Title
Ancient Plant Genomics in Archaeology, Herbaria, and the Environment

Authors
Logan Kistler[1]*, Vanessa C. Bieker[2], Michael D. Martin[2], Mikkel Winther Pedersen[3], Jazmín Ramos Madrigal[4], Nathan Wales[5]

Author Affiliations
1 – Smithsonian Institution, National Museum of Natural History, Department of Anthropology. Washington, DC 20650.
KistlerL@si.edu

2 – The Norwegian University of Science and Technology University Museum, Department of Natural History. 7491 Trondheim, Norway.
vanessa.bieker@ntnu.no; mike.martin@ntnu.no

3 – Lundbeck Foundation GeoGenetics Centre, GLOBE Institute, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen K, Denmark
mwpedersen@snm.ku.dk

4 – Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen, Denmark.
jazmin.ramos@snm.ku.dk

5 – Department of Archaeology, University of York, King's Manor, York, YO1 7EP, UK.
nathan.wales@york.ac.uk

*Correspondence to: KistlerL@si.edu; +1 202 633 1908

Abstract

The ancient DNA revolution of the past 35 years has driven an explosion in breadth, nuance, and diversity of questions approachable using ancient biomolecules, and plant research has been a constant, indispensable facet of these developments. Using archaeological, paleontological, and herbarium plant tissues, researchers have probed plant domestication and dispersal, plant evolution and ecology, paleoenvironmental composition and dynamics, and other topics across related disciplines. Here, we review the development of the ancient DNA discipline and the role of plant research in its progress and refinement. We summarize our understanding of long-term plant DNA preservation and the characteristics of degraded DNA. In addition, we discuss challenges in ancient DNA recovery and analysis, and the laboratory and bioinformatic strategies used to mitigate them. Finally, we review recent research applications of ancient plant genomics.

Keywords
Archaeogenomics; Paleogenomics; Ancient Plant DNA; Environmental Ancient DNA; Archaeobotany; Herbarium collections

1. Introduction

The ancient DNA revolution

The analysis of DNA from long-deceased tissues was first reported in 1984 with the publication of 213bp of mitochondrial DNA from a museum-preserved quagga, a recently-extinct zebra relative (57). This foundational study demonstrated that DNA from historic tissues had significant potential for long-term survival and recovery through the tools of modern molecular genetics. The PCR revolution in the late 1980s led to rapid expansion of the burgeoning ancient DNA (aDNA) field, when researchers simultaneously began to probe the limits of DNA survival and understand the nuance and difficulty involved in working with archaeological and paleontological biomolecules (95). The primary challenges that emerged were extreme sensitivity to contamination by modern DNA sources, and the degraded nature of ancient DNA molecules.

Early reports of DNA from Mesozoic bone (163), Miocene plant remains (46), and remains in amber (reviewed by (143)) are among the formative results later learned to have originated with contamination (131). Collectively, the growing pains of the new methodology guided the ancient DNA field toward a set of rigorous experimental criteria for ensuring authenticity of ancient DNA (26, 44). These criteria included physical isolation of ancient DNA labwork from other molecular facilities and stringent, specialized laboratory procedures (e.g. (42)), independent replication of experimental results, and predictions about the preservation and behavior of ancient biomolecules (26). The revision and formalization of ancient DNA procedures ultimately shaped the field for the ensuing two decades of rigorous research.

Ancient DNA methodologies matured during the PCR era. Early after the emergence of genomic technologies, however, the aDNA field began a protracted shift to massively parallel high-throughput sequencing (HTS) that reflected the broader trend in molecular genomics. Early genomic applications to ancient specimens involved highly multiplexed PCR and amplicon sequencing (73), low-coverage mammoth genomic screening and divergence estimation from

elephants (108), and cytoplasmic genome assembly from ancient organisms (73). The first complete genome of an ancient organism—a ~20,000 year-old woolly mammoth from Siberia—was published in 2008 using the now-defunct 454 sequencing platform (86). In the subsequent ~decade, aDNA methods have been refined and optimized for the genomic era, genomic data collection costs have dropped by orders of magnitude, and the early promise that ancient specimens could be completely integrated into nuanced analytical frameworks using large datasets is being realized (109, 133).

The role of plants through the aDNA era.

Plant research has been integral to each phase of the ancient DNA revolution. Closely following the quagga proof-of-concept for aDNA survival, researchers began assessing the biomolecular preservation potential of herbarium specimens (115) and amplifying short fragments of maize DNA (120), eventually gaining early insights about the process of domestication (47). PCR was deployed extensively throughout the 1990s with archaeological and paleoecological plant remains including charred and desiccated seeds, maize cobs, fossil plants, archaeological textiles, herbarium sheets, and sediment cores (96). As with other study systems, the oldest of these materials—pre-quaternary fossil remains (46)—are now widely considered early examples of contaminated experiments, and helped shape the rigorous procedures and the maturation of the field. Plant research has kept pace with the paleogenomic revolution, beginning with genome-wide analysis of cotton evolution in 2012 (97), and continuing with a wide range of species across diverse study areas to the present, discussed in the following.

Additionally, plant research was at the forefront of ancient biomolecular developments in several cases. The first aDNA target sequence capture was carried out with maize (10); the first ancient RNA reported by hybridization in cress (119) and by sequencing in maize (37); the first ancient RNA viruses were recovered from barley (136); and the earliest ancient epigenomic sequencing was carried out on well-preserved barley specimens (137) (Figure 1). Furthermore, the domestication of plants was an early research framework for tests for selection using aDNA, providing a real-time evolutionary account of specific human-influenced traits in crop species (59). Lastly, plant research led the way for applying population genetic methods to ancient DNA datasets, again motivated by questions of domestication and crop movements ((39, 59); and see below).

