded the best results so this kit was used for *a*DNA extraction from the charred grape seed. The details of *a*DNA extraction, PCR procedures, and authenticity of study materials using targeted gene regions were described in Ciftci et al.

Fig. 1 View of ancient charred *Vitis vinifera* seed (sample no. 4; 99.526.178.2; courtesy of the Kaymakçı Archaeological Project)

(2019). Furthermore, *a*DNAs from all four charred grape seeds were also verified as the *Vitis vinifera* specific (with 99% coverage and 90% similarity) by partial amplification of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) gene (*rbc*L) located in chloroplast DNA (Uluğ et al. 2025).

Considering our previous experiences with ancient animal and human DNA studies, it was decided that the whole genome shotgun approach of NGS technology would be more appropriate to cover the whole genome of ancient grape.

Library Preparation and Initial Sequencing

Since the double index system increases the number of samples that can be sequenced in a single sequencing pool and eliminates the possibility of sample misidentification, these kinds of libraries can be used to screen and authenticate *a*DNA. Thus, double-stranded DNA libraries were prepared using 20 mL of DNA extract using the blunt-end ligation method as described in Myles et al. (2010) and in Gunther et al. (2015). Each library was amplified via PCR in six replicates, each with a 25 mL volume, using specific indexing primers (15 single- and seven double-indexing) (Meyer and Kircher 2010). Negative controls were included in both the library preparation and PCR steps. Each PCR reaction volume had 3 mL DNA library, 1X AmpliTaq Gold Buffer, 2.5 mM $MgCl_2$, 250 nM of each dNTP, 2.5U AmpliTaq Gold (Life Technologies), and 200 nM each of the IS4 primer and an indexed P7 primer (Myles et al. 2010) while a negative control was included in both the library preparation and PCR steps. The PCR cycling steps included 94 °C for 10 min followed by 10–14 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 45 s and a final extension at 72 °C for 10 min. Amplified libraries were pooled and purified with AMPure XP beads (Agencourt). The library quantifications were made with a 2100 Bioanalyzer using the High Sensitivity DNA Kit (Agilent Technologies).

Libraries were pooled in equimolar concentrations (final vol of 10 nM total pool) for the initial sequencing (prescreening) process and sequenced on Illumina HiSeq 2500, HiSeq X, and NovaSeqS4 platforms at SciLife, Stockholm, with 100–150 bp paired-end reads on single or several lanes. Libraries that yielded sufficient reads from the initial screening process were then sequenced deeper in pools of three to six libraries per lane.

Sequencing and Data Analysis

The sequencing data from each library was processed as described by Koptekin et al. (2023). Because *a*DNA molecules are particularly small, most of the pair-mate reads are expected to be merged, the residual adapter sequences in FASTQ file which was received from the sequencing process were subjected to trim of adapters and truncated where they are necessary using the *AdapterRemoval v.2.12*, with the following settings: 'quality base33-gzip-trimns' (Schubert et al. 2016). After removing residual adapter sequences, the pair-end reads were merged, requiring at least 11 bp overlap between the pairs using additional "-collapse-misalignments length 11". The merged reads were mapped to *Vitis vinifera* (PinotNoir, clone ENTAV115, version PN40024.v4) reference genome file (version v2.1; Velasco et al. 2007) as well as *V. vinifera (*PN40024) reference genome (Jaillon et al. 2007) using the program "*BWA aln/samse*" (version 0.7.15; Li and Durban 2009) with the parameters "-n 0.01 -o 2" and by disabling the seed with "-l 16,500" (Kircher 2012). To eliminate PCR duplicates with identical start and end coordinates, "Filter UNiqueSAMCons.py" was used (Kircher 2012). We filtered the reads with greater than 10% mismatches to the *Vitis* reference genomes, less that 35 base pairs. Postmortem deamination patterns were estimated using PMDtools with the "-deamination" parameter. Average genome coverage was calculated including only reads with mapping quality greater than 30 and using "genomeCoverageBed" implemented in "bedtools" (Quinlan and

Table 1 Annotation of 44 CMap DNA sequences of ancient-grapevine

The CMap ancient *Vitis vinifera* sequences	Number of CMap sequences	Size range (bp)	Average (bp)	Number of genes	Annotated gene (gene products)
CMap nuclear DNA sequences	28	60–140	97.96	17	-VITISV_012700 (hypothetical protein/CAN66105.1) -LOC100250340 (crocetin glucosyltransferase, chloroplastic) -LOC100245570(aspartic proteinase nepenthesin-1) VITISV_001694 (hypothetical protein/CAN81735.1) -VITISV_006448 (hypothetical protein/CAN63602.1) -LOC104880017 (hypothetical protein/CAN83945.1) -LOC104880017 (hypothetical protein/CAN83945.1) -LOC100854426 (protein phosphate starvation response 2) -VITISV_033547 (hypothetical protein/CAN79935.1) - LOC109122575 (ncRNA) -LOC100260866 (auxin response factor 6) -VITISV_039748(hypothetical protein/CAN78516.1) -VITISV_014576(hypothetical protein/CAN69963.1) - LOC100262957 (isoform X3 3-dehydroquinate synthase II) -VITISV_013532(hypothetical protein/CAN81964.1) -VITISV_026457(hypothetical protein/CAN67695.1) -LOC100244983(E3 ubiquitin-protein ligase UPL5)
CMap Chloroplast DNA (*cp*DNA) sequences	9	32–157	90.25	9	−16S ribosomal RNA gene, part of -*rpo*C2 -*ndh*B gene (NADH dehydrogenase subunit 2) - *psb*I (Photosystem II reaction center protein I) - *psb*C (photo system 2 C region) - *rpo*C2 (RNA polymerase beta) -*trn*G-GCC (tRNA-Gly) -*rpl*2 (ribosomal protein L2) -*psb*C (photosystem II CP43 chlorophyll apoprotein) -*psa*B (Photosystem II 700 Chlorphyll a Apoprotein A2)
CMap Mitochondrion DNA (*mt*DNA) sequences	7	48–170	93.71	3	-*ccm*FN (cytochrome c biogenesis protein) -loc104878239 & cox3 (cytochrome c oxidase subunit 3/LOC104878239) -LOC100247105 (LRR receptor-like serine/threonine-protein kinase At1 g53420)

Hall 2010). The reads with mapping quality greater than 30 were obtained by using Samtools program (Li et al. 2009). Consensus-mapped (CMap) sequences of ancient grape were obtained by selecting the mapped sequences which were common in both *Vitis vinifera* genomes (Velasco et al. 2007; Jaillon et al. 2007). Subsequently, the CMap sequences specific to nuclear, chloroplast, and mitochondria genomes will be also referred as the CMap ancient grape *n*DNA, CMap *cp*DNA, and CMap *mt*DNA sequences.

