Journal of Archaeological Science 72 (2016) 57–70

Contents lists available at ScienceDirect

Journal of Archaeological Science

journal homepage: http://www.elsevier.com/locate/jas

The limits and potential of paleogenomic techniques for reconstructing grapevine domestication

Nathan Wales [a, b, *], Jazmín Ramos Madrigal [a], Enrico Cappellini [a], Aldo Carmona Baez [a, c], José Alfredo Samaniego Castruita [a], J. Alberto Romero-Navarro [a, d], Christian Carøe [a, e], María C. Ávila-Arcos [a, f], Fernando Peñaloza [a, c, g], J. Víctor Moreno-Mayar [a], Boris Gasparyan [h], Diana Zardaryan [h], Tamara Bagoyan [h], Alexia Smith [b], Ron Pinhasi [i], Giovanna Bosi [j], Girolamo Fiorentino [k], Anna Maria Grasso [k], Alessandra Celant [l], Guy Bar-Oz [m], Yotam Tepper [m], Allan Hall [n], Simone Scalabrin [o], Mara Miculan [p], Michele Morgante [p], Gabriele Di Gaspero [p], M. Thomas P. Gilbert [a, q, r]

[a] Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Denmark
[b] Department of Anthropology, University of Connecticut, Storrs, CT, USA
[c] Undergraduate Program on Genomic Sciences, National Autonomous University of Mexico, 62210 Cuernavaca, Mexico
[d] School of Integrative Plant Sciences, Section of Plant Breeding and Genetics, Cornell University, Ithaca, NY, USA
[e] Center for Biological Sequence Analysis, Technical University of Denmark, Kgs. Lyngby, Denmark
[f] International Laboratory for Human Genome Research, National Autonomous University of Mexico, 76230 Querétaro, Qro., Mexico
[g] Leibniz Institute for Zoo and Wildlife Research, 10315 Berlin, Germany
[h] Institute of Archaeology and Ethnology, National Academy of Sciences, Yerevan, Armenia
[i] School of Archaeology and Earth Institute, University College Dublin, Dublin 4, Ireland
[j] Laboratorio di Palinologia e Paleobotanica, Dipartimento di Scienze della Vita, Università di Modena e Reggio Emilia, Italy
[k] Laboratory of Archaeobotany and Palaeoecology, University of Salento, Italy
[l] Dipartmento di Bilogia Ambientale, "Sapienza" University of Rome, Italy
[m] Zinman Institute of Archaeology, University of Haifa, Haifa, Mount Carmel 3498837, Israel
[n] Department of Archaeology, University of York, York, UK
[o] IGA Technology Services, Udine, Italy
[p] Istituto di Genomica Applicata (IGA), Udine, Italy
[q] Trace and Environmental DNA Laboratory, Department of Environment and Agriculture, Curtin University, Perth, Western Australia, Australia
[r] NTNU University Museum, N-7491 Trondheim, Norway

ARTICLE INFO

Article history:
Received 15 April 2016
Received in revised form
25 May 2016
Accepted 31 May 2016

Keywords:
Ancient DNA
Domestication
Archaeobotany
Grapevine
Wine
Chloroplast DNA
High-throughput sequencing

ABSTRACT

In ancient DNA (aDNA) research, evolutionary and archaeological questions are often investigated using the genomic sequences of organelles: mitochondrial and chloroplast DNA. Organellar genomes are found in multiple copies per living cell, increasing their chance of recovery from archaeological samples, and are inherited from one parent without genetic recombination, simplifying analyses. While mitochondrial genomes have played a key role in many mammalian aDNA projects, including research focused on prehistoric humans and extinct hominins, it is unclear how useful plant chloroplast genomes (plastomes) may be at elucidating questions related to plant evolution, crop domestication, and the prehistoric movement of botanical products through trade and migration. Such analyses are particularly challenging for plant species whose genomes have highly repetitive sequences and that undergo frequent genomic reorganization, notably species with high retrotransposon activity. To address this question, we explored the research potential of the grape (*Vitis vinifera* L.) plastome using targeted-enrichment methods and high-throughput DNA sequencing on a collection of archaeological grape pip and vine specimens from sites across Eurasia dating ca. 4000 BCE–1500 CE. We demonstrate that due to unprecedented numbers of sequence insertions into the nuclear and mitochondrial genomes, the grape plastome provides limited intraspecific phylogenetic resolution. Nonetheless, we were able to assign archaeological specimens in

* Corresponding author. Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade, 5-7 1350 Copenhagen K, Denmark.
E-mail address: nathan.wales@snm.ku.dk (N. Wales).

http://dx.doi.org/10.1016/j.jas.2016.05.014

the Italian peninsula, Sardinia, UK, and Armenia from pre-Roman to medieval times as belonging to all three major chlorotypes A, C, and D found in modern varieties of Western Europe. Analysis of nuclear genomic DNA from these samples reveals a much greater potential for understanding ancient viticulture, including domestication events, genetic introgression from local wild populations, and the origins and histories of varietal lineages.

1. Introduction

Through the analysis of aDNA from archaeobotanical remains, it is possible to shed light on long-standing archaeological and evolutionary questions, including the geographic origins of crops, rates of change during domestication, human migration patterns, past trade networks, and paleoecology (as reviewed by Brown et al., 2015; Palmer et al., 2012; Wales et al., 2013). One key question for plant aDNA research is the suitability of the chloroplast genome to address archaeological and evolutionary questions. This genetic compartment, more generally referred to as the "plastome" because all plastids in an individual contain the same genome (Herrmann and Possingham, 1980), plays a large role in genetic research of modern plants due to its ability to act as a genetic barcode, differentiating between species and lineages at several genetic markers (CBOL Plant Working Group et al., 2009; China Plant BOL Group et al., 2011). Like mitochondria, plastids are the descendants of endosymbiotic bacteria (Gray, 1993; Sagan, 1967), and are present in many copies per living cell, each containing a copy of their genome (Bock, 2007). Plastids are almost always inherited uniparentally, most commonly from the mother, and very rarely undergo genetic recombination (Birky, 2001), although there are exceptions (Wolfe and Randle, 2004). These characteristics, coupled with expansive DNA barcoding databases of living plant accessions, makes the plastome an obvious target for aDNA research. In fact, the same traits of uniparental inheritance and high copy number have led aDNA researchers to study mitochondrial genomes from extinct mammalian species (Hofreiter et al., 2001), including wooly mammoths (Rogaev et al., 2006), wooly rhinoceros (Willerslev et al., 2009), Neanderthals (Green et al., 2008), and Denisova hominins (Krause et al., 2010).

To date, analysis of plastome sequences from archaeological and environmental samples has been primarily focused on short DNA barcoding loci (Palmer et al., 2012). This method has been especially useful in reconstructing plant communities of ancient environments using sediment samples (Haile et al., 2007; Willerslev et al., 2003), as well as for inferring ancient human diets from coprolites and gut contents (Rasmussen et al., 2009; Rollo et al., 2002). In these studies, the polymerase chain reaction (PCR) was used to amplify hypervariable stretches of DNA flanked by conserved regions. While this approach can distinguish distantly related plant taxa, closely related species often have identical barcodes, due to their recent evolutionary divergence (Taberlet et al., 2007). This limitation is magnified in archaeological samples, because DNA is highly fragmented, necessitating the use of "mini-barcodes", small portions of the conventional DNA barcodes with reduced phylogenetic resolution (Little, 2014).

Considering plastid genomes range in size from 120,000 to 160,000 nucleotides in most plant species (Bock, 2007), the DNA barcodes within them represent <1% of the total sequence. Because genetic polymorphisms are not limited to the barcoding regions, by studying complete plastomes of different plant species or individuals within a species, researchers can observe many more polymorphic loci and investigate phylogenetic questions that are not feasible using conventional barcode markers. In modern DNA-based projects, full plastome sequences have been used to clarify deep evolutionary splits in the plant kingdom (Karol et al., 2010), resolve evolutionary relationships of closely related species (e.g. Sherman-Broyles et al., 2014; Yang et al., 2013), and explore divergence and evolution of intraspecific lineages (Schelkunov et al., 2015; Whittall et al., 2010). Relationships within a species are of significant interest for archaeobotanical aDNA research, as they can shed light on important questions about domestication of a crop and transportation by past humans. In two aDNA projects focused on Cucurbitaceae, the gourd family, researchers recovered significant portions of plastomes from archaeological samples using high throughput DNA sequencing technologies, permitting new understandings of the movement of the bottle gourd (*Lagenaria siceraria* (Molina) Standl.) to the Americas and the domestication of gourds and squashes (*Cucurbita* spp. L.) (Kistler et al., 2014, Kistler et al., 2015). Despite the success of these aDNA studies, it is unclear whether complete plastome sequences are equally useful for other aDNA projects, especially in species with repetitive genomic sequences and high levels of retrotransposons—mobile genetic elements that may compromise >50% of a plant genome (Kumar and Bennetzen, 1999). Indeed, even in *Lagenaria siceraria*, horizontal gene transfer from the plastome to mitochondrial genome makes some plastome loci uninformative for ancestry inferences (Kistler et al., 2014), a finding which undermines earlier PCR-based findings for an Asian origin of bottle gourds in the Americas (Erickson et al., 2005).

To more broadly investigate the extent to which plastomes can be used to answer archaeological questions, we explored the genomic locus in ancient specimens and in modern Eurasian grapevines (*Vitis vinifera* L.), covering the current geographical range of viticulture and representing all known modern chlorotypes. The grape genome presents a significant analytical challenge, due to high levels of horizontal gene transfer between the nuclear, plastid, and mitochondrial genomes (Michalovova et al., 2013). Plastome inserts in the nuclear (nupt) and mitochondrial (mtpt) genomes are here referred to collectively as *o*rganellar-*d*erived *in*serted *s*equences or "odins" (Samaniego Castruita et al., 2015). At least 42% of the grape plastome has been integrated into the mitochondrial genome (Goremykin et al., 2009), where selection is relaxed and mutations occur more frequently than in the plastome, potentially making it very difficult to distinguish plastome aDNA molecules from odins, especially in degraded specimens. Still, the high copy number of plastids in living cells increases the chance that the plastome can be recovered in archaeological samples, potentially making it a more practical target than nuclear loci.

The domestication of wild grapevines and their subsequent spread by humans around the world is a complicated archaeological topic for which aDNA investigation can provide important insights into the timing and pace of change. The wild grapevine (*Vitis vinifera* L. ssp. *sylvestris* (C.C. Gmelin) Hegi), hereafter *sylvestris*, is a dioecious plant yielding small, dark berries, with a natural range spanning the Mediterranean to west of the Himalaya Mountains. Archaeological evidence indicates that the domesticated grapevine (*Vitis vinifera* L. ssp. *sativa* Hegi), hereafter *sativa*, characterized by

larger berries and hermaphroditism (This et al., 2006), first appeared ca. 8000 years ago in Western Asia (Miller, 2008; Zohary, 1996; Zohary et al., 2012). Traces of grape-derived compounds on archaeological storage vessels and site features from this region strongly indicate an early use of grapes for consumption as a beverage (Barnard et al., 2011; McGovern et al., 1996), likely as wine, given that the juice spontaneously ferments when berries are crushed (Beck, 2004). Patterns in genetic diversity of modern *sylvestris* and *sativa* also point to the Near East as the probable center of domestication (Arroyo-García et al., 2006; Bacilieri et al., 2013; Myles et al., 2011). However, the timing of this domestication and *sativa's* precise geographic origins continue to be contested. Later events in the history of grape domestication are equally ambiguous, including a possible secondary domestication event in Western Europe (Arroyo-García et al., 2006) or introgression of Western *sylvestris* into introduced varieties of *sativa* (Myles et al., 2011), millennial-scale changes in vinicultural practices and vineyard management, and the origins of varietal lineages. These processes are currently examined by archaeobotanical studies using image analysis techniques (Terral et al., 2010), in which morphometric features distinguish *sylvestris* and *sativa* cultivars, leading to insights on ancient grapevine cultivation (Bouby et al., 2013; Ucchesu et al., 2015). Such morphometric approaches can be complemented by aDNA from archaeological samples, providing unparalleled snapshots of past genetic diversity to understand these complex processes. The plastome would at first glance be particularly suitable to study these issues in grapes because when vines are propagated vegetatively, a lineage exists as a genetic clone, and a particular plastome becomes inextricably linked to each variety. Given grapevine's venerable symbolic and ritualistic role in many societies (McGovern, 2003), as well as its position as today's most economically important fruit (Jenster et al., 2008), the situation is ripe for high-throughput sequencing of aDNA from archaeological grape specimens.