In total, the history of ancient plant DNA closely reflects broader advances in the aDNA field, with several key benefits to working with ancient specimens compared with only modern material. These advantages include integrating extinct variation to contextualize modern biodiversity and understand species dynamics through time, resolving taxonomic and compositional uncertainty in archaeobotanical and paleoecological contexts, and tracking in real-time processes such as crop domestication and Quaternary human impacts on diverse ecosystems. While fundamental limits to ancient DNA survival and obstacles to its recovery still exist, plant aDNA research in the genomic era is increasingly subtle and powerful owing to the ability to build complete ancient genomes into nuanced analytical frameworks.

2. DNA survival in plant tissues and the environment

Ancient DNA authenticity, and expectations of DNA breakdown.

The two most ubiquitous challenges to authentically ancient DNA analysis are 1) the rapid degradation of DNA in deceased tissues, and 2) the contamination of experiments and sample materials by ubiquitous sources of high-quality modern DNA. The latter is preempted by strict adherence to stringent protocols for aDNA analysis (26, 42, 44), but the former imposes fundamental limits on aDNA recoverability.

Isolated DNA in solution is known to spontaneously break down at a predictable rate governed by temperature and the chemical environment, so that the number of DNA molecules at a given size decreases through exponential decay with a measurable half-life (78). While this prediction has been previously applied to ancient DNA breakdown (6), a recent meta-analysis suggested instead that in animal hard tissues, most DNA fragmentation occurs rapidly after organism death (69), a result predicted by previous experimental work (22). In most cases, DNA is lost entirely, but in the rare conditions favorable to DNA preservation, fragmentation then appears to slow dramatically to some level of stability and then decay more slowly over long timeframes. Instead of sample age, this study found that DNA fragmentation was best predicted by thermal fluctuations and precipitation in the regional environment (69). Even within single archaeological sites, relative age was a poor predictor of which samples would be most fragmented. The oldest DNA widely accepted as authentic was recovered from a 500-700kya horse recovered in the North American permafrost under exceptional preservational conditions, underscoring the environmental factors involved with DNA survival (94).

In contrast to what is known about skeletal remains, sufficient ancient genomes for a similar study in plants have not yet been generated with sufficient time depth and environmental variability for this kind of analysis, and the highly variable composition of plant tissues compared with animal hard tissues warrant independent investigation for their preservational dynamics. Studies using herbarium leaf tissue have observed a more constant rate of DNA breakdown than that in bone (157), but have focused on a maximum timespan of ~300 years rather than the half million years possible with animal genomes. Moreover, the herbarium pattern of DNA breakdown is restricted to a single preservational method—desiccation and protected storage, sometimes with chemical treatments or heat. In total, much more work is necessary to establish empirical expectations of ancient plant DNA survival.

In addition to physical fragmentation, DNA accumulates chemical damage and compositional biases which can be used as markers of authentically ancient fragments. Most prominent among these, cytosine nucleotides tend to spontaneously lose an amine group and convert to deoxyuracil (18). During library preparation, this uracil residue is complemented by an adenine template on the opposite strand, and recovered during sequencing as a thymine on the original strand. The effect of these misincorporations when mapping reads to a reference sequence is a surplus of reference C to read T mismatches in ancient genomes. Further, this process occurs much more rapidly in the single-stranded overhangs occurring at the ends of fragments than in the double-stranded body of the molecule. As such, an abundance of C-to-T mismatches in the 5’ end of reads and complementary G-to-A mismatches in the 3’ end are treated as one gold standard for ancient DNA authenticity in the genomic era (69, 117). This pattern can also be used to discriminate endogenous from contaminant molecules, particularly in cases like human and hominin research where contamination is ubiquitous and difficult to isolate (134).

To summarize, we expect that DNA from historic and ancient tissues will be highly fragmented, chemically damaged, substantially depleted of endogenous DNA, and mixed with environmental DNA. Consistent with long-established criteria for ancient DNA authenticity (26, 44), these characteristics are useful for validating the results of ancient DNA studies and for discriminating authentic from contaminant DNA sources. The nature of DNA contamination and decay also has significant implications for the laboratory processing and informatic components of aDNA genomics, discussed below (Table 1).

Archaeobotanical tissues, and other sources of plant aDNA

Archaeological macrofossils

Archaeobotanical macrofossils—organic tissues including seeds, fruits, cobs, shells, and wood—are preserved by four key mechanisms: 1) charring, often in storage pits or hearth features; 2) desiccation in dry caves, rockshelters, or arid environments; 3) anaerobic waterlogging; and 4) partial or complete mineralization. By far the majority of archaeobotanical remains worldwide are charred, often with complete carbonization of organic tissue but good structural preservation allowing morphological examination. Charred remains are routinely recovered from archaeological sediments and features through flotation, and a range of analytical approaches can be applied to reconstruct aspects of past plant communities and human use (101).

Early attempts to amplify DNA fragments from charred cereal grains were successful in a low proportion of seeds across several experiments (2, 3, 19). Preserved pockets of organic matter were thought to remain in some incompletely carbonized tissues, facilitating recovery of trace levels of DNA (20, 96). Even in the genomic era, PCR remains the most sensitive method for recovery of a single target present at very low copy numbers, and early work with charred remains seems to have leveraged this sensitivity where minute levels of DNA persisted in a small fraction of burned seeds. However, attempts at archaeogenomic analysis with charred remains have been uniformly disappointing. A recent re-analysis of charred plant archaeogenomic attempts representing four species using and both metagenomic and target capture approaches revealed a total of only 26 authentically ancient reads across more than 200 million total fragments sequenced (91). Further, this study re-visited raw data from a previous genomic test case using charred cereal grains (21), and found that the previous results were more likely based on exogenous molecules homologous to the target reference genomes rather than successful recovery of endogenous DNA. In total, charred plant remains do not appear to be good candidates for ancient DNA research for genomics unless methodological strides substantially improve recovery success.