Exploration of Consensus-Mapped (CMap) Sequences by Comparing to the Modern Grapevine DNA Sequences and Their Gene Annotations

The file in BAM format, consisting of the CMap sequences were converted to FASTA format for further analysis using the BLAST tools of the NCBI nucleotide database (Sayers et al. 2022). The consensus-mapped sequences were individually subjected to a BLAST search using options of standard database-nucleotide collection (nr/nt), *Vitis* as organism id and highly similar sequences. The sequences retrieved through BLAST search were further selected by the criteria of 100% read coverage and, at least 90% identity scores. Then, each consensus-mapped sequence with high coverage and identities with the nuclear and organelle genomes (chloroplast and mitochondria) of *Vitis vinifera* sequences from the NCBI database were aligned with the BLAST-multiple sequence alignment option (Sayers et al. 2022). The BLAST search and multiple sequence alignments were repeated for each of the CMap sequences. Then, consensus-mapped and *Vitis vinifera* sequences in multiple sequence alignment file were viewed with the MSA viewer of BLAST to determine base changes (potential SNP sites) in modern sequences compared to the CMap sequences. Following the alignment of the CMap DNA sequences with the *Vitis vinifera* accessions from the NCBI database (Sayers et al. 2022), the CMap DNA sequences with potential SNPs were checked if any known gene was associated with it. The results from the analysis of the CMap DNA sequences, if they were associated with known genes or just related to non-coding regions of grapevine genome, were presented separately. Selective CMap sequences of ancient grape genome associated with known genes were further examined with respect to SNP sites and their gene products using the NCBI database search tools (Sayers et al. 2022).

Results

The removal and trimming of adapter sequences, quality checking for minimum quality scores of the raw DNA sequencing yielded 16,178,404 reads. Forty-nine (coverage: 0.0000065%) and 738 (coverage: 0.000015%) of these were mapped to the indexed *Vitis vinifera* (PinotNoir, clone ENTAV115) (version v2.1; Velasco et al. 2007) and *V. vinifera* (PN40024) reference genomes (Jaillon et al. 2007), respectively. The average lengths of the mapped sequences were 93.27 and 91.85 bp for *Vitis vinifera* (PinotNoir) and *V. vinifera* (PN40024) reference genomes, respectively. The forty-four mapped sequences (CMap sequences) were common in both reference genomes. The postmortem deamination patterns scores were estimated as 0.2 for C to T base changes and 0.18 for G to A base changes. For authentication of extracted *a*DNA, post-mortem modifications in extracted charred grape *a*DNA were assessed using the MapDamage2 (Jonsson et al. 2013). The results were provided in Supplementary Fig. 1. Although there are some indications for excess of T bases in 5′ end and G bases in 3′ end of the reads, nevertheless, the postmortem deamination patterns could not be reliably estimated due to low number of the CMap sequences. Among the 44 CMap sequences, there were 28 *n*DNA, 9 *cp*DNA, and 7 *mt*DNA genome specific sequences (Table 1).

Twelve of the 28 the CMap sequences, specific to nuclear genome, had no specific gene and gene product annotation while 16 of these were annotated to particular plant genes and their products (Tables 2, 3). Furthermore, the majority of the 28 CMap sequences maintained at least one SNP (23/28) related to a gene and product, although 7 of the 23 the CMap sequences with SNP sites were not associated with any known gene in the NCBI database (Tables 2, 3, 4). In general, the CMap sequences attributed to known gene and product including hypothetical proteins had lower mean SNP sites (1.66 ± 1.30; 1.660 ± 2.0), sequence coverage (93.33 ± 14.48; 94.16 ± 17.21) and higher identity (98.18 ± 2.06; 98.21 ± 1.35) than those CMap sequences which were not attributed to any gene or product (mean SNP site = 2.33 ± 1.87; mean coverage = 99.9 ± 0.32; mean identity = 96.44 ± 4.13) (Table 5).

When the CMap sequences were subjected to DNA blast search using the NCBI database, 9 of 44 CMap ancient grapevine DNA sequences showed high coverage and identity with the *Vitis vinifera* chloroplast genome. Four of 9 chloroplast genome associated the CMap sequences had no SNP sites, while the remaining ones had a variable number of sites, varying from one to four SNP sites. Based on multiple sequence alignment and annotated genes in the NCBI database, all nine CMap *cp*DNA sequences were found to be associated with known genes and gene products (Table 6). The CMap cpDNA sequences with known gene and products had mean SNP sites of 2.89 ± 6.51 with 97.5% mean sequence coverage and 94.89% mean identity with the grapevine *cp*DNA accessions available from the NCBI database (Table 5).

There were also seven CMap ancient grape *mt*DNA sequences. Five of the sequences were not related to known genes or gene product from the NCBI database; but two of them were recognized with known genes. Also, four of the CMap ancient grape *mt*DNA sequences presented three SNP sites (Table 7). The CMap ancient grape *mt*DNA sequences attributed to known mitochondrial gene and products had higher mean SNP sites (3.0 ± 0), lower sequence coverage (97.5 ± 3.54) and higher identity (94.89 ± 21.99) compared to those CMap *mt* DNA sequences which were not attributed to any gene or products (mean

Table 2 Variable sites in CMap sequences specific to the *Vitis vinifera* nuclear genome annotated according to known genes

The codes of the CMap ancient *Vitis vinifera* sequences	Modern *Vitis vinifera* sequences (NCBI codes)	SNP sites	Coverage for the CMap sequence (%)	e- value	Sequence identity of the CMap ancient sequence (%)	Base change from The CMap ancient sequences to modern	Gene (annotated)	Product of gene
ZK-AVV7	XM 0022667644 VV [b] + 6 more sequences	6	100	2e-55	94.3	G > A; C > T; 5(A > G); T > C; C > T; A > G; C > T; A > C	LOC100250340	Crocetin glucosyltransferase
ZK-AVV9	XM_002272766.3[b]	1	99	5e-64	99.3	G > A	LOC100245570	Aspartic proteinase nepenthesin-1
ZK-AVV12	XM_010655096.2[b]	1	100	1e-43	99.0	C > A; C > A	LOC104880017	None
ZK-AVV14	XM_019221549.1[b]	1	58	7e-14	97.8	G > A	LOC100854426	Protein phosphate starvation response 2
ZK-AVV19	XR_002030056.1[b]	1	100	2e-61	99.2	G > A	LOC109122575	ncRNA
ZK-AVV20	AM466203.1[a] XM_019223627.1[b]	none 1	100 83	3e-45 2e-34	100.0 98.8	none T > A	none LOC100260866	None auxin response factor 6
ZK-AVV25	*AM432113.1*[a] XM_002282954.4[b]	4	100	8e-32	95.2	C > T;T > C, T > C;C > T	none LOC100262957	None uncharacterized protein LOC100262957 isoform X3 (3-dehydroquinate synthase II)
ZK-AVV37	XM_002272023.4[b]	0	100	8e-32	100.0	None	LOC100244983	E3 ubiquitin-protein ligase UPL5

[a]Velasco et al. (2007), [b]Venturini et al. (2013)

SNP site = 1.20 ± 1.64; mean coverage = 99.9 ± 2.24; mean identity = 97.91 ± 2.87) (Table 5).