To date, few published studies have utilized grape aDNA, either for methodological studies (Wales et al., 2014, Wales et al., 2012), or in conventional PCR-based approaches (Cappellini et al., 2010; Foley et al., 2012; Hansson and Foley, 2008; Manen et al., 2003). Here, we circumvent the limitations of PCR-based studies (discussed by Knapp and Hofreiter, 2010; Rizzi et al., 2012) and attempt to characterize plastomes from 27 archaeological grape samples using a target-enrichment methodology and high-throughput DNA sequencing. To investigate the potential of the grape plastome for aDNA research, we enrich for plastome DNA using an in-solution capture method (Ávila-Arcos et al., 2011; Gnirke et al., 2009), and test genetic relationships between archaeological and modern samples. In addition, we examine nuclear genomic sequences from the archaeological samples to better understand the limitations of plastome data and determine if the nuclear genome may be a more practical target for future grape aDNA research.

2. Methods and materials

2.1. Archaeological samples

Twenty-seven archaeobotanical specimens were tested from 17 archaeological sites, spanning a range of geographic locations and time periods (Fig. 1 and Table 1). The majority of the specimens were waterlogged grape pips, excavated from wells, latrines, and in one case, a shipwreck. In addition to the waterlogged samples, three desiccated pips from a Byzantine dovecote near Shivta, Israel were tested (Hirschfeld and Tepper, 2006); we note that pigeons likely digested these seeds in antiquity. Three desiccated grapevine branches were tested from Areni-1, a dry cave in Armenia with the oldest known winemaking installation, dating to ca. 4000 BCE (Barnard et al., 2011; Wilkinson et al., 2012). All three branches from Areni-1 were of sufficient size to allow direct radiocarbon dating and aDNA analyses. The small sizes of pips did not permit direct dating and aDNA analysis, but closely associated AMS dates are available for Çamaltı Burnu, Colosseum, Sa Osa, Scorpo and Torre Montello. The age of the other pips is inferred from associated artifacts and archaeological contexts (Table 1). The stratigraphic integrity of the archaeological sites do not indicate any of the seeds are likely to be intrusive, but we note that the small size of pips renders them prone to vertical movement via bioturbation and other site formation processes (Reitz and Shackley, 2012). Nonetheless, the testing of many samples from multiple contexts places the focus of this study on the large patterns in past *Vitis* genetic diversity, minimizing the impact of individual samples. As far as could be ascertained, none of the specimens were charred or had otherwise been exposed to fire. Exposure to high temperatures, such as in charring, can be expected to heavily fragment DNA (Lindahl, 1993), with greater destruction with higher temperatures and longer exposure.

2.2. Modern samples

Two modern grapevine specimens were included in the study as positive controls for the experimental technique. Both leaf samples were collected in Armenia, and were tested to supplement the dataset for individuals that might be more closely related to samples from the Caucasus. One *sativa* individual was collected from a vineyard near the village of Hatis, in the Kotayk Province of central Armenia. Another individual, referred to as "Feral" was collected 3 km away from the Areni-1 site, upstream beside the Gnishik stream. This individual was growing among a stand of grapevines, thought to be feral descendants of a medieval vineyard, perhaps related to the nearby Noravank monastery.

2.3. aDNA extraction and library preparation

DNA was extracted from archaeological specimens in an aDNA facility at the University of Copenhagen that implements extensive measures to minimize contamination, including full body suits, nightly UV-irradiation, and isolation from post-PCR laboratories (Cooper and Poinar, 2000; Gilbert et al., 2005). Laboratory methods slightly varied across samples, as different specimens were processed over the course of the project, in accordance to improvements in methodologies. Detailed methods used for each sample are described in Table S1, but in general followed commonly employed techniques for plant aDNA. Seeds were rinsed in a 5% sodium hypochlorite (bleach) solution for 30 s, followed by a wash with molecular biology grade water, except when they appeared likely to absorb the solution, thereby destroying endogenous DNA. Sediment and other contaminants were removed from branches by removing the external bark with a sterile scalpel. All samples were extracted using a method which had previously been used on archaeological grape pips (Cappellini et al., 2010), and has been validated across a wide range of non-carbonized archaeobotanical remains (Wales et al., 2014). In addition to the standard extraction technique, portions of the two oldest vines were extracted using an enzymatic technique in an attempt to more thoroughly disrupt cell walls and release DNA (Manen et al., 2005).

Due to the limited sizes of most archaeobotanical remains when compared to many archaeological skeletal remains, it is open question whether researchers should extract seeds individually or in bulk. In theory, the extraction of a single seed is preferable because only one genetic signature is present. In practice, however, if DNA yields are very low, insufficient endogenous DNA may be available for library preparation or genetic characterization.

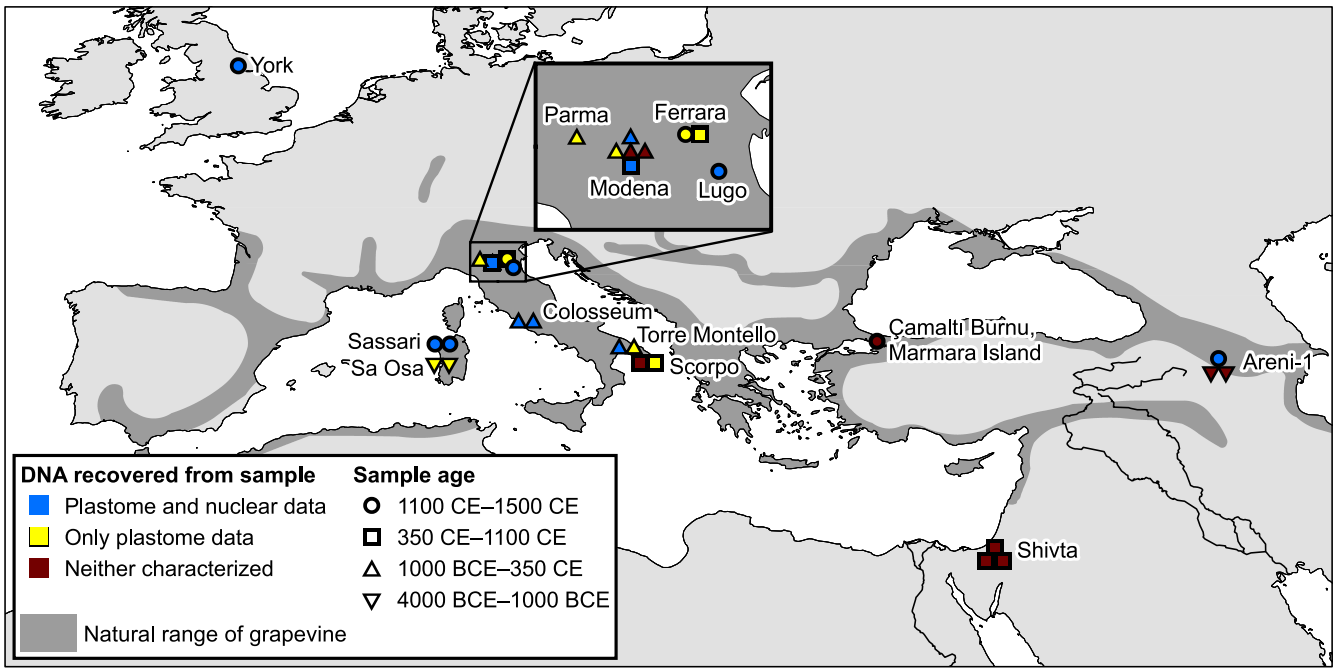

Fig. 1. Map of archaeological sites with number of samples processed. Marker color indicates whether samples yielded sufficient nuclear DNA or plastome data for inclusion in analyses and symbol shape represent age grouping. Natural range of grapevine based on Zohary (1996). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Therefore, for 10 of the 24 seed samples, DNA was extracted from multiple pips excavated from the same stratigraphic unit. These multi-seed samples may consequently have a mixed signal from multiple individuals. The other 14 seed samples were performed on a single grape pip, with the goals of achieving less ambiguous biological findings and providing new insights into whether single seeds can yield sufficient endogenous DNA for genetic characterization.

Extracted aDNA was converted to Illumina-compatible libraries using double-stranded DNA preparations, either with A-tailing and Y-shaped adapters (New England Biolabs NEBNext library kit E6040L), or blunt-end repair and blunt-adapters (Meyer and Kircher, 2010) (Table S1). DNA extracts of the two oldest vines were also converted to Illumina libraries using a single-stranded DNA library preparation to maximize library complexity and improve the recovery of very short DNA molecules (Gansauge and Meyer, 2013). Libraries were amplified using either Phusion or AmpliTaq Gold DNA polymerases (Thermo Fisher Scientific, Waltham, MA), the latter enabling a less biased representation of aDNA damage patterns. The number of PCR cycles used varied by sample in order to reach sufficient quantities of DNA for capture-enrichment experiments, and when necessary, an additional round of PCR was performed using high fidelity Phusion polymerase (Table S1).

2.4. Modern DNA extraction and library preparation

DNA from two modern grape samples was extracted in a facility at the University of Copenhagen designated for modern DNA, in a separate building from where aDNA is extracted and converted to sequencing libraries. DNA from leaf tissue was extracted using a protocol designed for samples rich in polyphenols, especially grapes (Japelaghi et al., 2011). In preparation for library building, genomic DNA was fragmented to an average length of ~200 bp using a Diagenode Bioruptor NGS, using 20 cycles of 30 s of sonication and 30 s of rest. Sheared DNA was converted to Illumina-compatible libraries using NEBNext library kit E6070L (New England Biolabs) and blunt-ended library adapters described by Meyer and Kircher (2010). The libraries were amplified with Phusion polymerase for a total of 14 cycles with sample-specific indexes included in the P7 primer (Meyer and Kircher, 2010).

2.5. Targeted capture of plastome

Polymorphisms and genetic diversity in the grape plastome have been characterized by several researchers, notably microsatellites (Arroyo-García et al., 2006), but also base insertions/deletions (indels) and single nucleotide polymorphisms (SNPs) (Beridze et al., 2011; Hunt et al., 2010; Pipia et al., 2012; Tröndle et al., 2010). Such reference data can allow ancient samples to be placed in a greater context, although like mitochondrial genomes, their unilineal inheritance allows only limited insights into genetic relationships.

PCR-based amplification of a small number of genetic markers is challenging for most archaeological samples, due to the fragmented nature of aDNA and the limited number of template molecules (reviewed by Shapiro and Hofreiter, 2014). In addition, archaeobotanical remains often contain enzymatic inhibitors, potentially leading to false negative results (Wales et al., 2014). In contrast, a hybridization enrichment methodology developed for modern DNA (Gnirke et al., 2009) has been shown to be suited for aDNA research, including archaeobotanical remains (Ávila-Arcos et al., 2011; da Fonseca et al., 2015; Kistler et al., 2014, Kistler et al., 2015). Enrichment of the grape plastome was performed using custom-designed RNA probes (MYcroarrary, Ann Arbor, MI). By design, the 120-mer probes covered 108,222 of the 160,928 bp genome, to the exclusion of the two inverted repeats, as recovered DNA could not unambiguously be mapped to these regions.

Targeted enrichment was performed on both modern samples and 23 of the 24 archaeological samples, according to the

Table 1
Archaeological sample information. Samples with radiocarbon ages are listed with AMS laboratory codes in parentheses, and are provided with 95.4% confidence intervals, calibrated using OxCal Ver. 4.2 web interface build 96 (Bronk Ramsey and Lee, 2013) and IntCal13 (Reimer et al., 2013). The vines from Areni-1 were dated directly. The AMS date for Çamaltı Burnu is a wood sample from the frame of the ship. The Colosseum radiocarbon date is from organic waste from cesspits in the east sewer. At Sa Osa and Scorpo, AMS dates were taken for grape pips excavated in the same stratigraphic layers as those tested for aDNA. The date for Torre Montello is a seed of *Lagenaria* sp. L. found in the same stratigraphic layer as the two grape seeds. Other dates are based on archaeological context. Repository abbreviations are IAE: Institute of Archaeology and Ethnology, National Academy of Sciences, Yerevan, Armenia; UH: University of Haifa, Israel; UNIMORE: University of Modena and Reggio Emilia, Italy; UNIROMA: "Sapienza" University of Rome; US: University of Salento, Lecce, Italy; and York: University of York, England. Context abbreviation "SU" represents "stratigraphic unit."