Desiccated and waterlogged archaeobotanical macrofossils, however, have proven to be excellent reservoirs for DNA survival over at least ten thousand years and in diverse environments. Dry caves and rockshelters (68, 82, 145, 152), arid Andean (65, 161) and Egyptian (139) sites, and the desert southwestern U.S. (29) have all shown excellent DNA preservation in plant remains, and all complete genome sequences from archaeological plants to date have been from dry-preserved remains. Waterlogged remains are less frequent, since they rely on serendipitous deposition in bodies of water where decomposition is impeded. However, waterlogged grape pips (23, 111), gourd rinds (66), squash seeds (67), and oak wood (151) have yielded high-quality datasets using target sequence capture or focusing on

cytoplasmic DNA analysis with genomic methods. Finally, no recent studies have tested mineralized remains for DNA preservation, and chances of success with these materials seem limited owing to the replacement of organic tissue with inorganic substrate.

Archaeological microfossils

Pollen, starch grains, and phytoliths constitute valuable resources for environmental and archaeological reconstruction, particularly in regions such as the humid tropics where conditions prevent macrofossil preservation (101, 150). Pollen contains genomic material for reproduction and, as with charred remains, there were some early PCR-based successes extracting DNA directly from ancient pollen grains (100). In the genomic era, this approach has largely been subsumed by sedimentary ancient DNA (sedaDNA) strategies, discussed below. Starch grains—tightly-bound amylose residues cross-linked with amylopection—are not expected to carry any nucleic acids, and have not been targeted as ancient DNA sources.

On the other hand, phytoliths—non-crystalline silica bodies in diverse plant taxa and tissues—often contain organic material sufficient for direct radiocarbon dating, suggesting significant incidental sequestration of cellular debris during phytolith formation (104). Although the only published attempt at phytolith DNA recovery was unsuccessful (33), it remains plausible that phytoliths could be a viable source of nucleic acids. Moreover, researchers exploring long-term data storage using synthesized DNA found that storing information-bearing DNA molecules in amorphous silica beads was extremely effective at combatting degradation, substantially outperforming other room-temperature DNA storage strategies including desiccation and chemical preservation (49). As such, if DNA is present in phytoliths during a plant's life cycle, it could possibly survive archaeological timespans in a chemical microenvironment optimal for DNA preservation.

Plant DNA in ancient sediments

The early 2000s saw the first genetic characterization of ancient plant communities preserved in permanently frozen soil and cave sediments (160). Using generic plant DNA markers (metabarcodes), Willerslev and colleagues (158) later demonstrated that diverse plant species could be retrieved from small amounts of bulk sediment samples over very long timespans (>400 kyr BP). This method does not rely on the preservation of identifiable plant fossils, which is advantageous where a large proportion of a plant's structures and many plant species do not preserve and therefore go undetected by traditional paleo-ecological methods. However, many plants leave genetic traces in the environment which can be preserved in geological and archaeological deposits. Several paleogenetic studies have since utilized bulk sedaDNA and plant metabarcodes to investigate past floristic assemblages from marine, lacustrine, terrestrial, and archaeological sediments as well as basal glacial ice (9, 25, 62, 88, 102, 158, 159).

More recently, shotgun metagenomic sequencing of lake and marine sediments has provided a new revolutionizing way to retrieve and analyze partial/whole plant genomes (98, 103). This approach was used for investigating past floras from Holocene archaeological middens (129) and in Pleistocene cave sediments from hominin sites in Europe (135). Recent bulk extraction and shotgun metagenomic sequencing from stalagmites yielded ~80 kyr BP old plant DNA (142), and sedaDNA-based research in marine deposits provides a faithful correlate to pollen and macrofossil remains (102, 138). The shotgun metagenomic approach allows

simultaneous analysis of all organismal DNA for a more comprehensive understanding of the past environment, and improves verification and quantification of DNA degradation (e.g. cytosine deamination) as a measure of DNA authenticity (103).

Despite recent methodological advances and expanded possibilities in both metabarcoding and sedaDNA, we still have limited knowledge about the taphonomy of environmental DNA, such as the pathway of deposition and the conditions which encourage long-term preservation in the different depositional environments (99). From the current literature, we can observe that only a subset of plants and other organisms in the environment leave detectable genetic traces in the associated sediment. While this observation is likely a combination of a multitude of factors, key taphonomic processes may include distance from source to deposit and individual biomass production of living organisms (7). Another important limitation is the lack of complete reference databases and full reference genomes available, while this is constantly expanding only a small fraction today of all plants have their full genome sequenced. Bias-correcting approaches such as phylogenetic intersection analysis can be invoked to mitigate effects of database gaps, such as false positive hits to over-represented species (138). But as the number of species in databases grow with lower costs for genome sequencing, this bias will decrease to improve taxonomic resolution of metagenomic analysis and usher in possibilities beyond the current presence-absence taxonomic identification.