SNP polymorphisms in CMap ancient grape DNA sequences specific to nuclear, chloroplast, and mitochondrial genomes were further examined on selected sequences associated with known genes. Among the CMap *n*DNA sequences of the ancient grape, those coded as ZK-AVV7, ZK-AVV9, and ZK-AVV14 with 140, 140, and 79 bp in length, respectively, were selected. The CMap *n*DNA sequences of ZK-AVV7, ZK-AVV9, and ZK-AVV14 which were attributed to the genes with the products of *crocetin glucosyltransferase*, *aspartic proteinase nepenthesin-1*, and *protein phosphate starvation response 2* had 12, 1, and 1 SNP sites, respectively. The aspartic proteinase is member of a large gene family. The aspartic proteinase nepenthesin-1 is an important enzyme in digesting insect proteins by carnivorous plants; this enzyme has hydrolysing function of peptide bonds (Takahashi et al. 2008). The DHR2 gene involved in *protein phosphate starvation response 2* and had transcription and transcription regulation function in plants (Zhang et al. 2024). Interestingly, the *crocetin glucosyltransferase* functions in the synthesis of apoprotenoids which are responsible for the colour and bitter taste of saffron from *Crocus sativus* (Moraga et al. 2004). In the CMap ancient grapevine DNA sequences, the G base at the position of 107 th in ZK-AVV9 was changed to A while the A base at the 5 th position of ZK-AVV14 was modified as G. In the CMap ZK-AVV7 DNA sequence, the G base at the 10 th, C at the15th, A at the 32nd, 34 th, 38 th, and 56 th, G at the 84 th, T at the 88 th, C at the 97 th, at the 98 th, C at the 111 th, and A at the 137 th were altered as A, T, G, G, G, G, A, C, T, G, T, and C in modern varieties, respectively (Table 8).

All CMap grapevine DNA sequences belonging to the chloroplast genome were associated with known genes from the NCBI database. Two CMap *cp*DNA sequences, namely ZK-AVV16 of 25 bp and ZK-AVV17 of 157 bp in length, were selected from the 9 CMap grapevine *cp*DNA sequences. The CMap *cp*DNA sequences coded as ZK-AV16 and ZK-AVV17, related to *rrn16* and *ycf1* genes, had 3 and 2 SNP sites, respectively. The *rrn16* gene codes for the

Table 3 Variable sites in CMap sequences specific to *Vitis vinifera* nuclear genome associated with the loci coding for hypothetical proteins

The codes of the CMap ancient *Vitis vinifera* sequences	Modern *Vitis vinifera* sequences (NCBI codes)	SNP polymorphic sites	Coverage for the CMap ancient Sequences (%)	e- value	Sequence identity of the CMap ancient sequence (%)	Base change from the CMap ancient sequences to modern accessions	Attributes gene	Gene product
ZK-AVV1	AM474208.1[a]	3	99	2e-40	97.0	**A** > T; **A** > G; **C** > G	VITISV_012700	hypothetical protein/ CAN66105.1
ZK-AVV10	AM432627.2[a]	3	100	7e-41	97.0	C > T; A > G; A > G	VITISV_001694	hypothetical protein/ CAN81735.1
ZK-AVV11	AM451670.1[a]	4	100	7e-41	97.0	CCC > TTT C > T; G > A; A > G	VITISV_006448	hypothetical protein/ CAN63602.1
ZK-AVV18	AM423468.2[a] AM475788.2[a]	6	100 100	3e-18 3e-18	89.0 88.7	A > C; G > A; A > G; C > A; C > T; A > T; T > C	VITISV_033547	hypothetical protein/ CAN79935.1
ZK-AVV21	AM442408.2[a]	1	100	7e-33	98.7	C > A	VITISV_039748	hypothetical protein/ CAN78516.1
ZK-AVV23	AM473838.2[a]AM453184.2[a]	1 1 insert	100 100	4e-44 4e-43	99.0 99.0	T > C A(63 and 64)	*none* VITISV_014576	*none* hypothetical protein/ CAN69963.1
ZK-AVV26	AM487161.2[a]	1	100	1e-43	99.01	A > G	VITISV_013532	hypothetical protein/ CAN81964.1
ZK-AVV36	AM451884.2[a]	0	100	2e-23	100.0	None	VITISV_026457	hypothetical protein/ CAN67695.1

[a]Velasco et al. (2007)

ribosomal 16S rRNA while the *ycf1* is with unknown function, although it is in the open-reading frames gene group of the chloroplast genome (Dong et al. 2015). In the CMap *cp*DNA sequence of ZK-AVV16, the T base at the 6 th, C at the 15 th, and A at the 16 th base positions were changed to G, C, and T in modern grape vine accessions, respectively. The CMap *cp*DNA sequences coded as ZK-AVV17 had base substitutions at the 113 th and 145 th base positions where A and G in the CMap *cp*DNA sequences were replaced with G and T in the modern accessions, respectively (Table 9).

In the CMap ancient grapevine *mt*DNA sequences, there were only two sequences associated with known genes in the NCBI database. These were the CMap ancient grapevine DNA sequences coded as ZK-AVV38 and ZK-AVV39, of 46 and 81 bp in length and related to the genes involved in *cytochrome c biogenesis* (ccmFn) and *cytochrome c oxidase subunit 3* (cox3), respectively. The *ccmFn* involved in the cytochrome c biogenesis has the *heme transmembrane transporter activity* and functions in assembly of cytochrome complex while *cytochrome c oxidase subunit 3* functions as a *cytochrome c oxidase activity* and has an important role in aerobic and cellular respirations in plants (Mansilla et al. 2018). Both CMap ancient *mt*DNA sequences had three SNP sites. In ZK-AVV38, the bases of A at the 7 th, G at the 25 th, and C at the 29 th base positions were replaced with G, A, and A bases in modern accession, respectively. In the CMap ancient *mt*DNA sequences of ZK-AVV39, the bases of C at the 39 th, T at the 49 th, and T at the 60 th base positions were changed to A, G, and C bases in modern grape vine *mt*DNA sequences, respectively (Table 10).

Table 4 Variable sites in the CMap Sequences specific to *Vitis vinifera* nuclear genome not associated with any known genes

The codes of the CMap ancient *Vitis vinifera* sequences	Modern *Vitis vinifera* sequences (NCBI codes) Modern Vitis vinifera sequences (NCBI codes)	SNP polymorphic sites (number)	Coverage for the CMap ancient Sequences (%)	e- value	Sequence identity of the CMap ancient sequence (%)	Base change from the CMap ancient sequences to modern accessions	Attributes gene	Gene product
ZK-AVV2	AM438689.1[a]	0	100	4e-28	100	None	none	none
ZK-AVV3	AM42642.1[a] *AM454124.1*[a]	2 0	98 100	1e-57 2e-61	98.7 100.0	A > T; T > A *None*	none	none
ZK-AVV4	AM474155.2[a]	3	100	5e-42	97.0	G > T; C > T T > C	none	none
ZK-AVV13	AM427957.2[a]	4	100	3e-52	96.8	T > C; G > A T > A; G > A	none	none
ZK-AVV15	AM438003.1[a] AM466476.2	2	92 92	3e-38, 4e-37	97.8 96.8	A > C; C > T	none	none
ZK-AVV24	AM450786.2[a]	1	40	0.011	96.0	G > T	none	none
ZK-AVV27	AM481249.2[a]	0	100	5e-34	100.0	None	none	none
ZK-AVV28	AM487704.1[a]	0	100	3e-45	100.0	None	none	none
ZK-AVV29	AM487704.1[a]	3	100	5e-42	97.0	G > T; C > T T > C	none	none
ZK-AVV34	AM432265.2[a]	2	100	1e-43	98.1	A > G;A > G	none	none
ZK-AVV35	The same as ZK34-AVV34	2	100	1e-43	98.1			
ZK-AVV45	AM480767.1[a]	1	100	1e-43	99.0	G > A	none	none