Site	Location	Sample	Tissue	Preservation	Age	Repository	Context	Reference
Areni-1	Areni, Armenia	Areni-1	Vine	Desiccated	1166–1258 CE (OxA-24048)	IAE	Trench 1, square N16, locus 29	Wilkinson et al., 2012; Smith et al., 2014
Areni-1	Areni, Armenia	Areni-2	Vine	Desiccated	2201–1961 BCE (AA-95788)	IAE	Trench 1, unit 3, square R18, locus 19, spit 3	Wilkinson et al., 2012; Smith et al., 2014
Areni-1	Areni, Armenia	Areni-3	Vine	Desiccated	3982–3715 BCE (AA-95789)	IAE	Trench 3, square J26, spit 11	Wilkinson et al., 2012; Smith et al., 2014
Çamaltı Burnu, Marmara Island	Turkey	Çamaltı Burnu	3 seeds	Waterlogged	1263–1397 CE (LTL4984A)	US	Çamalti Burnu-1 shipwreck, sample ATB	Nergis, 2010; Grasso and Fiorentino, 2010
Colosseum	Rome, Italy	Colosseum-1	1 seed	Waterlogged	125–322 CE (Ua-42355)	US/ UNIROMA	Rome Colosseum, CCE 75	Ghini, 1988; Celant, 1998
Colosseum	Rome, Italy	Colosseum-2	1 seed	Waterlogged	125–322 CE (Ua-42355)	US/ UNIROMA	Rome Colosseum, CCE 75	Ghini, 1988; Celant, 1998
Ferrara: corso Porta Reno/via Vaspergolo	Ferrara, Italy	Ferrara-1	6 seeds	Waterlogged	1001–1050 CE	UNIMORE	SU 2599, rubbish bin in kitchen gardens	Bosi, 2000
Ferrara: piazza Municipale	Ferrara, Italy	Ferrara-2	4 seeds	Waterlogged	1451–1500 CE	UNIMORE	SU 1050, brick rubbish pit of Este's family	Bosi et al., 2009
Lugo: piazza Baracca	Lugo, Italy	Lugo	4 seeds	Waterlogged	1501–1550 CE	UNIMORE	SU 144, well 138, homes and artisan workshops	
Modena: area Novi Sad	Modena, Italy	Modena-1	1 seed	Waterlogged	100 BCE–100 CE	UNIMORE	SU 230, tub (probably used for fish farming)	Bosi et al., Submitted
Modena: ex Cassa di Risparmio	Modena, Italy	Modena-2	3 seeds	Waterlogged	1–25 CE	UNIMORE	SU 29, reclaimed channel	Bosi et al., Submitted
Modena: ex cinema Capitol	Modena, Italy	Modena-3	1 seed	Waterlogged	300–101 BCE	UNIMORE	SU 12, pre-domus (layer very anthropized)	Bosi et al., 2015b
Modena: ex cinema Capitol	Modena, Italy	Modena-4	1 seed	Waterlogged	1–200 CE	UNIMORE	Cut 2, Imperial domus	Bosi et al., 2015b
Modena: palazzo Vaccari	Modena, Italy	Modena-5	4 seeds	Waterlogged	401–600 CE	UNIMORE	SU 9, landfill (complete abandonment domus)	Rinaldi et al., 2013; Bosi et al., 2015a
Parma: piazza Garibaldi	Parma, Italy	Parma	4 seeds	Waterlogged	250–175 BCE	UNIMORE	SU 320, pit (votive offerings)	Bosi et al., 2011
Sa Osa	Cabras (Oristano), Italy	Sa Osa-1	1 seed	Waterlogged	1286–1115 BCE (OxA-25106)	US	SU 171, well N	Ucchesu et al., 2015
Sa Osa	Cabras (Oristano), Italy	Sa Osa-2	1 seed	Waterlogged	1276–1088 BCE (OxA-25107)	US	SU 171, well N	Ucchesu et al., 2015
Sassari: via Satta	Sassari, Italy	Sassari-1	1 seed	Waterlogged	1301–1350 CE	UNIMORE	SU 1336, well in an urban courtyard	Bosi and Bandini Mazzanti, 2013
Sassari: via Satta	Sassari, Italy	Sassari-2	1 seed	Waterlogged	1301–1350 CE	UNIMORE	SU 1336, well in an urban courtyard	Bosi and Bandini Mazzanti, 2013
Scorpo	Supersano (Lecce), Italy	Scorpo-1	3 seeds	Waterlogged	686–876 CE (LTL2562A)	US	SU 204, well	Arthur et al., 2008; Calcagnile et al., 2011
Scorpo	Supersano (Lecce), Italy	Scorpo-2	3 seeds	Waterlogged	686–875 CE (LTL2563A)	US	SU 204, well	Arthur et al., 2008; Calcagnile et al., 2011
Shivta	Shivta, Israel	Shivta-1	1 seed	Desiccated	501–600 CE	UH	G-8/04, structure VI	Hirschfeld and Tepper, 2006; Ramsay and Tepper, 2010
Shivta	Shivta, Israel	Shivta-2	1 seed	Desiccated	501–600 CE	UH	G-8/04, structure VI	Hirschfeld and Tepper, 2006; Ramsay and Tepper, 2010
Shivta	Shivta, Israel	Shivta-3	1 seed	Desiccated	501–600 CE	UH	G-8/04, structure VI	Hirschfeld and Tepper, 2006; Ramsay and Tepper, 2010
Torre Montello	Taranto, Italy	Torre Montello-1	1 seed	Waterlogged	357–164 BCE (RTD-8238)	US	SU 14, well	Cinquantaquattro, 2012
Torre Montello	Taranto, Italy	Torre Montello-2	1 seed	Waterlogged	357–164 BCE (RTD-8238)	US	SU 19, well	Cinquantaquattro, 2012
York: 62–68 Low Petergate	York, England	York	3 seeds	Waterlogged	1301–1400 CE	York	Basal fill of a cobble-lined pit	Reeves, 2006; Cappellini et al., 2010

MYcroarrary's MYbaits manual version 1.3.8. Between 193 and 546 ng (mean = 384 ng) of amplified library was captured in an individual reaction (Table S1). The hybridization reactions were run in an Applied Biosystems 2720 Thermal Cycler for 24–53 h at 65 °C, with a 105 °C lid to prevent condensation. Following hybridization with RNA probes and streptavidin-coated magnetic beads, libraries were washed with supplied buffers, and DNA was released from the RNA probes using alkaline treatment. To reach sufficient quantities of DNA for sequencing, enriched libraries were amplified using Phusion polymerase for 12–24 PCR cycles.

Direct sequencing of the non-enriched library, i.e., shotgun sequencing, was performed on seven archaeological samples (Table S1). Shotgun sequencing allows an investigation into the enrichment of plastome DNA in the samples, although in some cases the captured and shotgun libraries do not originate from the same DNA extract, such as Areni-2. Shivta-3 was only shotgun

sequenced, and not captured. To achieve a general impression of capture success for samples that were not shotgun-sequenced, the amount of plastome-like DNA was compared to that obtained from shotgun sequencing of a modern grape sample.

Libraries were sequenced on Illumina GAIIx and HiSeq 2500 platforms, in single read mode with read lengths of 76 or 100 bp (Table S1).

2.6. Bioinformatic analyses

2.6.1. Read trimming and mapping

Sequencing adapters, low quality runs, and trailing Ns were trimmed from raw Illumina reads using AdapterRemoval v1.5 (Lindgreen, 2012). Trimmed reads were mapped with BWA aln v0.7.12 (Li and Durbin, 2009) to the reference grape plastome (Jansen et al., 2006) and the entire grape genome (Jaillon et al., 2007). The two mapping approaches were necessary because plastome-like DNA in the nuclear and mitochondrial genomes can also be enriched, as discussed below in the context of odins. In BWA, seeding (−l parameter) was disabled in order to prevent 5′ terminal substitutions characteristic to ancient DNA to bias the mapping (Schubert et al., 2012). PCR duplicates were identified by library and removed using Picard tools 1.130 (http://broadinstitute.github.io/picard). Reads were discarded if they mapped ambiguously in the genome or if they were assigned a mapping quality <30. Finally, reads were realigned around indels using GATK 3.3 (McKenna et al., 2010) and the MD-tag was calculated using SAMtools 1.2 (Li et al., 2009) before conducting subsequent analyses (Table 2).

2.6.2. Authenticating endogenous DNA

To further investigate the biological origin of DNA recovered from the samples, 10,000 trimmed reads from each library were randomly selected for querying against a local copy of NCBI's nucleotide database with the BLAST algorithm (Altschul et al., 1997; Camacho et al., 2009). The BLAST results were analyzed in MEGAN 5 using the default parameters to achieve an overview of the biological origins of DNA from the samples (Huson et al., 2011) (Table S2). While most archaeological samples demonstrate complex metagenomic signatures with high levels of microbial DNA (Carpenter et al., 2013; Poinar et al., 2006), we note most of our samples are only sequenced after target enrichment, leading to significant increases in endogenous content compared to shotgun sequencing.

DNA irreversibly degrades over time and with higher temperatures, becoming more fragmented and more chemically damaged, largely consisting of abasic sites and deaminated cytosine residues at the ends of the molecules (Dabney et al., 2013). Characterization of these features in DNA recovered from archaeological samples provides one means of determining if DNA likely is from an ancient/historic source as opposed to modern contamination. Reads mapping to the grape genome were analyzed for aDNA damage with bamdamage (Malaspinas et al., 2014).

2.6.3. Analysis of plastome data

As discussed previously, the high frequency of odins related to the grape plastome complicates bioinformatic analyses, mandating very strict filtering protocols. Therefore, phylogenetic analyses were restricted to portions of the plastome where reads can be mapped unambiguously. The "mappability" (Li and Freudenberg, 2014) of the plastome was assessed *in silico* by mapping fragmented portions of the grape genome using the above-described methods. Based on this experiment, 40 bp and 100 bp reads can be unambiguously mapped to only 23.4 and 44.3% of the plastome, respectively. Due to the high level of genetic diversity in grapevine (Bacilieri et al., 2013; Myles et al., 2011), it would be unsurprising if different grape individuals had varying frequencies of odins, a factor that may cause an overestimate of how much of the plastome can reliably be characterized in specimens with highly fragmented DNA.

To investigate the relationship of archaeological samples to modern specimens, plastome sequences were acquired from a *Vitis* sequencing project initiated by IGA comprising shotgun sequencing of 52 *sativa* varieties and seven accessions of spontaneous grapevines, purportedly *sylvestris*. Further sequencing reads of Sultanina and Tannat varieties and another *Vitis* species (*Vitis rotundifolia* Michx.) were retrieved from the NCBI Sequence Read Archive (SRR924200, SRR863595, SRR863618, SRR094945). Reads were mapped with BWA aln v0.7.12 (Li and Durbin, 2009) to the reference grape plastome (Jansen et al., 2006). Reads best mapping to the plastome were retrieved and filtered using the same parameters used for the ancient samples. Consensus sequences were generated for both ancient and modern samples using a simple majority rule as implemented in ANGSD (Korneliussen et al., 2014). Bases with quality <20 and sites with a depth of coverage <5 were not considered in the consensus sequences. Consensus sequences and a set of reference sequences (4 *sativa* cultivars (AB856289.1, AB856290.1, AB856291.1, DQ424856.1) and 2 *Vitis rotundifolia* (NC_023790.1, KF976463.1)) were aligned using MAFFT v7.221 (Katoh and Standley, 2013) with default parameters (Data S1). To avoid erroneous interpretations caused by missing data, analyses were restricted to samples with <50% missing data and to plastome nucleotides with <20% missing data across samples. A maximum-likelihood tree was generated based on the final alignment (48,750 nucleotides) using MEGA 6 (Tamura et al., 2013). We note that we did not proactively remove damaged nucleotide residues at the 5′ and 3′ ends of molecules, as done in some studies (Schubert et al., 2012), but rather relied on the 5 × depth of coverage requirement to minimize biases introduced by DNA damage and sequencing errors. The same phylogenetic relationships between samples were observed when putative DNA damage was removed from sequencing reads (data not shown), but this approach reduced the length of the final alignment.

Modern and ancient plastome data were directly compared with Arroyo-García et al.'s (2006) dataset to explore their hypothesis that many Western European *sativa* lineages descend from Western European *sylvestris* populations. Following Arroyo-García et al.'s chlorotype naming conventions, we determined the expected chlorotype of modern varieties in our dataset (Table S3), assuming that a particular plastome sequence would be maintained in a vegetatively propagated lineage, such as chlorotype A in all Pinot noir/blanc/gris accessions. Using a median joining network constructed in PopART (Bandelt et al., 1999; Leigh, 2016), we then investigated whether archaeological samples matched chlorotype A, the group purportedly introduced into *sativa* from Western European *sylvestris* populations, and other major chlorotypes (B, C, and D).

2.6.4. Analysis of nuclear DNA bycatch

Targeted capture experiments enrich for specific regions of interest, and despite a sequence of washing steps in which off-target DNA is discarded, some non-target DNA inevitably is retained and sequenced. For samples with relatively high endogenous content, this "bycatch" is prone to remain high, making it possible to investigate endogenous nuclear aDNA.

To explore how nuclear DNA might be exploited in future grape aDNA research, nuclear aDNA was compared to a *Vitis* SNP database (Myles et al., 2010), consisting of data for 10 *sativa* varieties, and 7 wild *Vitis* species. Raw data were obtained from the NCBI sequence read archive (SRA) and mapped to the grape reference genome

Table 2
Sequencing data for ancient and modern samples. The total number of trimmed and mapped reads are reported for aggregated capture and shotgun sequencing data for each sample. Nuclear and plastome read counts after filtering signify reads with mapping quality ≥30 and after PCR duplicate removal. Due to enrichment, large numbers of PCR clones are expected for on-target DNA, leading to relatively few unique reads after removal of PCR duplicates. Plastome analyses were calculated on samples with an average plastome depth of coverage (DOC) > 5 and with <50% missing data for the portion of the plastome that can be characterized with short reads, equivalent to >30% coverage. Nuclear genome analyses were performed on samples >0.05 DOC on the nuclear genome.