Coprolites

Archaeological and paleontological remains of ancient feces—coprolites—provide an exceptional opportunity to investigate past diets and health (50). For decades archaeologists have dissected coprolites to reveal their macroscopic and microscopic contents, such as plant fibers, seeds, pollen, and intestinal parasites (114). With the advent of PCR, researchers began investigating paleofeces with genetic markers for intestinal bacteria (40), parasitic worms (58, 79), gut microbiomes (125), and in the case of ancient human feces, the chromosomal sex of the defecator (144). In addition to these lines of inquiry, several groups inferred the plant component of coprolites though PCR-amplification of plastome barcodes. For example, Poinar et al. (107) tested a 20 kya fecal bolus of an extinct giant ground sloth and found genetic evidence of eight orders of plants, only four of which were observed through macroscopic analysis. The researchers argue this molecular detection is advantageous as it captured a greater range of the ground sloth's diet, although one additional plant taxon was only identified through macroscopic examination, suggesting a combined approach may be the most fruitful. Other researchers have used plastome markers to examine the plant component of human diets, including the colon contents of Ötzi the Iceman (121) and a pre-Clovis coprolite from Paisley Caves, Oregon (113).

Advances in DNA sequencing technologies allow researchers to treat coprolites as metagenomic samples, much like is now done with ancient sediment samples. Rather than amplifying genetic markers for plants, shotgun sequencing can be used to take a genetic snapshot of all taxa present in a coprolite, including gastrointestinal pathogens, beneficial species of the gut microbiome, and dietary plants and animals. So far, this direction has been limited, and in fact recent studies of plant DNA from coprolites continue to follow a metabarcoding approach, where PCR amplicons are sequenced on an Illumina platform (e.g. New Zealand birds (17); New Zealand dogs (162)). However, this may change in the future due

to the recognition that PCR-amplification of genes yields a biased picture of ancient microbiomes due to gene length polymorphisms and variations in primer binding sites (167). Even though coprolites are relatively uncommon in the archaeological and paleontological records, often restricted to dry cave contexts for long-term preservation, we anticipate increased interest in shotgun metagenomics of paleofeces, as this approach should provide the most accurate characterization of diet and gut microbiomes.

Herbaria

Global herbarium collections contain approximately 350 million specimens amassed during the last ca. 400 years and are increasingly viewed as valuable repositories for genome-scale biodiversity data (14). As herbaria digitize their collections at a quickening pace (12), photos and metadata associated with each specimen become available through online databases such as the Global Biodiversity Information Facility (GBIF) (146). The collections thus become more accessible and the number and diversity of potential scientific applications expands, especially with advances in DNA extraction and HTS (48, 110). Genomic sequencing approaches successfully applied to historical herbarium specimens include shotgun deep-sequencing (81, 165), genome skimming (11, 87, 92), targeted DNA capture (53, 56, 124), SNP assays (149), and *de novo* organellar genome assembly (11, 123). In addition to their well established and common usage providing material for studies of molecular systematics (93), herbarium collections, often containing large numbers of well-preserved, expertly curated conspecific specimens that can be grouped in time and space according to their associated metadata, and are therefore particularly well suited for population genomic studies of temporal evolution and ecological change (15).

Herbarium specimens can be used if samples cannot be collected in the wild, for example if species are endangered or extinct, or if collection from the wild is not feasible due to high cost and time restrictions. For example, Konrade et al. (70) tested about 500 herbarium specimens of black cherry (*Prunus serotina*) to sample the whole Eastern North American range of this widespread species. They were able to detect a weak signal of isolation-by-distance by using 15 microsatellite loci. This pattern would be difficult or even impossible to detect without having samples from the whole range. In another study, Olofsson et al. (93) analyzed 28 herbarium specimens collected between 1872 and 2013 in their phylogenetic study of the Oleeae tribe (Oleaceae, Lamiales), some of which originated from remote areas in tropical Asia and Australasia, and neotropical America. They showed that it is possible to infer the phylogeny even with low-depth genome skimming data when using high copy regions such as the plastome or nuclear ribosomal clusters.

Herbarium specimens are also increasingly used for generating DNA barcodes, either to improve barcode reference databases or to verify the identification of a specimen. It is notable that the barcode success is highly dependent on the sample age. For example, Korpelainen and Pietiläinen (72) used standard barcode protocols to obtain plastid barcodes of Finnish angiosperm species from over 3000 herbarium specimens collected between 1867 and 2013 as part of the Finnish Barcode of Life initiative (FinBOL; www.finbol.org). The success rate of barcoding was only 35% in specimens collected between 1867 and 1899, compared to about 90% in those collected between 2001 and 2013, underscoring the increasing difficulty of long fragment recovery in aging specimens. Target sequence capture using reliable orthologous

panels of targets across plants has potential as a powerful method for herbarium systematics and data collection similar to barcoding strategies (e.g. (60)).

3. *Challenges and characteristics of ancient plant genomics*

Plant-specific considerations of aDNA research

Ancient plant genomics combines the biological complexities of plant genomes with the practical challenges of ancient genomics. The idiosyncrasies of plant genomes include massive genome size variation; domestic grasses range from ~400Mbp (rice) to 16Gbp (bread wheat) haploid genome size, and flowering plants span at least 68Mbp to 148Gbp, a >2000-fold range of variation (75). Compared with a ~4-fold variation in mammalian genomes, for example (63), this extensive range presents data collection challenges in many species, especially when DNA is highly fragmented and tissues are contaminated by environmental DNA. Thus target sequence capture (45) presents a useful method to reduce data collection requirements in many cases that may be intractable using shotgun sequencing. Ploidy variation complicates plant research broadly (76), and alloploidy through hybridization presents additional challenges in short fragment mapping and target capture experimental design when paralogous targets are rampant. Moreover, hemizygosity and genomic structural variation have been shown to evolve substantially even over the Holocene timespans of domestication (16, 166), introducing fundamental variation in genomic composition among individuals. The variable breeding systems of plants—inbreeding and outcrossing, plus vegetative propagation and diverse ecological breeding contexts—render many population genetic assumptions unreliable, so that plant population analyses must rely on modified and purpose-built approaches. Finally, the three cellular sources of DNA in plants—plastids, mitochondria, and the nucleus—have the potential for substantial horizontal transfer (66, 155), posing risks of variable copy number, non-specific read mapping, and confounded variant calling. Plant aDNA research, in total, must consider these and other fundamental difficulties of plant biology in the context of the specific research design and analytical requirements necessary for ancient DNA broadly (Table 1).