[a]Velasco et al. (2007)

Discussion

The study yielded a low number of CMap ancient grape DNA sequences which were aligned with high coverage and identity scores with the grapevine reference genome. The reasons for this result are low quality (fragmentation of and damages to DNA) and amount of ancient DNA extracted from a charred grape seed. When the filtering of whole genome shotgun DNA sequences by restricting the sequences greater than 35 bp in length and less than 10% mismatches, and imposed mapping parameters to map with the reference genomes, the great majority of ancient DNA reads were eliminated from the analysis. It was clear from the results of our previous studies (Çiftci et al. 2019; Değirmenci et al. 2022; Uluğ et al. 2025) that the extracted ancient DNA needed to be subjected to whole genome DNA amplification treatment to provide high quantity and quality *a*DNAs for a better amplification of ancient DNA regions in PCRs or next-generation sequencing experiment. If the whole genome shotgun sequencing experiment was carried out with the whole genome DNA amplified ancient grape DNA genome, a greater coverage of ancient grape genome would be explored in comparisons with contemporary genome sequences of grapevine available from previous studies (Velasco et al. 2007; Jaillon et al. 2007; Guo et al. 2020; Venturini et al. 2013; Goremykin et al. 2009). The other issue working with archaeological materials is limitation and quality of available ancient seeds. As it was the case in the present study, only four seeds were unearthed from the Kaymakçı archaeological site. Three of the seeds were used in the previous study (Çiftci et al. 2019) while one of the seeds, with high DNA concentration and quality, was used in the present study. Furthermore, imposed restrictions and regulations on archaeobotanical materials prevented analyses on additional charred seed material. This situation prevented us from repeating the whole genome sequencing with additional ancient seed samples.

Considering the number of CMap ancient grapevine DNA sequences generated from a charred grape seed excavated from Kaymakçı, the results should be treated as an outcome from subsampling of the ancient grapevine variety genome and interpreted that it has limitations. We believe that the study provides important information on genetic changes occurred in contemporary grape varieties compared to the ancient variety. The CMap ancient grapevine DNA sequences were a fair sample of nuclear (63.6% of reads), chloroplast (20.5%), and mitochondrial (15.9%) genomes. With respect to coding vs noncoding regions of nuclear

Table 5 Single nucleotide polymorphic sites (SNP) of the CMap ancient *Vitis vinifera* DNA sequences compared with modern *Vitis vinifera* sequences submitted to the NCBI database

Sample size (number of the CMap ancient grape DNA sequences)	Mean SNP sites ± SDEV	Range for SNP sites	Mean coverage for the ancient sequences (%) ± SDEV	Mean e- value ± SDEV	Mean sequence identity with ancient sequences (%) ± SDEV	Minimum SNP site	Maximum SNP site
The CMap ancient grape nuclear DNA sequences attributed to known gene and gene product (Venturini et al. 2013)							
8	1.660 ± 2.0	0–6	93.33 ± 14.48	7.78E-15 ± 2.33e-14	98.18 ± 2.06	0 (LOC100244983/ E3 ubiquitin-protein ligase UPL5)	6 (LOC100250340/ crocetin glucosyltransferase)
The CMap ancient grape nuclear DNA sequences attributed to a gene and a hypothetical protein (Velasco et al. 2007)							
8	1.66 ± 1.30	0–6	94.16 ± 17.21	0.0009 ± 0.0032	98.21 ± 1.35	0 (VITISV_026457/hypothetical protein CAN67695.1)	6 (VITISV_033547/hypothetical protein CAN79935.1)
The CMap ancient grape nuclear DNA sequences not attributed to a gene and product (Velasco et al. 2007)							
12	2.33 ± 1.87	0–4	99.9 ± 0.32	6.0e-19 ± 1.26e-18	96.44 ± 4.13	0	4
The CMap ancient grape CpDNA sequences attributed to a gene and gene product							
9	2.89 ± 6.51	0–20	99.0 ± 3.0	7.4e-07 ± 1.98e-06	95.67 ± 8.23	0 (*psb*C/photosystem II CP43 chlorophyll apoprotein)	20 (psaB/Photosystem II 700 Chlorophyll a Apoprotein A2)
The CMap ancient grape mtDNA sequences attributed to a gene and gene product							
2	3 ± 0	3	97.5 ± 3.53	5e-12 ± 7.07e-12	94.89 ± 1.99	3 (*ccm*FN/ccmFN); (*cox*3/ cytochrome c oxidase subunit 3)	3 (ccmFN/ccmFN); (cox3/ cytochrome c oxidase subunit 3)
The CMap ancient grape mtDNA sequences not attributed to a gene and gene product							
5	1.2 ± 1.64	0–3	99 ± 2.24	4e-18 ± 8.93e-18	97.91 ± 2.87	0	3

genome, 57.1% of the CMap ancient *n*DNA sequences were related to previously identified genes in the grape genome. With regard to organelle genomes, 42.8% of the CMap ancient grapevine DNA sequences were associated with mitochondrial genes while all the CMap *cp*DNA sequences were linked to known genes. The CMap ancient *n*DNA, *cp*DNA, and *mt*DNA sequences showed variation when they were compared to similar contemporary grapevine variety DNA sequences. On average, 1.86 ± 1.76 SNP site per CMap ancient grape *n*DNA sequences, the mean SNPs were 2.89 ± 6.51, and 1.71 ± 1.61 for the CMap *cp*DNA and *mt*DNA sequences, respectively. However, the mean SNPs were low for the CMap ancient grape DNA sequences associated with genes except for the CMap *cp*DNA sequences which were found to be related with chloroplast genes. The high SNP sites in the CMap ancient grape *cp*DNA sequences could be explained with some sequences available from the NCBI database belonging to the *Vitis vinifera* subsp. *sylvestris* from Georgia (unpublished data) and large germplasm collections from China (Guo et al. 2020). In fact, only one CMap ancient grape cpDNA sequence with 20 SNPs caused to estimate high average SNPs. Otherwise, a trend of low SNP sites in the CMap ancient DNA sequences with ascribed genes was observed. These results support previous findings that there are large numbers of structural variations and SNPs reported in various grape varieties (Velasco et al. 2007; Di Genova et al. 2014; Da Silva et al. 2013) compared to the reference grapevine genome (Jaillon et al. 2007). For instance, significant inter-cultivar genetic variations were reported by Venturini et al. (2007). In their study, a local grape cultivar (Corvina) and the PN40024 reference genome were comparatively analysed, finding that 2321 potentially novel protein-coding genes resided in non-annotated or unassembled regions of the reference genome (Velasco et al. 2007). The presence of many single nucleotide variations in CMap ancient grape *n*DNA sequences could be due to domestication processes and the presence of a considerable percent of introns (12.4%). These introns had transposable