Sample	Trimmed reads	Reads mapping grape genome	Plastome reads after filtering	Nuclear reads after filtering	Mean length of endogenous DNA	Percent of plastome covered	DOC on plastome sites	Percent of nuclear genome covered	DOC on nuclear genome	Analyses
Areni-1	84,754,174	35,139,859	32,395	702,313	57.8	35.1%	39.3	6.43%	0.074	Plastome and nuDNA
Areni-2	87,441,116	276,377	1492	9521	71.3	22.6%	3.3	0.09%	0.001	
Areni-3	105,200,858	15,838	1269	3756	77.0	20.3%	3.2	0.03%	0.000	
Çamaltı Burnu	31,831,414	480,287	4314	4687	71.8	29.3%	6.6	0.00%	0.000	
Colosseum-1	72,093,339	7,929,373	13,648	373,428	78.1	39.6%	18.2	4.90%	0.057	Plastome and nuDNA
Colosseum-2	86,844,481	14,870,870	19,604	2,366,792	86.3	41.1%	25.7	23.73%	0.414	Plastome and nuDNA
Feral	76,954,110	70,743,134	100,153	19,164,801	89.0	50.8%	114.6	66.91%	3.432	Plastome and nuDNA
Ferrara-1	43,413,589	9,074,197	15,963	55,035	59.9	32.2%	22.5	0.38%	0.005	Plastome only
Ferrara-2	59,611,554	820,721	18,004	48,319	80.7	33.1%	28.4	0.44%	0.005	Plastome only
Hatis	52,319,129	47,811,471	99,526	12,931,345	88.8	50.5%	114.5	63.87%	2.300	Plastome and nuDNA
Lugo	71,653,561	29,219,840	72,089	1,459,686	81.5	44.1%	93.8	14.18%	0.226	Plastome and nuDNA
Modena-1	65,272,320	18,157,046	27,164	1,672,173	78.3	41.6%	33.8	18.41%	0.264	Plastome and nuDNA
Modena-2	35,891,843	54,954	407	535	65.6	10.0%	1.7	0.00%	0.000	
Modena-3	69,492,085	2,929,118	10,630	60,582	59.1	32.1%	14.0	0.55%	0.006	Plastome only
Modena-4	57,051,083	1,486,662	2628	137,805	71.4	25.6%	4.8	1.81%	0.020	
Modena-5	65,603,045	45,384,917	83,072	2,389,154	70.3	45.1%	106.6	19.69%	0.326	Plastome and nuDNA
Parma	48,739,726	1,797,117	26,825	64,148	71.7	37.5%	36.8	0.35%	0.004	Plastome only
Sa Osa-1	27,371,201	960,670	7553	70,993	70.2	33.0%	11.2	0.83%	0.009	Plastome only
Sa Osa-2	53,823,743	959,413	18,022	106,429	73.9	35.6%	24.0	1.23%	0.013	Plastome only
Sassari-1	14,651,811	10,033,212	80,448	786,978	82.0	43.9%	106.0	8.16%	0.111	Plastome and nuDNA
Sassari-2	122,280,788	14,624,903	46,237	2,460,081	83.5	43.2%	60.6	22.39%	0.411	Plastome and nuDNA
Scorpo-1	58,095,854	1,303,288	11,957	186,118	79.0	30.1%	19.8	2.47%	0.028	Plastome only
Scorpo-2	31,436,322	320,113	261	3335	67.0	6.9%	1.6	0.04%	0.000	
Shivta-1	70,607,424	375,529	4439	27,746	52.2	25.6%	6.9	0.22%	0.002	
Shivta-2	77,745,215	33,244	1044	2414	68.2	19.9%	2.4	0.01%	0.000	
Shivta-3	28,370,120	7798	92	760	45.8	1.9%	1.4	0.00%	0.000	
Torre Montello-1	37,885,454	2,788,571	21,044	97,849	75.7	39.3%	27.0	0.91%	0.010	Plastome only
Torre Montello-2	79,412,754	3,655,461	5504	363,097	71.3	31.9%	8.0	4.54%	0.052	Plastome and nuDNA
York	20,187,997	14,476,700	61,278	758,503	67.0	38.1%	75.0	7.24%	0.087	Plastome and nuDNA

v12x0 as described by Myles et al. (2010). Genotypes were called using GATK HaplotypeCaller in sites with depth of coverage >3. After removing non-polymorphic sites, the remaining 4602 SNPs were used as a reference panel for our samples. As a result of the low amount of endogenous DNA and the bias caused by enrichment of plastid DNA, the number of unambiguously mapped reads to the nuclear genome in the samples is variable. Depth of coverage ranged from 0.414× to <0.001× for ancient samples. In order to merge low depth data with the reference panel, a random sampling approach was used. For each site in the reference panel, a random base with quality >20 was sampled, and only samples with depth of coverage >0.05× were considered. Multidimensional scaling (MDS), as implemented in bammds (Malaspinas et al., 2014), was used to explore the relationships between the reference panel each ancient sample. Due to the limited amount of data for the SNP loci, it was not possible to correct for DNA damage by truncating the ends of reads. However, because ancient samples were tested against the modern reference database individually, we have avoided introducing an inherent bias that could cause ancient samples to appear more closely related to one another. Individual MDS plots (not shown) were merged into a single plot using a Procrustes transformation implemented in the R package Vegan (Oksanen et al., 2016).

To understand the effect of missing data and to determine how much data are required for reliable genetic characterizations, an *in silico* downsampling experiment was conducted using publicly available sequencing data from a Tannat cultivar (Da Silva et al., 2013) and the two modern Armenian samples from our dataset. Raw data for the Tannat sample was downloaded from the NCBI SRA repository and mapped using the same parameters used for ancient samples. The three samples were evaluated at different depths of coverage of the nuclear genome (0.01–10×) in MDS (Fig. S1). We observed that a depth ≥0.05×, the samples become correctly oriented in MDS-space along a principal axis separating Eastern and Western European *sativa* varieties. At much greater sequencing depths (≥0.5×) the placement within MDS begins to become stable, permitting more confident genetic affiliations. However, given that the ancient samples only reached a maximum of 0.414×, we only infer whether individual samples are more closely related to Eastern or Western European cultivars in Myles et al.'s (2010) dataset.

3. Results and discussion

3.1. aDNA authenticity

BLAST results of 10,000 reads from each library demonstrate that DNA recovered from the specimens comes from a variety of taxa, particularly bacteria, archaea and fungi. The amount of plant DNA was highly variable in archaeological samples, ranging from 0.08 to 73.16% (mean = 14.95%) (Table S2). These results include both shotgun sequencing and capture enrichment. However, the enrichment process skews the composition of the libraries toward plant DNA, and potentially genetic sequences of organisms with a shared evolutionary history, such as cyanobacteria (McFadden, 2001; Sagan, 1967). Therefore, the most relevant metagenomic data come from samples that were shotgun-sequenced (Fig. 2). For these samples, the amount of plant DNA is low, consistent with expectations for archaeobotanical remains that have been preserved via desiccation and waterlogging. An unusually high proportion of animal DNA occurs in Ferrara-1 (15% of reads), and deeper investigation determined 89% of this animal DNA matches *Protopolystoma xenopodis*, a flatworm. Given the waterlogged preservation of Ferrara-1, and the fact that many flatworm species are aquatic, it is conceivable this is a reliable signal of the source of DNA in the sample and an important insight into the metagenomic complexity of archaeobotanical remains. The high level of unassigned DNA in the samples, up to 82%, is a factor of database limitations and a tendency for short DNA molecules to match the genomes of divergent taxa, preventing the software from reaching an unambiguous taxonomic assignment.

The average length of endogenous DNA in archaeological samples ranged 45.8–86.3 bp, values generally consistent with lengths observed in other aDNA projects. We note that the sequencing modes used limit the maximum read length that can be observed, but the lengths distributions suggest endogenous DNA follows the expected patterns for samples of their age and preservation environments (Fig. S2). Nucleotide misincorporation patterns in aDNA libraries are expected to differ according to the first polymerase used to amplify the library. aDNA libraries initially amplified with AmpliTaq Gold are expected to have high rates of C-to-T transition at 5′ ends of DNA and corresponding G-to-A transitions at 3′ ends (Briggs et al., 2007; Brotherton et al., 2007), and these expectations are confirmed for such samples (Fig. S2). Due to the stalling of Phusion polymerase on deaminated cytosine (Rasmussen et al., 2010), libraries initially amplified with this polymerase are not expected to have the full damage pattern. For example, Areni-1 was amplified with Phusion polymerase and the first and last positions of the DNA strands show a drop in DNA damage, as opposed to the highest levels seen in Areni-3, a sample only amplified with AmpliTaq Gold. Several samples (e.g. Çamaltı Burnu, Ferrara-1, Modena-2, Scorpo-2) show a noisy damage signal, with little evidence for elevated misincorporations at the 5′ and 3′ ends. All of these samples were processed toward the beginning of the project, with Y-shaped adapters and Phusion polymerase, factors that may have limited the amount of endogenous DNA recovered. However, noisy signals appear to also be common among samples with extremely low levels of endogenous DNA (Areni-2 and Areni-3). As expected, the two modern leaves show extremely low nucleotide misincorporation values (Fig. S2).

3.2. Enrichment of targeted loci

The most accurate assessment of capture success is provided by the five samples that were both shotgun sequenced and captured. For these samples, the portion of endogenous DNA that share homology with the probes increases for three specimens, with enrichments in on-target DNA ranging from 9.8 to 66-fold (Table 3). The two samples that do not show enrichment, Shivta-2 and Areni-3, started with extremely low levels of endogenous DNA (<0.05%), and may indicate a threshold exists where targeted capture is ineffective. An alternative interpretation is that these scant reads could be microbial in origin, with sequences similar to low complexity regions of the grape plastome, resulting in an inability to increase endogenous content. The precise enrichment factors are not available for the other samples and the objectives of this study were not focused on methodological improvements (see Ávila-Arcos et al., 2011; Ávila-Arcos et al., 2015; Enk et al., 2013, Enk et al., 2014). However, it is possible to infer whether the capture experiments had the desired effects of increasing the amount of plastome-like DNA in the enriched libraries beyond what would be expected in a modern grape sample. For example, the sequencing of a modern Tannat variety leaf yielded ~1% plastid DNA, while in our captured samples, 35–98% of the grape DNA is plastome-like (Fig. S3), suggesting that capture succeeded, albeit at varying degrees, in the ancient samples.

3.3. Relationships of modern and ancient plastomes

Placing archaeological grape plastomes into the larger context of

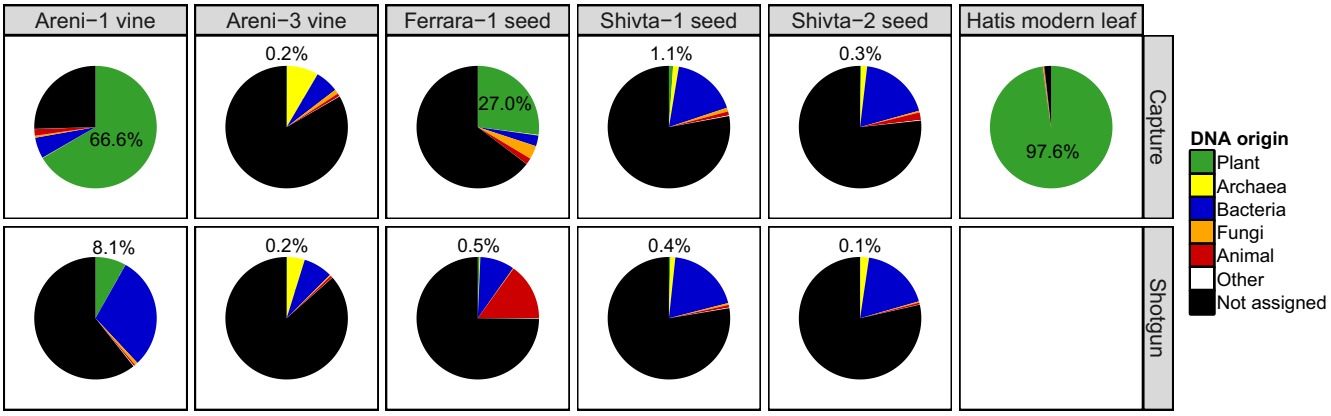

Fig. 2. BLAST determination of DNA sources in representative samples. The percentage of DNA assigned to plants (Viridiplantae) is displayed for each sample.

Table 3
Enrichment of plastome-like DNA as a proportion of reads mapping to grape genome. Two samples, Areni-3 and Shiva-2, do not demonstrate enrichment in plastome-like DNA, likely due to very low endogenous content.