Analytical and computational challenges

Laboratory methodologies

The fragmented, low-quantity, and fragile nature of ancient DNA requires strict protocols to mitigate contamination from modern sources and recover miniscule quantities of degraded DNA from diverse tissues. Isolating DNA from varied samples including herbarium vouchers, ancient seeds, coprolites, and sediment cores requires a wide range of off-the-shelf and purpose-built methods. These techniques are often based on protocols for fresh samples, with modifications to maximize the recovery of degraded DNA. Even as protocols continue to be optimized, here we highlight the most common approaches and issues in the laboratory component of plant archaeogenomics.

Of the substrates used in ancient plant genomic research, herbarium samples are peculiar in their often-pristine visual appearance. Thus, one might assume herbarium samples have biomolecular and biochemical traits much like living plants, and therefore could be treated as such in the laboratory. Researchers have frequently extracted DNA from herbarium specimens using techniques developed for fresh plant tissues. For example, cetyl-trimethyl ammonium bromide (CTAB) is one of the most popular methods for isolating DNA from living

plants (32), and many researchers have applied it to herbarium samples (8, 27, 28). Commercial kits have also been used to recover DNA from herbarium samples, such as the use of a Plant Mini Kit (Qiagen) on *Juniperus* accessions up to 80 years old (1) and successful testing of the 100-year-old *Sartidia perrieri* type specimen with a BioSpring 15 DNA Plant Kit (Qiagen) (13).

Despite some successes with conventional DNA extraction methodologies, DNA contained within herbarium specimens is far from pristine. In one study, newly prepared herbarium vouchers were found to contain only 3% the amount of DNA as fresh tissue, leading the authors to argue the vast majority of DNA is lost during herbarium specimen preparation (141). The amount of DNA destroyed through conservation treatments is variable, and some methods like air drying are more forgiving than alcohol drying (126). However, it is difficult to infer how much DNA is preserved in individual vouchers because treatment methods were rarely recorded during the conservation process. In addition to the net loss of DNA molecules, DNA in herbarium samples is much more fragmented than DNA in fresh tissues. In a set of experiments on herbarium specimens collected over the past 170 years, Weiß et al. (157) observed median DNA fragment lengths of 50–90bp, similar to that of archaeological bones hundreds to thousands of years old. The DNA in herbarium samples is also chemically damaged by cytosine deamination (51), similar to other ancient DNA substrates (see Section 2). However, cytosine deamination is a time-dependent process (69, 127), and therefore most herbarium samples exhibit very low levels of chemical damage.

Herbarium collections are a non-renewable resource, so researchers must minimize destructive sampling and maximize DNA recovery. Today, herbarium specimens are increasingly used for genome-wide characterization as opposed to PCR-based assays (15), and consequently it is critical to retain the short DNA fragments which are ideal for short read sequencing technologies. Gutaker et al. (52) performed one of the most recent and relevant investigations, using *Arabidopsis thaliana* accessions collected 1839–1898. After comparing multiple methods including CTAB, they ultimately recommend a digestion buffer containing N-phenacylthiazolium bromide (PTB) and dithiothreitol (DTT) used previously for ancient gourd rinds (66), followed by a specialized silica binding purification originally developed for ancient skeletal remains (30).

Regardless of the DNA extraction method, it is critical to recognize that herbarium specimens, though relatively young, require special handling and processing beyond what is appropriate for fresh plant tissue. Samples should be collected with sterile implements and care must be taken to avoid cross contamination in the herbarium and in the laboratory. Disposable gloves are essential, as they minimize DNA transfer between samples and also reduce skin contact with mercuric chloride, a toxin once used routinely as a conservation treatment (156). The impact of mercury salts on DNA preservation is an interesting but so far underexplored issue (35). Due to these challenges, Shepherd and Perrie (132) have argued that herbarium accessions should be processed in physically isolated laboratories with extensive controls to monitor contamination, following the principles of aDNA research (26). As many herbarium samples yield low amounts of endogenous DNA, extraction in dedicated paleogenomic laboratories is ideal (42). At a minimum, herbarium vouchers must not be processed in facilities where amplified DNA is present.

Researchers attempting to recover nucleic acids from archaeological macrofossils face similar challenges as with herbarium vouchers, and therefore implement some of the same laboratory strategies. In contrast to herbarium specimens, archaeobotanical samples are preserved by natural processes discussed above, most commonly charring, waterlogging, desiccation, and mineralization (168). The most successful archaeogenomic studies have been focused on desiccated (65, 82, 112, 139, 145, 147) or waterlogged (111) archaeological macrofossils. When possible, it is preferable to take samples directly at excavations as lengthy exposure to new environments could lead to microbial colonization or accelerated DNA degradation. However, most of the above projects tested desiccated macrofossils which had been stored in museums for decades.