Table 6 Variable sites in the CMap ancient *Vitis vinifera cp*DNA sequences compared with modern *Vitis vinifera* sequences submitted to the NCBI database

The codes of the CMap ancient *Vitis vinifera* sequences	Modern *Vitis vinifera* sequences (NCBI codes)	SNP polymorphic sites (number)	Coverage for the CMap ancient Sequences (%)	e- value	Sequence identity of the CMap ancient sequence (%)	Base change from the CMap ancient sequences to modern accessions	Gene (annotated)	Gene product
ZK-AVV5	AM429143.2[a] AB856290.1[b]	1	100	3e-45	99	T insert btw 30 & 31 none	None *rpo*C2	RNA polymerase beta
ZK-AV**V6**	MN561034.1[c] + many other	0	100	2e-62	100	0	*t*RNA	tRNA-Glu
ZK-AVV8	MN561034.1[c]	0	100	9e-31	100	0	*psb*I	Photosystem II reaction center protein I
ZK-AVV16	*MN561034.1*[c] *MW592542.1*[c]	3	100	6e-06	91.43	A > G;T > C;C > T	*rrn*16	16S ribosomal RNA
ZK-AVV17	MN561034.1[c] MW592542.1[c]	2	100	1e-72	98.73	A > G; G > T	*ycf*1	hypothetical chloroplast RF1
ZK-AVV30	MN561034.1c	1	91	5e-07	96.88	T > C	*trn*G-GCC	tRNA-Gly
ZK-AVV31	MN561034.1[c] AM463477.2[a]	0 3	100 98	8e-32 4e-23	100 −92	C > A; A > T G > A; C > T	*rpl*2 None	ribosomal protein L2 -none
ZK-AVV33	MN561034.1[c] LC510289.1	0 0	100% 100	9e-53 9e-53	100 100	None	None *psb*C	photosystem II CP43 chlorophyll apoprotein
ZK-AVV46	LC510289.1[d]	0 20	100	2e-34 2e-07	100 75	None See the alignment	*psa*B	psaB (photosystem II 700 chlorphyll a apoprotein A2)

[a]Velasco et al. (2007), [b]BeridzeT. and TabidzeV (unpublished), [c]Guo et al. (2020), [d]Goremykin et al. (2009)

elements (TE) which were long interspersed retrotransposons (75% of the TE) (Jaillon et al. 2007) that TEs might have responsible for finding high SNPs in noncoding ancient grape DNA reads.

The three genes associated with the CMap ancient grape *n*DNA sequences with SNP sites, namely of *crocetin glucosyltransferase*, *aspartic proteinase nepenthesin-1*, and *protein phosphate starvation response 2*, were further examined with regard to their possible functions in grapevines. The *protein phosphate starvation response 2* (*DHR2*) is found in cytoplasm and expressed roots, stems, leaves, and fruits, participating in transcription. It functions as transcription and transcription regulations. However, the remaining two genes and their functions are interesting since they are related to fruit colour and aroma. *Aspartic proteinase nepenthesin-1* is active extracellularly with aspartic type endopeptidase activity. Remarkably, the gene for *aspartic proteinase nepenthesin-1* was linked to berry colours in grape (Guo et al. 2020). It is likely that the gene sequence of aspartic proteinase nepenthesin-1 in the ancient grapevine was changed in modern varieties to improve fruit colour and in turn, anthocyanin content. Similarly, the gene for *crocetin glucosyltransferase*, which was described in *Crocus sativus* (saffron), has the function of carotenoid synthesis. The enzyme of *crocetin* glucosyltransferase is involved in synthesis of apocarotenoids responsible for the colour and bitter taste of saffron. The gene and its products are most likely to have some roles in colour and aroma of vine grape fruits as well as in providing resistance to fungal infection (He et al. 2018). This is likely why the base change in modern grape vine gene sequences occurred compared to the CMap sequences of ancient grape gene. A comprehensive study dealing with large number of grape accessions by Magris et al. (2021) supports this assumption. They reported that certain genomic regions in cultivated varieties had a reduced genetic diversity and high homozygosity. In particular, three regions of grape genome with reduced genetic diversity involved in two candidate genes that gained berry-specific

Table 7 Variable sites in the CMap ancient *Vitis vinifera mt*DNA sequences compared with modern *Vitis vinifera* sequences submitted to the NCBI database

The codes of the CMap ancient *Vitis vinifera* sequences	Modern *Vitis vinifera* sequences (NCBI codes)	SNP polymorphic sites (number)	Coverage for the CMap ancient Sequences (%)	e- value	Sequence identity of the CMap ancient sequence (%)	Base change from the CMap ancient sequences to modern accessions	Gene (annotated)	Gene product
ZK-AVV22	FM179380.1[a]	0	100	2e-82	100	None	None	None
ZK-AVV32	GQ220325.1[b] AM427945.2[c]	0 3	100 95	1e-43 6e-35	100 94.79	None A > T; G > A C > A; Between 45&46 CT in AM427945.2 instead T	None	None
ZK-AVV38	GQ220324.1[b] FM179380.1[a] AM446904.1[c]	3	95	1e-11	93.48	A > G; G > A, C > A	*ccm*FN	ccmFN
ZK-AVV39	*XM_019218225.1*[d], GQ220323.1[b]	3	100	4e-30	96.3	C > A; T > A T > C	*none* loc104878239;*cox*3	*none* low quality protein; uncharacterized protein loc104878239 cytochrome c oxidase subunit 3
ZK-AVV40	GQ220324.1[b] AM470652.2[c]	3	100	2e-17	94.74	T > C; T > G, T > A	None	None
ZK-AVV41	GQ220324.1[b]	0	100	3e-45	100	None	None	None
ZK-AVV42	GQ220324.1[b]	0	100	3e-45	100	None	None	None

[a]Goremykin et al. (2009); [b]Picardi et al. (2010); [c]Velasco et al. (2007); [d]Venturini et al.(2013)

Table 8 Single nucleotide sites of the CMap ancient grape DNA sequences of genes involved in crocetin glucosyltransferase (A) aspartic proteinase nepenthesin-1 (B) and protein phosphate starvation response 2 (C). After giving the first and last bases of the CMap ancient DNA sequences, the only base positions with SNP sites were presented in the table. ZK-AVV prefix sequences represent the CMap ancient grape DNA sequences from this study

Nuclear genome/gene:LOC100250340 (*crocetin glucosyltransferase*)															
Base positions → DNA sequence names↓	1	10	15	32	34	38	56	84	88	97	98	111	137	140	
ZK-AVV7	**C**	**G**	**C**	**A**	**A**	**A**	**A**	**G**	**T**	**C**	**A**	**C**	**A**	**C**	
XM 0022667644 VV		A	T	G	G	G	G	A	C	T	G	T	C		
FQ3805331 VV		A	T	G	G	G	G	A	C	T	G	T	C		
FQ3841261 VV		A	T	G	G	G	G	A	.C	T	G	T	C		
FQ3838231 VV		A	T	G	G	G	G	A	C	T	G	T	C		
AM437128 VV		A	T	G	G	G	G	A	C	T	G	T	C		
XM 0022665924 VVM		A	T	G	G	G	G	A	C	T	G	T	C		
Nuclear genome/gene: LOC100245570 (*aspartic proteinase nepenthesin-1*)															
Base positions → DNA sequence names↓	1	6	39	56	59	60	70	77	78	89	92	107	112	128	140
ZK-AVV9	**T**	**A**	**T**	**T**	**G**	**A**	**T**	**A**	**C**	**G**	**T**	**G**	**G**	**T**	**T**
refXM 0022727663												A			
AM4578162												A			
XM 0022640203			A	C		T			T	C	C	A	C		
AM4716552		T	A	C		T	C		T	C		A	C	C	
Nuclear genome/gene: LOC100854426 (product-protein phosphate starvation response 2)															
Base positions → DNA sequence names↓	1	2	3	4	5	6	35	78	79						
ZK-AVV14	**A**	**T**	**G**	**G**	**A**	**C**	**T**	**A**	**T**						
AM424652.2					G										
XM 019221549.1					G										

expression in domesticated varieties contributed to the change in berry size and morphology.