Sample	Percent of endogenous reads on target		Fold enrichment
	Shotgun	Capture	
Areni-1	4.8%	93.8%	19.40
Areni-3	57.0%	47.2%	0.83
Ferrara-1	1.4%	91.1%	66.05
Shivta-1	8.2%	80.3%	9.83
Shivta-2	79.6%	76.6%	0.96

modern *sativa* and *sylvestris* provides an opportunity to investigate Arroyo-García et al.'s (2006) hypothesis of a second domestication of grapevine in Western Europe and Myles et al.'s (2011) hypothesis of introgression of Western *sylvestris* into modern *sativa*. In Arroyo-García et al.'s study, the researchers tested nine chloroplast microsatellite markers in 1201 *sylvestris* and *sativa* accessions, leading to the definition of eight chlorotype groups (network reproduced in Fig. 3A). In their analysis, the frequency of chlorotype A was high among Western European *sylvestris* and *sativa* specimens, but not high among Eastern *sylvestris* populations, suggesting a secondary domestication event or hybridization with local wild grapes. More recently, the presence of chlorotype A was also reported in spontaneous grapevines from Eastern Europe and the Caucasus, although at lower frequencies (Dzhambazova et al., 2015; Pipia et al., 2012). Our data are not based on plastome microsatellites, partly because short tandem repeat markers cannot be effectively retrieved from most ancient samples (Arandjelovic and Thalmann, 2001). However, a median joining network generated on the modern samples in our dataset results in a similar topology for plastome data (Fig. 3B), with a major axis separating the A chlorotype and the D and E groups. The DNA alignments used in our dataset should theoretically provide increased phylogenetic resolution, but various factors may lead to inconsistencies between Fig. 3A and B. For example, either dataset may contain unrecognized odins that would be interpreted as novel haplotypes, or our dataset may fail to describe diversity in portions of the plastome that cannot be recovered with short DNA read technologies.

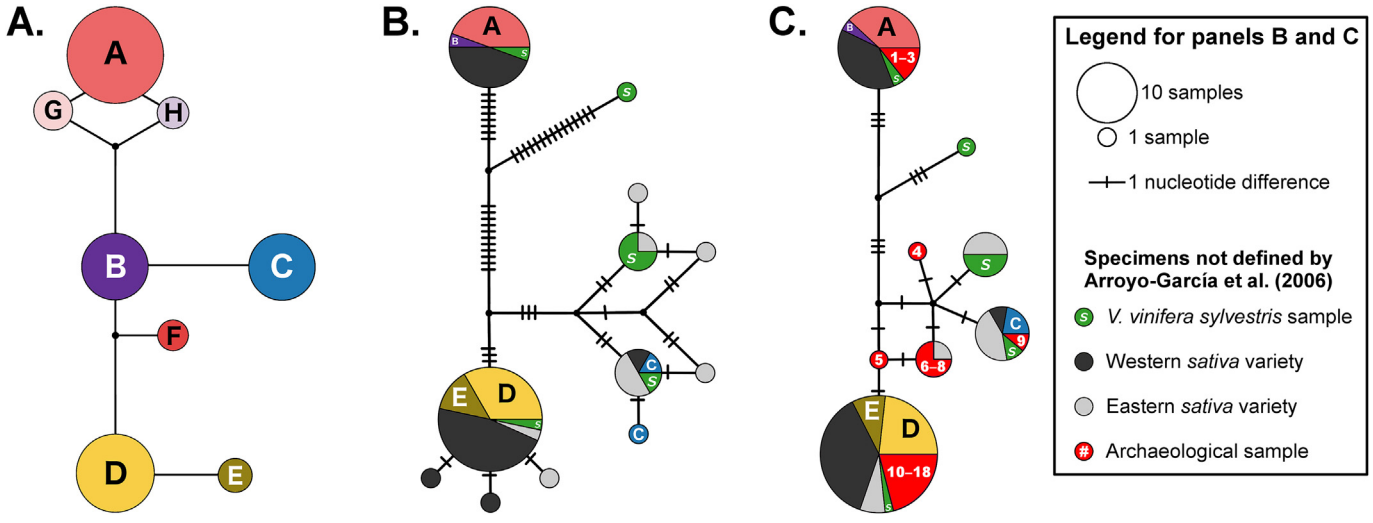

Fig. 3. A. Median network of modern *sativa* and *sylvestris* chlorotypes reproduced from Arroyo-García et al. (2006), defined by plastome microsatellites using their naming conventions. B. Median joining network on modern plastome data for modern samples in this study, with chlorotypes determined for varieties also present in Arroyo-García et al.'s (2006) dataset; no varieties in our dataset had been defined as belonging to the F, G, or H chlorotypes. Eastern and western *sativa* varieties are defined in Table S3. C. Median joining network with modern and ancient samples, with reduced resolution in the tree typology due to greater amounts of missing data. Archaeological samples were only included if >30% of the plastome was recovered and are listed by number. 1–3: Ferrara-1, Sassari-2, and York; 4: Scorpo-1; 5: Torre Montello-2; 6–8: Ferrara-2, Parma, and Sa Osa-2; 9: Areni-1; 10–18: Colosseum-1, Colosseum-2, Lugo, Modena-1, Modena-3, Modena-5, Sa Osa-1, Sassari-1, and Torre Montello-1.

Regardless, the central question of whether archaeological samples closely aligned with the A chlorotype can be confidently addressed (Fig. 3C). Indeed, samples from Northern Italy, Sardinia and UK (Ferrara-1, Sassari-2, and York) fall in this divergent clade, each of which dates to the medieval era, and may represent evidence of the hypothesized genetic introgression of *sylvestris* into *sativa* in Western Europe. Our specimens from the Italian Peninsula predating medieval times do not show evidence of the A chlorotype. A vine from Areni-1, Armenia, dating back to 1200 CE, has the chlorotype C. Several ancient samples from Southern, Central, and Northern Italy and Sardinia have the chlorotype D, including Roman-age seeds from the Colosseum and a Sardinian pip dating to ca. 1200 BCE (Sa Osa-1). Three ancient accessions from Northern Italy and Sardinia (Ferrara-2, Parma, and Sa Osa-2) have the same chlorotype as a modern Armenian *sativa* specimen collected in Hatis, Armenia, but no other characterized varieties. Two ancient specimens (Scorpo-1 and Torre Montello-2) from the southeastern tip of the Italian peninsula do not have a counterpart in modern chlorotypes. The first, Scorpo-1 dates 686–876 CE and is most closely related to chlorotype C. The sample Torre Montello-2 is from 357–164 BCE and has a chlorotype ancestral to chlorotype D.

3.4. Practicability of Vitis plastome for aDNA research

While the above results demonstrate that some pieces of information can be gleaned from plastome sequences of archaeological grape specimens, we argue the genetic locus is of limited phylogenetic resolution for grape aDNA research. The plastome network generated from modern samples (Fig. 3B) displays a relatively limited amount of genetic diversity, suggesting phenotypically and genotypically divergent lineages of grapes are not differentiated at the plastome level, ultimately diminishing the value of the grape plastome as a suitable locus for intraspecific phylogenetic analyses.

Research conducted in conjunction with this study has shown that practically the entire grape plastome sequence has been transferred into the mitochondrial and nuclear genomes, often in stretches thousands of nucleotides in length (Samaniego Castruita et al., 2015). The ubiquity of these odins posed a serious challenge to the inference of ancient plastomes in this dataset, and the promiscuous nature of plastomes may be an underappreciated complication in many research projects. For example, the chlorotype network defined by Arroyo-García et al. (2006) is based on plastome microsatellite markers, but it is conceivable that the nuclear or mitochondrial copies of the markers also amplify with PCR, leading to erroneous interpretations. For modern grape specimens that have been shotgun sequenced, it can be possible to identify and exclude most odin-derived mutations based on the premise that odins are expected to be in lower frequency than plastome sequences (Samaniego Castruita et al., 2015). However, in the case of ancient samples, we found this approach ineffective, because minor changes in parameters thresholds could yield incompatible phylogenies (data not shown). The challenge likely is due to a number of factors, such as the relatively low sequencing depth achievable for many ancient samples, variability in the number of plastids in seeds and other tissues, and environmental contamination by species such as cyanobacteria. In addition, for samples enriched for plastome DNA, PCR of all captured loci and subsequent removal of duplicate reads can further obscure the true proportion of plastid DNA versus odins (see Supplemental Figs. 4 and 5 in Samaniego Castruita et al., 2015). Thus, we have taken a relatively conservative approach for the plastome characterization of ancient samples, ensuring that chlorotype assignments are well supported and unaffected by slight adjustments in thresholds for missing data and SNP detection.

3.5. Analysis of nuclear DNA bycatch

In contrast to the challenges posed by the plastome data, nuclear DNA recovered from the archaeological samples indicates the nuclear genome is far more useful and informative about grape genetic diversity and relationships, although data are limited since these loci were not purposefully enriched via targeted capture. Archaeological samples with >0.05× coverage on the nuclear genome were compared against a reference panel of *Vitis* species and *sativa* varieties published by Myles et al. (2010) using MDS (Fig. 4). The general pattern of *Vitis* species plotting along the x-axis and *sativa* varieties along the y-axis recapitulates Myles et al.'s (2010) figures. Ten of the archaeological samples had sufficient data to be tested against the SNP database, at a degree of certainty where samples can be classified as being more genetically similar to Eastern or Western European grapevines.

In several cases, samples from the same archaeological site plot in divergent locations in the MDS-space. For example, the medieval Sassari-1 and Sassari-2 samples associate with Eastern and Western European grapes, respectively, in accord with their chlorotypes. By investigating the timing and prevalence of archaeological grapes with a Western European signature in an extended set of samples, the hypothesis for a local domestication of *sativa* can be tested in detail. Unlike the plastome, which only traces one genetic linage, the relative proportion of Eastern and Western European ancestry can be understood from the nuclear genome, eventually differentiating between scenarios of secondary domestication, hybridization between Eastern and Western European lineages, and use of local, wild grapes.

Important data about the history of viticulture in Southern Italy comes from the Torre Montello-2 sample, recovered in the *Magna Graecia* settlement of Taranto and dating to 357–164 BCE. The sample plots among Eastern European grapes in the MDS-space, a finding that is consistent with ongoing morphological analyses of 32 archaeological pips from the site. Digital images of the archaeological samples have been compared against a modern reference collection using Elliptic Fourier Analyses and multivariate statistical methods. Preliminary results show an affinity of the archaeological sample with modern Greek cultivars (Grasso, Fiorentino, study in progress), suggesting a possible key role of Greek settlers in the spread of viticulture.

Another intriguing insight of the nuclear DNA comes from two seeds excavated at the Colosseum in Rome. Both samples plot among modern Western European *sativa* typical of northern cool climates. Yet they are found in different locations in MDS-space, one falling near the space delimited by Ehrenfelser, Gewurztraminer, and Riesling and the other situated closely with Pinot noir. However, the difference reflected by this MDS-space is minor, since Pinot noir and Gewurztraminer have a parent-offspring relationship (Myles et al., 2011). It remains clear that the ancient Colosseum specimens fall short of the MDS-space populated by the modern varieties typical of warmer climates in Mediterranean shores and the Balkans, sampled by Myles et al. (2010). Future work with expanded datasets and deeper sequencing are necessary to explore how much data are necessary to reach reliable genetic distances to modern specimens. Still, nuclear DNA from these seeds gives a preliminary indication that it is possible to genetically observe ancient cultivation of different varieties. While the use of different varieties of grapes by the Romans is known from historic accounts (Bowers et al., 1999), such accounts are often incomplete and oral