As with herbarium samples, laboratory methods must be optimized to recover short DNA fragments while minimizing contamination. DNA extraction and all pre-amplification steps must be performed in a dedicated aDNA laboratory, as this is the best practice to avoid and monitor sources of contamination. Various DNA extraction methods have been successfully used by different research groups, but in comparing extraction protocols on archaeobotanical remains, Wales *et al.* (153) reported optimal success with a digestion buffer containing sodium dodecyl sulfate (SDS), proteinase K, and DTT, followed by a phenol-chloroform extraction. Wales and Kistler (154) provide an updated version of the method which implements an optimized silica purification for ultrashort DNA.

A key consideration is the great range of macrofossil tissue types—from maize cobs, cereal seeds, gourd rinds, sunflower heads, twigs, and hardwood. In contrast to herbarium samples where leaf tissue is normally available for testing and metabolite composition is often similar across taxa, the diversity of plant macrofossils requires researchers to investigate if tissue-specific compounds may interfere with DNA purification or library preparation. In addition to the plant-derived compounds, humic acids from the archaeological sediment may adhere to small macrofossil remains. Humic compounds often co-extract with DNA due to their similar chemical properties, and they create downstream problems through enzymatic inhibition (83). Based on these issues, it is unlikely one DNA extraction method will consistently outperform the others across all archaeological macrofossils, and small pilot studies are often advisable to determine the best option for a given set of samples.

Extraction of plant and other DNA from coprolites and archaeological soil follows similar approaches as plant macrofossils. For coprolites, Poinar et al. (107) demonstrated plant DNA could be recovered from 20,000-year-old dung of the extinct Shasta ground sloth using a PTB digestion buffer. PTB cleaves cross-links between DNA and macromolecules, and it is thought this may release trapped DNA from ancient feces. However, experiments by Rohland and Hofreiter (118) indicate PTB can have a negative effect on DNA recovery, and other researchers have isolated plant DNA from ancient human and other coprolites without PTB (113, 162). Given recent success recovering ultrashort DNA from herbarium vouchers with PTB (52), there is a need to reinvestigate the efficacy of the additive. For sedimentary DNA studies, a major concern is the coextraction of humic acids, and a range of extraction methods have been explored (98). In some cases, modification of commercial kits for soil DNA have been useful, such as Pedersen et al.'s (103) implementation of inhibitor-removing buffers from the PowerSoil Kit (MoBio/Qiagen). Rohland et al. (116) provide the most recent optimization,

showing a simple lysis buffer (Tween-20 and proteinase K) paired with silica binding can recover ultrashort DNA from sediment.

After DNA is isolated from herbarium vouchers, archaeological macrofossils, coprolites, or sediment, the subsequent laboratory steps are identical. For high-throughput sequencing, raw DNA must be converted to DNA libraries taking special care to retain short fragments and maximize complexity in light of low input concentrations. One of the most common library preparation methods relies on blunt-end ligation of custom adapters (85), which is more efficient and less biasing for ancient DNA than the AT-overhang adapter ligation methods common in commercial kits (130). Modified single-tube protocols have built on this fundamental strategy and optimized library complexity and DNA yield from ancient sources (24). A more sophisticated "single-stranded" library preparation protocol starts by denaturing double-stranded DNA molecules so that each strand can be individually sequenced (43), offering advantages with highly degraded template DNA. Optionally, deaminated cytosine sites can be enzymatically repaired using uracil-DNA glycosylase before library preparation, although the presence of deaminated sites is a useful indicator of DNA authenticity (117). Following library preparation, samples can be treated similar to highly fragmented modern DNA for shotgun-sequenced or target enrichment using hybridization capture (45).

Bioinformatic strategies

Paleogenomics has helped dissect evolutionary processes such as population bottlenecks (152), recent local adaptation (29, 139, 145), ancient admixture (82) and dispersal (29, 65) in various plant species. Notably, genetic studies based on archaeobotanical remains—as well as any other paleogenomic study—are complicated by post-mortem DNA degradation (18). The degradation and chemical damage in ancient DNA has the effect that paleogenomic datasets consist of very short sequencing reads, increased apparent error rates, and low genomic coverage. These features have the potential to affect bioinformatic analyses from initial data processing to end-stage evolutionary inferences.

In most studies, sequencing reads from an ancient plant sample are mapped to a reference genome from the same or a closely related species using modifications to standard algorithms (77, 128). The accuracy of this process decreases for shorter reads with high error rates, for which mapping locations become ambiguous in some cases. This issue is of particular relevance for plant genomes, which can be heterogeneous in size, even within the same species (16), as well as highly repetitive, which hinders retrieval of information from short reads. For instance, due to its high repeat content and recent whole-genome duplication, only 21% of the maize genome can be unambiguously mapped with 30nt reads (112). Moreover, larger genomes call for larger sequencing data volumes in order to reach useful genome coverage, thus increasing sequencing costs for samples with low endogenous content.

In genomic studies, once sequencing data have been mapped, information at specific genomic coordinates is transformed into called genotypes (90). However, this is usually not possible for paleogenomic studies, where low depth of coverage is a common feature of the data. Furthermore, higher error rates derived from post-mortem damage incorporate artificial variation that cannot be otherwise excluded through high-depth genotyping. Strategies to circumvent this issue include random sampling of one read at sites of interest (e.g. (65, 112)), or estimating genotype likelihoods that incorporate the uncertainty of the data (71). In addition, to

decrease the effect of aDNA-specific error, it is common to either 1) restrict analyses to transversion polymorphisms not subject to deamination error, 2) re-scale mapping quality before variant calling (61), or 3) hard-mask or remove ends of fragments (65) according to likelihood of a base call representing a deaminated site.