The CMap ancient grape *cp*DNA sequences which were examined in detail with regard to their associations with the known genes were part of *rrn16* and *ycf1* genes. The *rrn16* gene codes for ribosomal 16S *r*RNA and expresses in young cotyledons, while the *ycf1* gene is with unknown function, but it is a member of an open reading frame gene group in the grape chloroplast genome (Zhao et al. 2015).

The two CMap ancient grapevine *mt*DNA sequences were related to the *ccmFN* (cytochrome c biogenesis) and *cox3* gene (cytochrome c oxidase subunit 3). The product of the *ccmFN* gene functions in *heme transmembrane transporter activity* and cytochrome complex assembly while the product of *cox3*, which is the component of *cytochrome C oxidase-complex IV*, is involved in mitochondrial electron transport that is cytochrome C to oxygen. The COX of higher plants has over 10 ten subunits (Jansch et al. 1996). The three largest subunits, namely COX1, COX2, and COX3, are encoded by the mitochondrial and the others are encoded by the nuclear genome (Ohtsu et al. 2001).

The base substitutions occurring in modern grape varieties compared to this ancient grape variety certainly changed the gene products, but data from the current study could not address the outcome of these changes. Studies of modern grape varieties or cultivars report the presence of additional genes and diversity compared to the reference genome PN40024 (Jaillon et al. 2007). Da Silva et al. (2013), investigated RNA sequence data from *Vitis vinifera* cultivar "Tannat" for genome reannotation by using a combination of reference-guided annotation and de novo transcript assembly, were able to annotate 5901 previously unannotated or unassembled genes and to discover 873 new genes that were not shared with PN40024. A great proportion of the cultivar-specific genes (about 81%) were involved in the synthesis of phenolic and polyphenolic compounds that contribute additional genes and diversity compared to reference genome. They suggested that the grapevine reference genome lacks many genes that might be relevant for the varietal phenotype.

Another study by Zhou et al. (2017) also suggested that domestication reduced the effective population size rapidly, because of clonal propagation and perenniality of grapevine. In their study, it was found that accessions contained 5.2% more deleterious variants than wild individuals, and they were more often in a heterozygous state. In this way, clonally propagated grape vine varieties accumulated recessive deleterious mutations without decreasing fitness. These findings point out why there were high base differences between the CMap ancient grapevine DNA sequences and reference genome PN40024.

The previous reports suggest that grape varieties are highly heterozygous and harbor great polymorphism 4

Table 9 Single nucleotide sites of ancient grape DNA sequences of genes involved in rrn16 (A) and ycf1 (B). After giving the first and last bases of the ancient reads, the only base positions with SNP sites were presented in the table. ZK-AVV prefix sequences represent the CMap ancient grape DNA sequences from this study

Chloroplast/gene: *rrn*16						
Base positions → DNA sequence names↓	1	6	15	16	35	
ZK-AVV16	**A**	**A**	**T**	**C**	**A**	
MW592542.1-vv		G	C	T		
MW592542.2-vv		G	C	T		
FQ393717.1-vv		G	C	T		
AM435039.2-vv		G	C	T		
AM483961.2-vv		G	C	T		
MN561034.1		G	C	T		
MN561034.1		G	C	T		
LC510289.1		G	C	T		
LC508115.1		G	C	T		
LC507100.1		G	C	T		
LC495883.1		G	C	T		
LC495478.1		G	C	T		
LC494572.1		G	C	T		
LC492108.1		G	C	T		
MG586833.1		G	C	T		
LC722365.1		G	C	T		
LC687362.1		G	C	T		
AB856291.1		G	C	T		
AB856290.1		G	C	T		
AB856289.1		G	C	T		
Chloroplast/gene: *ycf*1						
Base positions → DNA sequence names↓	1	113	114	144	145	157
ZK17-AVV17	**C**	**A**	**C**	**T**	**G**	**A**
MN5610341		G			T	
LC5102891		G			T	
LC5081151		G			T	
LC5071001		G			T	
LC4958831		G			T	
LC4954781		G			T	
LC4945721		G			T	
LC4921081		G			T	
MG5868331		G			T	
LC7223651		G			T	
MW5925421		G			T	
LC6873621		G			T	
AB8562911		G			T	
AB8562901		G			T	

polymorphisms per kilobase (Velasco et al. 2007). Furthermore, Corso et al. (2015) report that many novel genes existing in ancient varieties were lost in modern varieties.

Although the results of present study have limitations for drawing concrete conclusions about genetic changes occurring from the ancient grape genome to modern variety genomes, the results of analyses of the ancient grape genome from Kaymakçı reveal important information about the evolution and domestication of the grapevine.

Future studies with ancient grape seeds or tissues should focus on specific gene regions of the ancient grape variety genome and compare with a more refined or more comprehensive reference genome including representative genomes currently available. For this kind of study, a

Table 10 Single nucleotide sites of ancient grape DNA sequences of genes involved in ccmFN gene (cytochrome c biogenesis) (A) and cox3 gene (cytochrome c oxidase subunit 3) (B). After giving the first and last bases of the ancient reads, the only base positions with SNP sites were presented in the table. ZK-AVV prefix sequences represent the CMap ancient grape DNA sequences from this study

Mitochondria/gene: ccmFN (cytochrome c biogenesis)					
Base positions → DNA sequence names↓	1	7	25	29	46
ZK-AVV38	T	A	G	C	G
GQ2203241		G	A	A	
FM1793801		G	A	A	
AM4469041		G	A	A	
Mitochondria/Gene: cox3 (cytochrome c oxidase subunit 3)					
Base positions → DNA sequence names↓	1	39	49	60	81
ZK-AVV39	A	C	T	T	G
XM 019218225.1		A	G	C	
GQ220323.1		A	G	C	
FM179380.1		A	G	C	
AM443817.2		A	G	C	

specific primer design for the selected gene regions along with ancient grape whole genome DNA amplification to get sufficient *a*DNA amount and quality as a template are needed considering that ancient DNA of charred seeds from Kaymakçı appeared to be highly damaged and fragmented. However, we believe that the gene associated with the CMap ancient grapevine DNA sequences could be used as a reference for the selection of gene regions to analyse the magnitude of genetic changes in a given gene or genes of modern varieties versus ancient varieties. The information that obtained from such studies will help to recover the lost alleles of important genes using the advance tools available from molecular genetics today.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11105-025-01573-7.

