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
Available online 14 June 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