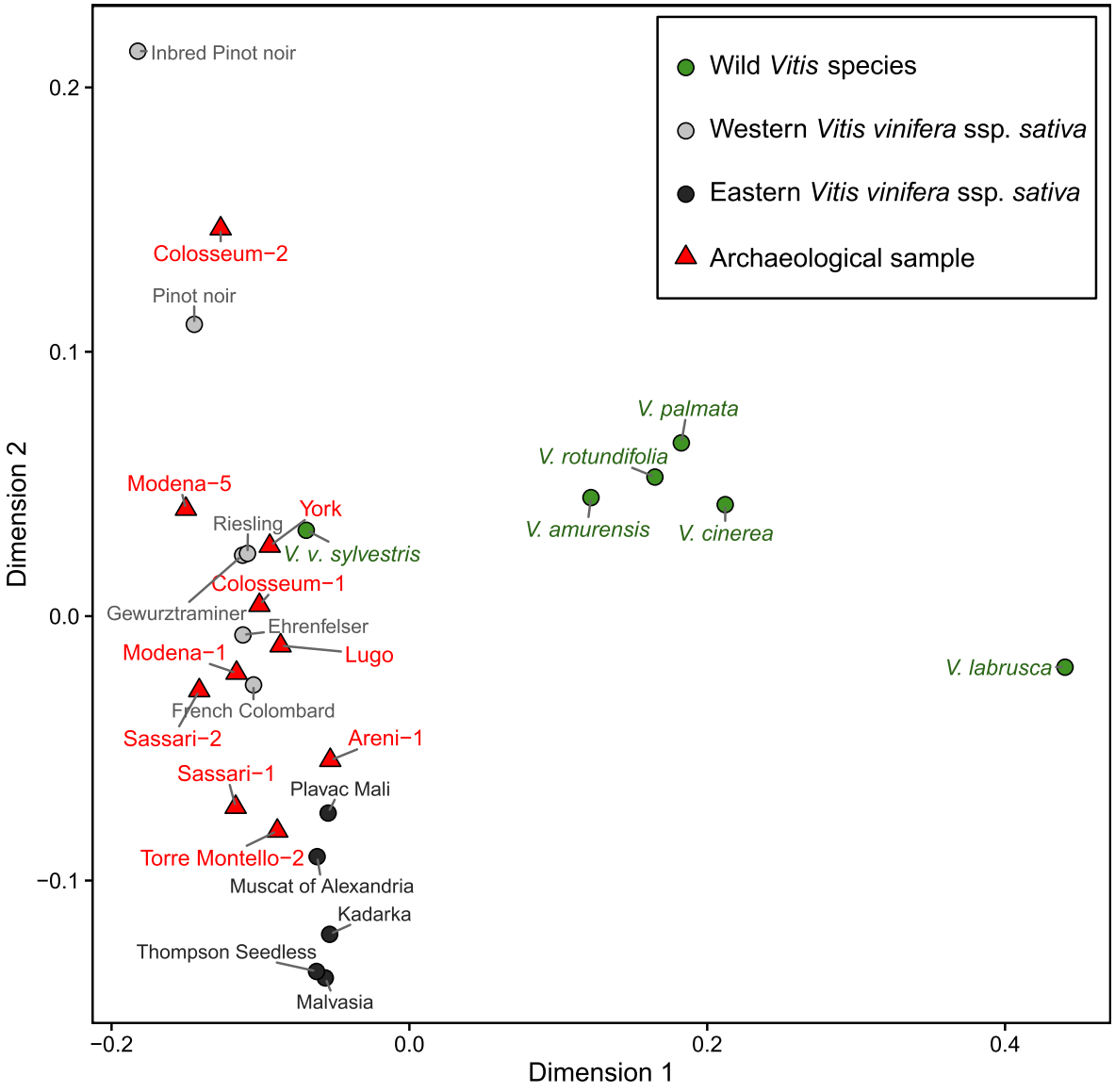

Fig. 4. Multidimensional scaling plot of informative nuclear SNPs in archaeological samples and *Vitis* species and cultivars tested by Myles et al. (2010).

traditions about cultivars are frequently overturned with pedigree analyses (Ibáñez et al., 2009; Myles et al., 2011; Sefc et al., 1998; Vouillamoz and Grando, 2006).

3.6. Individual seed vs. bulk extractions

We did not find any consistent differences in the quality of data resulting from archaeological seeds that were extracted individually or in bulk. With each approach, some samples yielded sufficient endogenous DNA for phylogenetic analyses while other samples were effectively devoid of endogenous DNA. This suggests that sample preservation is paramount, and increasing the amount of tissue in aDNA extractions will frequently have no positive effect. Furthermore, our data indicate that samples from the same stratigraphic context are often genetically distinct, such as Sassari-1 and Sassari-2, and would yield a mixed genetic signal if processed as one bulk sample. Therefore, we recommend archaeological grape seeds to be tested individually.

4. Conclusions

Organellar genomic loci have been a central focus in aDNA research because they are present in high copy numbers and inherited without recombination, making genetic lineages readily traceable in degraded specimens. Although complete plastome sequences are expected to be rewarding for aDNA research, we find the grape plastome hardly tractable for intraspecific aDNA studies, due to limited phylogenetic resolution and high rates of horizontal gene transfer. In a broader context, plastome research in other archaeological crops should be approached cautiously as these pitfalls are likely underappreciated in other species. In contrast to the limitations of the plastome, the grape nuclear genome shows great promise for archaeological samples, and preliminary analyses demonstrate that individual grape specimens can be compared against modern varieties, showing their genetic affiliations. With expanding databases of reference *sativa* and *sylvestris* accessions, it will become possible to characterize relationships between ancient samples and recent varieties, as well as chronicle the movement of

cultivated grapes away from their center(s) of domestication into regions where hybridization with wild *sylvestris* populations may have occurred.

Acknowledgments

This work was supported by the Lundbeck Foundation (R52-A5062), the Danish Council for Independent Research (10-081390), the Danish National Research Foundation (DNRF94), and the EU Marie Curie mobility program (237227). Illumina sequencing of modern varieties was funded by the Italian Ministry of Agriculture (Vigneto, Valorisation of the main Italian native vines and their terroir) and the European Research Council (Novabreed, Novel variation in plant breeding and the plant pan-genomes, 294780). Funds for C14 dating of samples from Areni-1 were provided by an NSF Faculty Early CAREER Award (No. 1136516) granted to Alexia Smith. Nathan Wales was funded by a dissertation fellowship from the American Scandinavian Foundation. The authors thank the Danish National High-throughput Sequencing Centre for assistance in generating Illumina data. Christina Warinner kindly provided the enzyme cocktail that was used to extract DNA from two grapevine samples. Nathan Wales thanks Daniel Adler, Sally McBrearty, Natalie Munro, and Linda Strausbaugh for their comments and assistance during a phase undertaken during his PhD research. Girolamo Fiorentino thanks the Soprintendenza Archaeologica della Puglia for the Taranto samples and the archaeologists of other sites for the archaeobotanical materials.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jas.2016.05.014.

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Arandjelovic, M., Thalmann, O., 2001. Microsatellites in Historic and Ancient DNA. eLS, John Wiley & Sons, Ltd.

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

Arthur, P., Fiorentino, G., Leo Imperiale, M., 2008. L'insediamento in Loc. Scorpo (Supersano, LE) nel VII-VIII secolo. La scoperta di un paesaggio di età altomedievale. Archeol. Mediev. 35, 365–380.

Ávila-Arcos, M., Cappellini, E., Romero-Navarro, J.A., Wales, N., Moreno-Mayar, J.V., Rasmussen, M., Fordyce, S.L., Montiel, R., Vielle-Calzada, J.-P., Willerslev, E., Gilbert, M.T.P., 2011. Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA. Sci. Rep. 1, 74.

Ávila-Arcos, M.C., Sandoval-Velasco, M., Schroeder, H., Carpenter, M.L., Malaspinas, A.-S., Wales, N., Peñaloza, F., Bustamante, C.D., Gilbert, M.T.P., 2015. Comparative performance of two whole genome capture methodologies on ancient DNA Illumina libraries. Methods Ecol. Evol. 6, 725–734.

Bacilieri, R., Lacombe, T., Le Cunff, L., Di Vecchi-Staraz, M., Laucou, V., Genna, B., Peros, J.-P., This, P., Boursiquot, J.-M., 2013. Genetic structure in cultivated grapevines is linked to geography and human selection. BMC Plant Biol. 13, 25.

Bandelt, H.-J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Barnard, H., Dooley, A.N., Areshian, G., Gasparyan, B., Faull, K.F., 2011. Chemical evidence for wine production around 4000 BCE in the Late Chalcolithic Near Eastern highlands. J. Archaeol. Sci. 38, 977–984.

Beck, H.T., 2004. Caffeine, alcohol, and sweeteners. In: Prance, G., Nesbitt, M. (Eds.), The Cultural History of Plants. Routledge, New York, pp. 173–190.

Beridze, T., Pipia, I., Beck, J., Hsu, S.-C., Gamkrelidze, M., Gogniashvili, M., Tabidze, V., This, P., Bacilieri, R., Gotsiridze, V., Glonti, M., Schaal, B., 2011. Plastid DNA sequence diversity in a worldwide set of grapevine cultivars (*Vitis vinifera* L. subsp. *Vinifera*). Bull. Georgian Natl. Acad. Sci. 5, 98–103.

Birky Jr., C.W., 2001. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. Annu. Rev. Genet. 35, 125–148.

Bock, R., 2007. Structure, function, and inheritance of plastid genomes. In: Bock, R. (Ed.), Cell and Molecular Biology of Plastids. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 29–63.

Bosi, G., Bandini Mazzanti, M., 2013. Informazioni etnobotaniche dai reperti carpologici del pozzo: risultati di un saggio preliminare. In: Rovina, D., Fiori, M. (Eds.), Sassari: Archeologia Urbana, Felici Editore. Pisa, Italy, pp. 86–92.

Bosi, G., Mercuri, A.M., Guarnieri, C., Bandini Mazzanti, M., 2009. Luxury food and ornamental plants at the 15th century a.d. Renaissance court of the Este family (Ferrara, northern Italy). Veg. Hist. Archaeobot. 18, 389–402.

Bosi, G., Bandini Mazzanti, M., Florenzano, A., Massamba N'siala, I., Pederzoli, A., Rinaldi, R., Torri, P., Mercuri, A.M., 2011. Seeds/fruits, pollen and parasite remains as evidence of site function: piazza Garibaldi – Parma (N Italy) in Roman and Mediaeval times. J. Archaeol. Sci. 38, 1621–1633.

Bosi, G., Benatti, A., Rinaldi, R., Dallai, D., Santini, C., Carbognani, M., Tomaselli, M., Bandini Mazzanti, M., 2015a. The memory of water: archaeobotanical evidence of wetland plants from Modena (Emilia-Romagna, northern Italy) and palaeoecological remarks. Plant Biosyst. 149, 144–153.

Bosi, G., Mercuri, A.M., Bandini Mazzanti, M., Florenzano, A., Montecchi, M.C., Torri, P., Labate, D., Rinaldi, R., 2015b. The evolution of Roman urban environments through the archaeobotanical remains in Modena – northern Italy. J. Archaeol. Sci. 53, 19–31.

Bosi, G., Bandini Mazzanti, M., Montecchi, M.C., Torri, P., Rinaldi, R., 2016. Life a Roman colony North. Italy: ethnobotanical information from archaeobotanical analysis. Submitted.

Bosi, G., 2000. Flora e ambiente vegetale a Ferrara tra il X e il XV secolo attraverso i reperti carpologici dello scavo di corso Porta Reno - via Vaspergolo nell'attuale centro storico (Ph.D thesis). Università di Firenze.

Bouby, L., Figueiral, I., Bouchette, A., Rovira, N., Ivorra, S., Lacombe, T., Pastor, T., Picq, S., Marinval, P., Terral, J.-F., 2013. Bioarchaeological insights into the process of domestication of grapevine (*Vitis vinifera* L.) during roman times in southern France. PLoS One 8, e63195.

Bowers, J., Boursiquot, J.-M., This, P., Chu, K., Johansson, H., Meredith, C., 1999. Historical genetics: the parentage of chardonnay, gamay, and other wine grapes of northeastern France. Science 285, 1562–1565.

Briggs, A.W., Stenzel, U., Johnson, P.L.F., Green, R.E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., Pääbo, S., 2007. Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl. Acad. Sci. 104, 14616–14621.

Bronk Ramsey, C., Lee, S., 2013. Recent and planned developments of the program OxCal. Radiocarbon 55, 720–730.

Brotherton, P., Endicott, P., Sanchez, J.J., Beaumont, M., Barnett, R., Austin, J., Cooper, A., 2007. Novel high-resolution characterization of ancient DNA reveals C > U-type base modification events as the sole cause of post mortem miscoding lesions. Nucleic Acids Res. 35, 5717–5728.

Brown, T.A., Cappellini, E., Kistler, L., Lister, D.L., Oliveira, H.R., Wales, N., Schlumbaum, A., 2015. Recent advances in ancient DNA research and their implications for archaeobotany. Veg. Hist. Archaeobot. 24, 207–214.

Calcagnile, L., Quarta, G., D'Elia, M., Maruccio, L., Gaballo, V., Braione, E., 2011. Datazione al radiocarbonio mediante AMS di resti vegetali. In: Arthur, P., Fiorentino, G., Grasso, A.M., Leo Imperiale, M. (Eds.), La Storia nel Pozzo. Ambiente ed economia di un villaggio bizantino in Terra d'Otranto - Supersano 2007. Unisalento Press, Lecce, pp. 42–45.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinforma. 10, 1–9.

Cappellini, E., Gilbert, M.T.P., Geuna, F., Fiorentino, G., Hall, A., Thomas-Oates, J., Ashton, P., Ashford, D., Arthur, P., Campos, P., Kool, J., Willerslev, E., Collins, M., 2010. A multidisciplinary study of archaeological grape seeds. Naturwissenschaften 97, 205–217.