Acknowledgements We express our gratitude to the Fulbright Scholar Program for offering Prof. Z. Kaya a research grant at UC-Davis, CA, USA in 2022 that made possible our collaboration on genomic variations in *Vitis vinifera*, from ancient (extant) to contemporary varieties. For permissions to conduct excavations at Kaymakçı and analyse its materials, we thank the Manisa Museum and the General Directorate of Cultural Heritage and Museums, Ministry of Culture and Tourism, Türkiye.

Author Contribution Z.K. carried out the data analysis and writing the manuscript, F.Ö.D. did the ancient DNA extraction and characterization, S.Y. prepared the DNA libraries for whole genome shotgun sequencing, B.K.V. and M.S.contributed on bioinformatics and sequencing, C..L and C.H.R. provided ancient grape seeds and archaeological review of the manuscript, D.B.N. had the final review and editing of the manuscript and hosting Z.K.

Data Availability Datasets which were not already publicly available at the time of submission will be deposited to The Species database of Tree Genes database (http://treegenesdb.org) upon the acceptance of the paper.

Code Availability The name, version, and parameters of the software used in this study are described in the "Materials and Methods" section.

Declarations

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Conflict of Interest The authors declare no competing interests.

References

Anderson-Carpenter LL, McLachlan JS, Jackson ST, Kuch M, Lumibao CY, Poinar HN (2011) Ancient DNA from lake sediments: bridging the gap between paleoecology and genetics. BMC Evol Biol 11:30

Arroyo-Garcia R, Ruiz-Garcia L, Bolling L, Ocete R, Lòpez MA, Arnold C, Martinez-Zapater JM (2006) Multiple origins of cultivated grapevine (Vitis vinifera L. ssp. sativa) based on chloroplast DNA polymorphisms. Mol Ecol. 15:3707–3714

Brown TA, Capellini E, Kistler L, Lister DL, Oliveira HR, Wales N, Schumbaum A (2015) Recent advances in ancient DNA research and their implications for archaeobotany. Veget Hist Archaeobot 24:207214

Cappellini E, Gilbert MTP, Geuna F, Fiorentino G, Hall A, Collins MJ (2010) A multidisciplinary study of archaeological grape seeds. Naturwissenschaften 97(2): 205–17

Çiftci A, Değirmenci FO, Roosevelt C, Luke C, Marston J, Kaya Z (2019) Ancient DNA (*a*DNA) extraction and amplification from 3500 year-old charred economic crop seeds from Kaymakçı in western Turkey: comparative sequence analysis using the 26S rDNA Gene. Genet Resour Crop Evol 66:1279–1294

Cohen P, Bacilieri R, Ramos-Madrigal J, Privman E, Boaretto E, Weber A, Meiri M (2023) Ancient DNA from a lost Negev

Highlands desert grape reveals a Late Antiquity wine lineage. PNAS 120(17): e2213563120
Corso M, Vannozzi A, Maza E, Vitulo N, Meggio F, Pitacco A, Lucchin M (2015) Comprehensive transcript profiling of two grapevine rootstock genotypes contrasting in drought susceptibility links the phenylpropanoid pathway to enhanced tolerance. J Exp Bot. 66: 5739±5752
Değirmenci FÖ, Ulaş B, Kansu C, Uluğ A, Caneva I, Asal R, Kaya Z (2022) The ancient DNA and archaeobotanical analysis suggest cultivation of Triticum aestivum subsp. spelta at Yumuktepe and Yenikapı Pottery Neolithic sites in Turkey. Genetic Resources and Crop Evolution 69 (7): https://doi.org/10.1007/s10722-022-01453-z
Di Genova A, Almeida AM, Muñoz-Espinoza C, Vizoso P, Travisany D, Moraga C, Pinto M, Hinrichsen P, Orellana A, Maass A (2014) Whole genome comparison between table and wine grapes reveals a comprehensive catalog of structural variants. BMC Plant Biol 14:7
Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S, Cheng T, Guo J, Zhou S (2015) ycf1, the most promising plastid DNA barcode of land plants. Sci Rep 5:8348
Dong Y, Duan S, Xia Q, Liang Z, Dong X, Margaryan K, …., Chen W, (2023) Dual domestications and origin of traits in grapevine evolution. Science 379:892
Focus OIV (2017) http://www.oiv.int/public/medias/5888/en-distribution-of-the-worlds-grapevinevarieties.pdf
Gökbayrak Z, Söylemezoğlu G (2010) Grapevine throughout the history of Anatolia. Int J Bot 6:465–472
Goremykin VV, Salamini F, Velasco R, Viola R (2009) Mitochondrial DNA of *Vitis vinifera* and the issue of rampant horizontal gene transfer. Mol Biol Evol 26:99–110
Gunther T, Valdiosera C, Malmstrom H, Urena I, Rodriguez-Varela R, Sverrisdottir OO, Daskalaki EA, Skoglund P, Naidoo T, Svensson EM, Jakobsson M (2015) Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. Proc. Natl. Acad. Sci. USA 112:11917–11922
Guo D, Li D, Li H, Wang S, Wang L, Niu J, Ma C (2020) The complete chloroplast genome sequence of *Vitis vinifera* Muscat Hamburg. Mitochondrial DNA B Resour 5(1):117–118
He Y, Ahmad D, Zhang Z, Zhang Y, Wu L, Jiang P, Ma H (2018) Genome-wide analysis of family-1 UDP glycosyltransferases (UGT) and identification of UGT genes for FHB resistance in wheat (*Triticum aestivum* L.). BMC Plant Biology 18:67
Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Wincker P (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463–467
Jansch L, Kruft V, Schmitz UK, Braun HP (1996) New insights into the composition, molecular mass and stoichiometry of the protein complexes of plant mitochondria. Plant J 9:357–368
Jonsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L (2013). mapDamage2.0
Karataş DD, Karataş H, Laucou V, Sarikamis G¸ Riahi L, Bacilieri R, This P, (2014) Genetic diversity of wild and cultivated grapevine accessions from southeast Turkey. Hereditas 151:73–80
Kılınç GM, Koptekin D, Atakuman C, Sümer AP, Dönertas HM, …., Somel M, (2016) The demographic development of the first farmers in Anatolia. Curr Biol 26(19):2659–2666
Kircher M (2012) Analysis of high-throughput ancient DNA sequencing data. Methods Mol Biol 840:197–228
Koptekin D, Yüncü E, Rodríguez-Varela R, Altınışık NE, Psonis N, Kashuba N, Somel M (2023). Spatial and temporal heterogeneity in human mobility patterns in Holocene Southwest Asia and the East Mediterranean. Current Biology 33(1):41–57
Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. Bioinformatics, 25(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352
Lister DL, Bower MA, Howe CJ, Jones MK (2008) Extraction and amplification of nuclear DNA from herbarium specimens of emmer wheat, amethod for assessing DNA preservation by maximum amplicon length recovery. Taxon 57:254–258
Luke C, Roosevelt CH (2017) Cup-marks and citadels: evidence for libation in the second-millennium BCE Marmara Lake Basin, Western Anatolia. Bull Am Sch Orient Res 378:1–23
Luke C, Roosevelt CH, Cobb P, Cilingiroglu C (2015) Composing communities: chalcolithic through iron age survey ceramics in the Marmara lake basin. Western Turkey J Field Archaeol 40(4):428–449
Magris G, Jurman I, Fornasiero A, Paparelli E, Schwope R, Marroni F, Di Gaspero G, Morgante M (2021) The genomes of 204 *Vitis vinifera* accessions reveal the origin of European wine grapes. Nat Commun 12:7240
Manen JF, Bouby L, Dalnoki O, Marinval P, Turgay M, Schlumbaum A (2003) Microsatellites from archaeological *Vitis vinifera* seeds allow a tentative assignment of the geographical origin of ancient cultivars. J Archaeol Sci 30:721–729
Mansilla N, Racca S, Gras DE, Gonzalez DH, Welchen E (2018) The complexity of mitochondrial complex IV: an update of cytochrome c oxidase biogenesis in plants. Int J Mol Sci 19:662
Marston JM, Çakırlar C, Luke C, Kováčik P, Slim FG, Shin N, Roosevelt CH (2021) Agropastoral economies and land use in Bronze Age Western Anatolia. Environ Archaeol 27(6):539–553
McGovern PE, Glusker DL, Exner LJ, Voigt MM (1996) Neolithic resinated wine. Nature 381:480–481
Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc 6:t5448
Moraga AR, Nohales PF, Pérez JAF, Gómez-Gómez L (2004) Glucosylation of the saffron apocarotenoid crocetin by a glucosyltransferase isolated from *Crocus sativus* stigmas. Planta 219:955–966
Myles S, Chia J-M, Hurwitz B, Simon C, Zhong GY, Buckler E, Ware D (2010) Rapid genomic characterization of the genus *Vitis*. PLoS ONE 5(1):e8219
Myles S, Boyko AR, Owens CL, Brown PJ, Grassi F, Aradhya MK, Prins B, Reynolds A, Chia JM, Ware D, Bustamante CD, Buckler ES (2011) Genetic structure and domestication history of the grape. PNAS 108:3530–3535
Ohtsu K, Nakazono M, Tsutsumi N, Hirai A (2001) Characterization and expression of the genes for cytochrome c oxidase subunit VIb (COX6b) from rice and *Arabidopsis thaliana*. Gene 264:233–239
Oliveira HR, Civan P, Morales J, Rodriguez-Rodriguez A, Lister DL, Jones MK (2012) Ancient DNA in archaeological wheat grains: preservation conditions and the study of pre-Hispanic agriculture on the island of Gran Canaria (Spain). J Archaeol Sci 39:828–835
Özgen M, Özdilek A, Birsin MA, Önde S, Şahin D, Açıkgöz E, Kaya Z (2012) Analysis of ancient DNA from in vitro grown tissues of1600-year-old seeds revealed the species as *Anagyris foetida*. Seed Sci Res 22:279–286
Parducci L, Alsos IG, Unneberg P, Pedersen MW, Han L, Lammers Y, Salonen JS, Väliranta MM, Slotte T, Wohlfarth B (2019) Shotgun environmental DNA, pollen, and macrofossil analysis of late glacial lake sediments from southern Sweden. Front Ecol Evol 7:189
Picardi E, Horner DS, Chiara M, Schiavon R, Valle G, Pesole G (2010) Large-scale detection and analysis of RNA editing in grape mtDNA by RNA deep-sequencing. Nucleic Acids Res 38(14):4755–4767
Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinform Oxf Engl 26:841–842