Different genetic markers from nuclear and organellar genomes can be targeted to learn about the evolutionary history of plants from archaeobotanical or herbarium samples, including pre-defined single nucleotide polymorphisms, specific genes, genomic tracts, or complete genomes. Historically, due to their high copy number and short length, uniparental organellar markers have been commonly typed in ancient DNA studies. In plants, the plastid genome has been frequently used in phylogenetic studies involving archaeobotanical and herbarium samples (67, 74, 148, 151, 152). However, for some plant species, obtaining useful information from the mitochondrial and plastid genomes from ancient samples can be challenging (155). Although their high copy number makes them attractive targets in degraded DNA sources, these genomes are prone to frequent genomic rearrangements and cytoplasmic-nuclear gene transfer, complicating and limiting analysis (155).

Evolutionary inference based on population allele frequencies represents an additional challenge in plant palaeogenomics. Due to their different reproduction and cultivation mechanisms (sexual or asexual, self-fertilizing or outcrossing, or annual or perennial), the properties of plant populations differ between species. As such, these mechanisms impact population histories and are relevant when considering the sampling strategies and potential analyses. This is particularly relevant for domesticated plant species, a major focus of plant palaeogenomics. Cultivated plants often have strong population structure and individuals from the same group are highly related (89). Thus, obtaining a sample that properly represents the species or population variability is complicated, as is analyzing the data in a standard population genetics framework. Furthermore, the sampling process is limited by specimen availability when working with archaeobotanical remains or herbarium samples. Thus, sampling multiple individuals from the same population is not always possible. To date, there are very few instances of studies with truly population-level data from ancient plant specimens. These include two genomic surveys of ancient maize from two archaeological sites in the US Southwest where signatures of local adaptation to high altitude and production of sugar and starch were identified and traced through time (29, 145).

4. Applications of ancient plant genomics

Domestication archaeogenomics

Ancient and historic plant genomic research is especially impactful for unraveling complex evolutionary trajectories by augmenting the present-day view of biodiversity with molecular data sampled through time. Plant domestication is one of the most formative processes in human history, emerging independently around the world in the early Holocene and critically important for anthropology, archaeology, ecology, and biodiversity studies. Domestic plants are well represented in the archaeobotanical record, and their dynamic relationships with humans over recent timescales provide opportunities for real-time evaluation of how domesticated systems evolve. As such, domestication is chief among target study areas

being approached with renewed perspective and empirical rigor in the archaeogenomic era (4), and has been a key focus of ancient plant DNA research throughout the field's brief history.

Maize, a staple crop produced at over one billion tons per year globally, has been the most frequent target of archaeogenomic domestication research owing to excellent preservation throughout much of its precolonial range in the Americas. Archaeobotanical microfossils and modern microsatellite-based genetic analysis suggest that domestication began ~9,000 years before the present (BP) in the lowland Balsas River valley of southern Mexico (84, 106). Sporadic macro-remains outside of the domestication center in Mexico and Central America begin at ~6,500 BP (105), and maize becomes the dominant staple of Mesoamerica only ~3,000 BP, 6ky after the onset of domestication (64). PCR-based studies with maize established the timing with which key domestication traits became acquired and fixed near the domestication center (59), and unraveled routes of dispersal and distribution of biodiversity in South America (38). Genome capture methods were later deployed with archaeological material in the southwest and Mexico to refine the model for human-mediated movement of plants northward from the domestication center, and regional adaptation over time mediated by adaptive introgression from crop-wild relatives and long-term habitation of the desert southwest environment (29).

As archaeogenomic studies have increasingly exploited genome-wide datasets, maize has been targeted for four ancient genome sequencing experiments to date. Two focused on ~5,300 year old maize from near the source region in Mexico (112, 147), each revealing a mix of wild-type and maize-like alleles at key domestication loci four millennia after the onset of domestication. These findings underscored the protracted arc of domestication as a long-term evolutionary process with extensive ongoing gene flow from wild populations and weak, landscape-level selection on suites of traits valuable to human cultivators. A third, analyzing maize genomes from the Turkey Pen site in Colorado, reconstructed aspects of growth phenotype to demonstrate a marginal level of adaptation to local conditions, shedding light on the selective stress involved with rapid human-driven dispersal of crop plants into new landscapes (145). Most recently, the fourth study analyzed traditional and archaeological biodiversity to suggest that the first maize carried into South America was in a state of partial domestication ~7,000 years ago (65). This study suggests a stratified model of domestication that began in a single large Mesoamerican gene pool and continued with parallel human selection pressures in multiple regions.

Near Eastern wheat domestication was another early focus area for an aDNA approach, but because archaeobotanical wheat is preserved almost exclusively by charring, results were infrequent and based entirely on PCR experiments (reviewed in (96)). Nonetheless these studies elucidate some early population dynamics of wheat traits important in domestication. The first genomic analysis of ancient plants was aimed at understanding genome evolution during cotton domestication (97). This experiment, showing changes in transposable element composition through archaeological time, suggested a punctuated equilibrium-like model for recent cotton evolution where massive genomic compositional events accompanied domestication with implications for diversity and biology. Other archaeological DNA studies have focused on domestication biogeography in barley (82), the mutation load and adaptive hybridization in sorghum through an Egyptian time-series (139), historical biogeography and biodiversity in North and South American *Chenopodium* species (68, 161), evolutionary ecology

and domestication in squashes (67) and bottle gourds (34, 66), the evolution and long-term management of domesticated grapevines (111), and the domestication gene pool of sunflowers (152).