Carpenter, Meredith L., Buenrostro, Jason D., Valdiosera, C., Schroeder, H., Allentoft, Morten E., Sikora, M., Rasmussen, M., Gravel, S., Guillén, S., Nekhrizov, G., Leshtakov, K., Dimitrova, D., Theodossiev, N., Pettener, D., Luiselli, D., Sandoval, K., Moreno-Estrada, A., Li, Y., Wang, J., Gilbert, M.Thomas P., Willerslev, E., Greenleaf, William J., Bustamante, Carlos D., 2013. Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. Am. J. Hum. Genet. 93, 852–864.

CBOL Plant Working Group, Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S., van der Bank, M., Chase, M.W., Cowan, R.S., Erickson, D.L., Fazekas, A.J., Graham, S.W., James, K.E., Kim, K.-J., Kress, W.J., Schneider, H., van AlphenStahl, J., Barrett, S.C.H., van den Berg, C., Bogarin, D., Burgess, K.S., Cameron, K.M., Carine, M., Chacón, J., Clark, A., Clarkson, J.J., Conrad, F., Devey, D.S., Ford, C.S., Hedderson, T.A.J., Hollingsworth, M.L., Husband, B.C., Kelly, L.J., Kesanakurti, P.R., Kim, J.S., Kim, Y.-D., Lahaye, R., Lee, H.-L., Long, D.G., Madriñán, S., Maurin, O., Meusnier, I., Newmaster, S.G., Park, C.-W., Percy, D.M., Petersen, G., Richardson, J.E., Salazar, G.A., Savolainen, V., Seberg, O., Wilkinson, M.J., Yi, D.-K., Little, D.P., 2009. A DNA barcode for land plants. Proc. Natl. Acad. Sci. 106, 12794–12797.

Celant, A., 1998. Ricerche Archeobotaniche nell'area Romana (Ph.D. thesis). Università degli Studi di Roma, La Sapienza.

China Plant BOL Group, Li, D.-Z., Gao, L.-M., Li, H.-T., Wang, H., Ge, X.-J., Liu, J.-Q., Chen, Z.-D., Zhou, S.-L., Chen, S.-L., Yang, J.-B., Fu, C.-X., Zeng, C.-X., Yan, H.-F., Zhu, Y.-J., Sun, Y.-S., Chen, S.-Y., Zhao, L., Wang, K., Yang, T., Duan, G.-W., 2011. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc.

Natl. Acad. Sci. 108, 19641–19646.
Cinquantaquattro, T.E., 2012. Attività della Soprintendenza per i Beni Archeologici della Puglia, Alle origini della Magna Grecia: mobilità migrazioni fondazioni. Atti del cinquantesimo Convegno di studi sulla Magna Grecia, Taranto, 1-4 ottobre 2010. L'Istituto per la Storia e l'Archeologia della Magna Grecia, Taranto, Italy, pp. 1209–1258.
Cooper, A., Poinar, H.N., 2000. Ancient DNA: do it right or not at all. Science 289, 1139.
da Fonseca, R.R., Smith, B.D., Wales, N., Cappellini, E., Skoglund, P., Fumagalli, M., Samaniego, J.A., Carøe, C., Ávila-Arcos, M.C., Hufnagel, D.E., Korneliussen, T.S., Vieira, F.G., Jakobsson, M., Arriaza, B., Willerslev, E., Nielsen, R., Hufford, M.B., Albrechtsen, A., Ross-Ibarra, J., Gilbert, M.T.P., 2015. The origin and evolution of maize in the Southwestern United States. Nat. Plants 1, 14003.
Da Silva, C., Zamperin, G., Ferrarini, A., Minio, A., Dal Molin, A., Venturini, L., Buson, G., Tononi, P., Avanzato, C., Zago, E., Boido, E., Dellacassa, E., Gaggero, C., Pezzotti, M., Carrau, F., 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.
Dabney, J., Meyer, M., Pääbo, S., 2013. Ancient DNA damage. Cold Spring Harb. Perspect. Biol. 5, a012567.
Dzhambazova, T., Tsvetkov, I., Atanassov, I., Rusanov, K., Martínez-Zapater, J.M., Atanassov, A., Hvarleva, T., 2015. Genetic diversity in native Bulgarian grapevine germplasm (*Vitis vinifera* L.) based on nuclear and chloroplast microsatellite polymorphisms, 48, 115–121.
Enk, J., Rouillard, J.-M., Poinar, H., 2013. Quantitative PCR as a predictor of aligned ancient DNA read counts following targeted enrichment. BioTechniques 55, 300–309.
Enk, J.M., Devault, A.M., Kuch, M., Murgha, Y.E., Rouillard, J.-M., Poinar, H.N., 2014. Ancient whole genome enrichment using baits built from modern DNA. Mol. Biol. Evol. 31, 1292–1294.
Erickson, D.L., Smith, B.D., Clarke, A.C., Sandweiss, D.H., Tuross, N., 2005. An Asian origin for a 10,000-year-old domesticated plant in the Americas. Proc. Natl. Acad. Sci. 102, 18315–18320.
Foley, B.P., Hansson, M.C., Kourkoumelis, D.P., Theodoulou, T.A., 2012. Aspects of ancient Greek trade re-evaluated with amphora DNA evidence. J. Archaeol. Sci. 39, 389–398.
Gansauge, M.-T., Meyer, M., 2013. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nat. Protoc. 8, 737–748.
Ghini, G., 1988. Prime indagini archeologiche. In: Reggiani, A.M. (Ed.), Anfiteatro Flavio: Immagine, Testimonianze, Spettacoli. Edizioni Quasar, Rome, pp. 100–105.
Gilbert, M.T.P., Bandelt, H.-J., Hofreiter, M., Barnes, I., 2005. Assessing ancient DNA studies. Trends Ecol. Evol. 20, 541–544.
Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E.M., Brockman, W., Fennell, T., Giannoukos, G., Fisher, S., Russ, C., Gabriel, S., Jaffe, D.B., Lander, E.S., Nusbaum, C., 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat. Biotechnol. 27, 182–189.
Goremykin, V.V., 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.
Grasso, A.M., Fiorentino, G., 2010. Çamaltı Burnu I: Archaeobotanical Analysis (Internal Report). Laboratory of Archaeobotany and Palaeoecology, University of Salento, Italy.
Gray, M.W., 1993. Origin and evolution of organelle genomes. Curr. Opin. Genet. Dev. 3, 884–890.
Green, R.E., Malaspinas, A.-S., Krause, J., Briggs, A.W., Johnson, P.L.F., Uhler, C., Meyer, M., Good, J.M., Maricic, T., Stenzel, U., Prüfer, K., Siebauer, M., Burbano, H.A., Ronan, M., Rothberg, J.M., Egholm, M., Rudan, P., Brajković, D., Kućan, Ž., Gušić, I., Wikström, M., Laakkonen, L., Kelso, J., Slatkin, M., Pääbo, S., 2008. A complete neandertal mitochondrial genome sequence determined by high-throughput sequencing. Cell 134, 416–426.
Haile, J., Holdaway, R., Oliver, K., Bunce, M., Gilbert, M.T.P., Nielsen, R., Munch, K., Ho, S.Y.W., Shapiro, B., Willerslev, E., 2007. Ancient DNA chronology within sediment deposits: are paleobiological reconstructions possible and is DNA leaching a factor? Mol. Biol. Evol. 24, 982–989.
Hansson, M.C., Foley, B.P., 2008. Ancient DNA fragments inside Classical Greek amphoras reveal cargo of 2400-year-old shipwreck. J. Archaeol. Sci. 35, 1169–1176.
Herrmann, R.G., Possingham, J.V., 1980. Plastid DNA — The plastome. In: Reinert, J. (Ed.), Chloroplasts. Springer, Berlin Heidelberg, pp. 45–96.
Hirschfeld, Y., Tepper, Y., 2006. Columbarium towers and other structures in the environs of Shivta. Tel Aviv 33, 83–116.
Hofreiter, M., Serre, D., Poinar, H.N., Kuch, M., Pääbo, S., 2001. Ancient DNA. Nat. Rev. Genet. 2, 353–359.
Hunt, H.V., Lawes, M.C., Bower, M.A., Haeger, J.W., Howe, C.J., 2010. A banned variety was the mother of several major wine grapes. Biol. Lett. 6, 367–369.
Huson, D.H., Mitra, S., Ruscheweyh, H.-J., Weber, N., Schuster, S.C., 2011. Integrative analysis of environmental sequences using MEGAN4. Genome Res. 21, 1552–1560.
Ibáñez, J., Vargas, A.M., Palancar, M., Borrego, J., de Andrés, M.T., 2009. Genetic relationships among table-grape varieties. Am. J. Enol. Vitic. 60, 35–42.
Jaillon, O., Aury, J.M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., Vezzi, A., Legeai, F., Hugueney, P., Dasilva, C., Horner, D., Mica, E., Jublot, D., Poulain, J., Bruyère, C., Billault, A., Segurens, B., Gouyvenoux, M., Ugarte, E., Cattonaro, F., Anthouard, V., Vico, V., Del Fabbro, C., Alaux, M., Di Gaspero, G., Dumas, V., Felice, N., Paillard, S., Juman, I., Moroldo, M., Scalabrin, S., Canaguier, A., Le Clainche, I., Malacrida, G., Durand, E., Pesole, G., Laucou, V., Chatelet, P., Merdinoglu, D., Delledonne, M., Pezzotti, M., Lecharny, A., Scarpelli, C., Artiguenave, F., Pè, M.E., Valle, G., Morgante, M., Caboche, M., Adam-Blondon, A.F., Weissenbach, J., Quétier, F., Wincker, P., 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449, 463–467.
Jansen, R., Kaittanis, C., Saski, C., Lee, S.-B., Tomkins, J., Alverson, A., Daniell, H., 2006. Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids. BMC Evol. Biol. 6, 32.
Japelaghi, R., Haddad, R., Garoosi, G.-A., 2011. Rapid and efficient isolation of high quality nucleic acids from plant tissues rich in polyphenols and polysaccharides. Mol. Biotechnol. 49, 1–9.
Jenster, P.V., Smith, D.E., Mitry, D.J., Jenster, L.V., 2008. The Business of Wine: a Global Perspective. Copenhagen Business School Press, Copenhagen.
Karol, K.G., Arumuganathan, K., Boore, J.L., Duffy, A.M., Everett, K.D., Hall, J.D., Hansen, S.K., Kuehl, J.V., Mandoli, D.F., Mishler, B.D., Olmstead, R.G., Renzaglia, K.S., Wolf, P.G., 2010. Complete plastome sequences of *Equisetum arvense* and *Isoetes flaccida*: implications for phylogeny and plastid genome evolution of early land plant lineages. BMC Evol. Biol. 10, 1–16.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Kistler, L., Montenegro, Á., Smith, B.D., Gifford, J.A., Green, R.E., Newsom, L.A., Shapiro, B., 2014. Transoceanic drift and the domestication of African bottle gourds in the Americas. Proc. Natl. Acad. Sci. 111, 2937–2941.
Kistler, L., Newsom, L.A., Ryan, T.M., Clarke, A.C., Smith, B.D., Perry, G.H., 2015. Gourds and squashes (*Cucurbita* spp.) adapted to megafaunal extinction and ecological anachronism through domestication. Proc. Natl. Acad. Sci. 112, 15107–15112.
Knapp, M., Hofreiter, M., 2010. Next generation sequencing of ancient DNA: requirements, strategies and perspectives. Genes 1, 227–243.
Korneliussen, T.S., Albrechtsen, A., Nielsen, R., 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinforma. 15, 1–13.
Krause, J., Fu, Q., Good, J.M., Viola, B., Shunkov, M.V., Derevianko, A.P., Paabo, S., 2010. The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. Nature 464, 894–897.
Kumar, A., Bennetzen, J.L., 1999. Plant retrotransposons. Annu. Rev. Genet. 33, 479–532.
Leigh, J., 2016. PopART, http://popart.otago.ac.nz.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25, 1754–1760.
Li, W., Freudenberg, J., 2014. Mappability and read length. Front. Genet. 5, 381.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Subgroup, G.P.D.P, 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079.
Lindahl, T., 1993. Instability and decay of the primary structure of DNA. Nature 362, 709–715.
Lindgreen, S., 2012. AdapterRemoval: easy cleaning of next-generation sequencing reads. BMC Res. Notes 5, 337.
Little, D.P., 2014. A DNA mini-barcode for land plants. Mol. Ecol. Resour. 14, 437–446.
Malaspinas, A.-S., Tange, O., Moreno-Mayar, J.V., Rasmussen, M., DeGiorgio, M., Wang, Y., Valdiosera, C.E., Politis, G., Willerslev, E., Nielsen, R., 2014. *bammds*: a tool for assessing the ancestry of low-depth whole-genome data using multidimensional scaling (MDS). Bioinformatics 30, 2962–2964.
Manen, J.-F., 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.
Manen, J.-F., Sinitsyna, O., Aeschbach, L., Markov, A.V., Sinitsyn, A., 2005. A fully automatable enzymatic method for DNA extraction from plant tissues. BMC Plant Biol. 5, 1–9.
McFadden, G.I., 2001. Chloroplast origin and integration. Plant Physiol. 125, 50–53.
McGovern, P.E., Glusker, D.L., Exner, L.J., Voigt, M.M., 1996. Neolithic resinated wine. Nature 381, 480–481.
McGovern, P.E., 2003. Ancient Wine: the Search for the Origins of Viniculture. Princeton University Press, Princeton, NJ.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303.
Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. doi: 10.1101/pdb.prot5448.
Michalovova, M., Vyskot, B., Kejnovsky, E., 2013. Analysis of plastid and mitochondrial DNA insertions in the nucleus (NUPTs and NUMTs) of six plant species: size, relative age and chromosomal localization. Heredity 111, 314–320.
Miller, N.F., 2008. Sweeter than wine? The use of the grape in early western Asia. Antiquity 82, 937–946.
Myles, S., Chia, J.-M., Hurwitz, B., Simon, C., Zhong, G.Y., Buckler, E., Ware, D., 2010. Rapid genomic characterization of the genus *vitis*. PLoS One 5, e8219.
Myles, S., Boyko, A.R., Owens, C.L., Brown, P.J., Grassi, F., Aradhya, M.K., Prins, B., Reynolds, A., Chia, J.-M., Ware, D., Bustamante, C.D., Buckler, E.S., 2011. Genetic