Rogers SO, Kaya Z (2006) DNA from ancient cedar wood from King Midas' Tomb, Turkey, and Al-Aksa Mosque. Israel Silvae Genetica 55(2):54–62

Roosevelt CH, Luke C (2017) The story of a forgotten kingdom? Survey archaeology and the historical geography of Central Western Anatolia in the second millennium BC. Eur J Archaeol 20(1):120–147

Roosevelt CH, Luke C, Ünlüsoy C, Çakırlar C, Marston JM, O'Grady CR, Pavuk P, Pieniazek M, Mokrisova J, Scott C, Shin N, Slim F (2018) Exploring space, economy, and interregional interaction at a second-millennium BCE citadel in Central Western Anatolia: 2014–2017 Research at Kaymakçı. Am J Archaeol 122(4):645–688

Sabina J, Leamon JH (2015) Bias in whole genome amplification: causes and considerations. In: Whole Genome Amplification: Methods and Protocols, Thomas Kroneis (ed.). Methods in Molecular Biology 1347:15–37, Springer Science+Business Media, New York

Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, …, Sherry ST. (2022) Database resources of the national center for biotechnology information. Nucleic Acids Res 50, Database issue: D20-D26

Schubert M, Lindgreen S, Orlando L (2016) AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res Notes 9:88

Shin N, Marston JM, Luke C, Roosevelt CH, Riehl S (2021) Agricultural practices at Bronze Age Kaymakçı, Western Anatolia. J Archaeol Sci Rep 36:102800

Da Silva C, Zamperin G, Ferrarini A, Minio A, Dal Molin A, Venturini L, Delledonne M (2013) The high polyphenol content of grapevine cultivar tannat berries is conferred primarily by genes that are not shared with the reference genome. Plant Cell 25: 4777–4788

Takahashi K, Niwa H, Yokota N, Kubota K, Inoue H (2008) Widespread tissue expression of nepenthesin-like aspartic protease genes in *Arabidopsis thaliana*. Plant Physiol Biochem 46(7):724–729

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

Uluğ A, Değirmenci FO, Kansu Ç, Luke C, Roosevelt CH, Kaya Z (2025) Archaeogenetics in Anatolia: genetic alterations in 2nd millennium BCE crops through cultivation and breeding activities. (Manuscript submitted for publication). Department of Biological Sciences, Middle East Technical University, Ankara, Türkiye

Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, Viola R (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS One 2(12): e1326

Venturini L, Ferrarini A, Zenoni S, Tornielli GB, …, Delledonne M (2013) De novo transcriptome characterization of Vitis vinifera cv. corvina unveils varietal diversity. BMC Genomics14:41

Willerslev E, Hansen AJ, Poinar HN (2004) Isolation of nucleic acids and cultures from fossil ice and permafrost. Trends Ecol Evol 19(3):141–147

Zhang G, Wang H, Ren X, Xiao Y, Liu D, Tong H (2024) Brassinosteroid-dependent phosphorylation of phosphate starvation response2 reduces its DNA-binding ability in rice. The Plant Cell 36(6):2253–2271

Zhao Y, Yin J, Guo H, Zhang Y, Xiao W, Xiao J (2015) The complete chloroplast genome provides insight into the evolution and polymorphism of Panax ginseng. Front Plant Sci 5:696. https://doi.org/10.3389/fpls.2014.00696

Zhou Y, Massonnet M, Sanjak, JS, Cantu D, Gaut BS (2017) Evolutionary genomics of grape (Vitis vinifera ssp. vinifera) domestication. PNAS 114 (44):11715–11720

Zohary D, Spiegel-Roy P (1975) Beginnings of fruit growing in the old world. Science