structure and domestication history of the grape. Proc. Natl. Acad. Sci. 108, 3530–3535.
Nergis, G., 2010. The construction of a monastic ship (?) in the Sea of Marmara: Çamaltı Burnu I wreck. In: Pomey, P. (Ed.), Transferts technologiques en architecture navale méditerranéenne de l'Antiquité aux temps modernes: identité technique et identité culturelle. Actes de la Table Ronde d'Istanbul 19-22 mai 2007. Institut Français d'Études Anatoliennes-Georges Dumézil, Istanbul, pp. 151–153.
Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2016. Vegan: Community Ecology Package. R package version 2.3–4 ed.
Palmer, S.A., Smith, O., Allaby, R.G., 2012. The blossoming of plant archaeogenetics. Ann. Anat. - Anatomisch. Anzeiger 20, 146–156.
Pipia, I., Gogniashvili, M., Tabidze, V., Beridze, T., Gamkrelidze, M., Gotsiridze, V., Melyan, G., Musayev, M., Salimov, V., Beck, J., Schaal, B., 2012. Plastid DNA sequence diversity in wild grapevine samples (*Vitis vinifera* subsp. *sylvestris*) from the Caucasus region. Vitis 51, 119–124.
Poinar, H.N., Schwarz, C., Qi, J., Shapiro, B., MacPhee, R.D.E., Buigues, B., Tikhonov, A., Huson, D.H., Tomsho, L.P., Auch, A., Rampp, M., Miller, W., Schuster, S.C., 2006. Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. Science 311, 392–394.
Ramsay, J., Tepper, Y., 2010. Signs from a green desert: a preliminary examination of the archaeobotanical remains from a Byzantine dovecote near Shivta, Israel. Veg. Hist. Archaeobot. 19, 235–242.
Rasmussen, M., Cummings, L.S., Gilbert, M.T.P., Bryant, V., Smith, C., Jenkins, D.L., Willerslev, E., 2009. Response to comment by Goldberg et al. on "DNA from Pre-Clovis human coprolites in Oregon, North America. Science 325, 148-d.
Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J.S., Albrechtsen, A., Moltke, I., Metspalu, M., Metspalu, E., Kivisild, T., Gupta, R., Bertalan, M., Nielsen, K., Gilbert, M.T.P., Wang, Y., Raghavan, M., Campos, P.F., Kamp, H.M., Wilson, A.S., Gledhill, A., Tridico, S., Bunce, M., Lorenzen, E.D., Binladen, J., Guo, X., Zhao, J., Zhang, X., Zhang, H., Li, Z., Chen, M., Orlando, L., Kristiansen, K., Bak, M., Tommerup, N., Bendixen, C., Pierre, T.L., GrÃ¸nnow, B., Meldgaard, M., Andreasen, C., Fedorova, S.A., Osipova, L.P., Higham, T.F.G., Ramsey, C.B., Hansen, T.v.O., Nielsen, F.C., Crawford, M.H., Brunak, S., Sicheritz-Pontén, T., Villems, R., Nielsen, R., Krogh, A., Wang, J., Willerslev, E., 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature 463, 757–762.
Reeves, B., 2006. 62–68 Low Petergate, York: Assessment Report on an Archaeological Excavation. York Archaeological Trust, York, England.
Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Bronk Ramsey, C., Buck, C.E., Cheng, H., Edwards, R.L., Friedrich, M., Grootes, P.M., Guilderson, T.P., Haflidason, H., Hajdas, I., Hatté, C., Heaton, T.J., Hoffmann, D.L., Hogg, A.G., Hughen, K.A., Kaiser, K.F., Kromer, B., Manning, S.W., Niu, M., Reimer, R.W., Richards, D.A., Scott, E.M., Southon, J.R., Staff, R.A., Turney, C.S.M., van der Plicht, J., 2013. IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 Years cal bp. Radiocarbon 55, 1869–1887.
Reitz, E., Shackley, M., 2012. The Processes by Which Archaeological Sites Form, Environmental Archaeology. Springer, New York, pp. 41–68.
Rinaldi, R., Bandini Mazzanti, M., Bosi, G., 2013. Archaeobotany in the urban sites: the case of Mutina. Ann. Bot. 3, 217–230.
Rizzi, E., Lari, M., Gigli, E., De Bellis, G., Caramelli, D., 2012. Ancient DNA studies: new perspectives on old samples. Genet. Sel. Evol. 44, 21.
Rogaev, E.I., Moliaka, Y.K., Malyarchuk, B.A., Kondrashov, F.A., Derenko, M.V., Chumakov, I., Grigorenko, A.P., 2006. Complete mitochondrial genome and phylogeny of pleistocene mammoth *Mammuthus primigenius*. PLoS Biol. 4, e73.
Rollo, F., Ubaldi, M., Ermini, L., Marota, I., 2002. Ötzi's last meals: DNA analysis of the intestinal content of the Neolithic glacier mummy from the Alps. Proc. Natl. Acad. Sci. 99, 12594–12599.
Sagan, L., 1967. On the origin of mitosing cells. J. Theor. Biol. 14, 225–274.
Samaniego Castruita, J.A., Zepeda Mendoza, M.L., Barnett, R., Wales, N., Gilbert, M.T.P., 2015. Odintifier - a computational method for identifying insertions of organellar origin from modern and ancient high-throughput sequencing data based on haplotype phasing. BMC Bioinforma. 16, 232.
Schelkunov, M.I., Shtratnikova, V.Y., Nuraliev, M.S., Selosse, M.-A., Penin, A.A., Logacheva, M.D., 2015. Exploring the limits for reduction of plastid genomes: a case study of the mycoheterotrophic orchids *Epipogium aphyllum* and *Epipogium roseum*. Genome Biol. Evol. 7, 1179–1191.
Schubert, M., Ginolhac, A., Lindgreen, S., Thompson, J., AL-Rasheid, K., Willerslev, E., Krogh, A., Orlando, L., 2012. Improving ancient DNA read mapping against modern reference genomes. BMC Genom. 13, 178.
Sefc, M.K., Steinkellner, H., Glössl, J., Kampfer, S., Regner, F., 1998. Reconstruction of a grapevine pedigree by microsatellite analysis. Theor. Appl. Genet. 97, 227–231.
Shapiro, B., Hofreiter, M., 2014. A paleogenomic perspective on evolution and gene function: new insights from ancient DNA. Science 343.
Sherman-Broyles, S., Bombarely, A., Grimwood, J., Schmutz, J., Doyle, J., 2014. Complete plastome sequences from *Glycine syndetika* and six additional perennial wild relatives of soybean. G3 Genes, Genomes Genet. 4, 2023–2033.
Smith, A., Bagoyan, T., Gabrielyan, I., Pinhasi, R., Gasparian, B., 2014. Late chalcolithic and medieval archaeobotanical remains from Areni-1 (birds' cave), Armenia. In: Gasparian, B., Arimura, M. (Eds.), Stone Age of Armenia: a Guide-book to the Stone Age Archaeology in the Republic of Armenia. Center for Cultural Resource Studies, Kanazawa University, Kanazawa, Japan, pp. 233–260.
Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., Willerslev, E., 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Res. 35, e14.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
Terral, J.-F., Tabard, E., Bouby, L., Ivorra, S., Pastor, T., Figueiral, I., Picq, S., Chevance, J.-B., Jung, C., Fabre, L., Tardy, C., Compan, M., Bacilieri, R., Lacombe, T., This, P., 2010. Evolution and history of grapevine (*Vitis vinifera*) under domestication: new morphometric perspectives to understand seed domestication syndrome and reveal origins of ancient European cultivars. Ann. Bot. 105, 443–455.
This, P., Lacombe, T., Thomas, M.R., 2006. Historical origins and genetic diversity of wine grapes. Trends Genet. 22, 511–519.
Tröndle, D., Schröder, S., Kassemeyer, H.-H., Kiefer, C., Koch, M.A., Nick, P., 2010. Molecular phylogeny of the genus *Vitis* (Vitaceae) based on plastid markers. Am. J. Bot. 97, 1168–1178.
Ucchesu, M., Orrù, M., Grillo, O., Venora, G., Usai, A., Serreli, P.F., Bacchetta, G., 2015. Earliest evidence of a primitive cultivar of *Vitis vinifera* L. during the bronze age in Sardinia (Italy). Veg. Hist. Archaeobot. 24, 587–600.
Vouillamoz, J.F., Grando, M.S., 2006. Genealogy of wine grape cultivars: 'Pinot' is related to 'Syrah'. Heredity 97, 102–110.
Wales, N., Romero-Navarro, J.A., Cappellini, E., Gilbert, M.T.P., 2012. Choosing the best plant for the job: a cost-effective assay to prescreen ancient plant remains destined for shotgun sequencing. PLoS ONE 7, e45644.
Wales, N., Allaby, R.G., Willerslev, E., Gilbert, M.T.P., 2013. Ancient plant DNA. In: Elias, S.A. (Ed.), Encyclopedia of Quaternary Science, vol. 2. Elsevier, Oxford, pp. 705–715.
Wales, N., Andersen, K., Cappellini, E., Ávila-Arcos, M.C., Gilbert, M.T.P., 2014. Optimization of DNA recovery and amplification from non-carbonized archaeobotanical remains. PLoS ONE 9, e86827.
Whittall, J.B., Syring, J., Parks, M., Buenrostro, J., Dick, C., Liston, A., Cronn, R., 2010. Finding a (pine) needle in a haystack: chloroplast genome sequence divergence in rare and widespread pines. Mol. Ecol. 19, 100–114.
Wilkinson, K.N., Gasparian, B., Pinhasi, R., Avetisyan, P., Hovsepyan, R., Zardaryan, D., Areshian, G.E., Bar-Oz, G., Smith, A., 2012. Areni-1 cave, Armenia: a chalcolithic–early bronze age settlement and ritual site in the southern Caucasus. J. Field Archaeol. 37, 20–33.
Willerslev, E., Hansen, A.J., Binladen, J., Brand, T.B., Gilbert, M.T.P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D.A., Cooper, A., 2003. Diverse plant and animal genetic records from holocene and pleistocene sediments. Science 300, 791–795.
Willerslev, E., Gilbert, M.T.P., Binladen, J., Ho, S.Y., Campos, P.F., Ratan, A., Tomsho, L.P., da Fonseca, R.R., Sher, A., Kuznetsova, T.V., Nowak-Kemp, M., Roth, T.L., Miller, W., Schuster, S.C., 2009. Analysis of complete mitochondrial genomes from extinct and extant rhinoceroses reveals lack of phylogenetic resolution. BMC Evol. Biol. 9, 1–11.
Wolfe, A.D., Randle, C.P., 2004. Recombination, heteroplasmy, haplotype polymorphism, and paralogy in plastid genes: implications for plant molecular systematics. Syst. Bot. 29, 1011–1020.
Yang, J.-B., Tang, M., Li, H.-T., Zhang, Z.-R., Li, D.-Z., 2013. Complete chloroplast genome of the genus *Cymbidium*: lights into the species identification, phylogenetic implications and population genetic analyses. BMC Evol. Biol. 13, 1–12.
Zohary, D., Hopf, M., Weiss, E., 2012. Domestication of Plants in the Old World: the Origin and Spread of Domesticated Plants in South-west Asia, Europe, and the Mediterranean Basin. Oxford University Press, Oxford.
Zohary, D., 1996. The domestication of the grapevine *Vitis vinifera* L. in the Near East. In: McGovern, P.E., Fleming, S.J., Katz, S.H. (Eds.), The Origins and Ancient History of Wine. Gordon & Breach, Amsterdam, pp. 23–30.