Genetic analysis of herbarium specimens has been important for our understanding of the domestication history of both the potato (*Solanum tuberosum*) and the sweet potato (*Ipomoea batatas*). Roullier et al. (122) performed genetic analyses on 57 historical herbarium collections of sweet potato, along with over 1000 modern samples, finding evidence of pre-Columbian movement of the species from South America to Oceania. To further elucidate the plant's evolutionary relationships and domestication history, Muñoz-Rodríguez et al. (87) used genome skimming and target capture approaches to sequence complete plastid genomes and over 600 nuclear genes from 75 herbarium specimens and 125 germplasm repository tissue samples of sweet potato and its crop wild relatives. With this data, the authors conclude that the modern domestic crop evolved solely from an *Ipomoea trifida* ancestor. More recently, Gutaker et al. (53) used a targeted capture approach to deep-sequence selected regions of the potato genome, including the plastid genome and over 330 genes related to photoperiod response. This approach was applied to generate data from 88 samples representing contemporary and historical diversity of the domestic potatoes, including some herbarium specimens originally collected by Charles Darwin and one collected 359 years ago. The authors report a major population turnover occurred in European potato crops after the year 1750, during which Chilean genotypes admixed into the existing diversity derived from Andean landraces. In addition, they detected adaptation of European potatoes to longer days by quantifying temporal change in allele frequencies at the StCDF1 gene.

In addition to regionally and taxonomically focused domestication studies, the interface of archaeology and genetics is being explored to re-think commonalities of domestication across diverse ecosystems, and the fundamental evolutionary dynamics involved. For example, the frequency of simple Mendelian traits such as a non-shattering rachis in wheat and increased seed size in barley can be traced through thousands of years of archaeobotanical deposits, and selection coefficients can be inferred from the rate of change (41). The results demonstrate very weak selection pressures over extremely protracted periods for fundamental domestication traits, commensurate with natural selection rather than focused breeding efforts. Although not based on ancient DNA, this archaeological and genetic finding establishes predictions testable through ancient crop genomics. For example, weak selection on many loci spanning thousands of years requires that a substantial amount of the biodiversity in wild populations must be maintained in crop progenitors for efficacious selection and a robust domestic population. This prediction runs counter to traditional predictions of a substantial genetic bottleneck during domestication. However, initial analysis across three grain crops suggests that a classical bottleneck may not be required or beneficial during the domestication process, and that gradual genetic erosion through serial founder effects and post-domestication breeding are the primary drivers of diversity loss in crop species (5).

Domestication has been a topic of special evolutionary interest beginning with the opening chapter of *On the Origin of Species* because of the dramatic outcomes when plants and animals adapt to the human environment with mutualistic benefits. Ancient DNA provides the unique opportunity to interrogate the domestication process in archaeological real time, rather than inferring its complexities solely from modern plants shaped by the post-industrial

world. The genomic era in particular has introduced unprecedented levels of nuance to our biological and anthropological understanding of crop domestication.

Evolution and ecology

Ancient and historic DNA can integrate a temporal perspective with other long-term or transient ancient processes. For example, the analysis of invasive plant species offers to help elucidate fundamental evolutionary questions, such as convergent adaptation to similar environments, and whether *de novo* mutations or standing variation contribute most to adaptation to new ranges. The incorporation of historic herbarium samples offers an opportunity to study these processes as population structure, allele frequencies, and *de novo* mutations can be directly observed over time (36). In a study of *Arabidopsis thaliana*, a self-fertilizing plant that is native to Eurasia and recently colonized North America, the authors used herbarium samples collected between 1863 and 1993 in combination with modern specimens from North America and were able to identify several *de novo* mutations that had risen to immediate or high frequencies (36). Quantitative variation in root traits could be explained by a small number of *de novo* mutations of intermediate frequencies. The authors were also able to estimate the substitution rate by using the collection dates of the herbarium specimens. They estimated the time of the most recent common ancestor, which relates to the colonization of the plant in North America.

Herbarium specimens not only contain DNA from the specimens themselves, but also the metagenomic DNA of microorganisms that populated the sample (164). Using samples from different times and locations, can help elucidate origins and introduction pathways as well as population dynamics of plant pathogens (15). Shotgun-sequenced genomes of the pathogenic oomycete *Phytophthora infestans*, which causes late potato blight, showed the genetic structure within Europe changed dramatically since the potato's introduction. The *Phytophthora* lineage that caused the Irish famine (HERB-1) was completely replaced by the US-1 lineage in Europe, which later caused the twentieth-century global outbreak of potato famine (80, 165). By using collection dates of historic herbarium specimens to calibrate branch length and estimate divergence times in a Bayesian phylogenetic framework, the diversification of the pathogen could be related to the Spanish Conquest of Central and South America.

Further, herbicides play an important role in modern-day agriculture, but herbicide resistance evolved in crop fields is increasingly problematic. It has been proposed that resistance alleles already present in a population are adaptively advantageous over *de novo* mutations due to lower negative pleiotropic effects on the life cycle of weeds. In a study of the grass weed *Alopecurus myosuroides* using herbarium specimens collected between 1788 and 1975, amplicon sequencing revealed that a herbicide resistance allele was already present in a specimen from 1888, predating the use of herbicides (31). Whole genome sequencing with HTS technologies can further enable the identification of other resistance alleles, especially non-target-site resistance (NTSR), in the future to better understand the population genetic context and consequences of herbicide resistance (123).

Finally, the characterization of past landscapes, ecosystems, and human practices through sedaDNA analysis has expanded our biomolecular toolkit to allow for the detection of organisms that leave few fossil traces. This avenue of research was initially explored through metabarcoding but later fully demonstrated by shotgun sequencing ancient metagenomes (103,

135). For the sedaDNA community, the development of shotgun sequencing for metagenomics is arguably as important as the innovation of PCR for the larger aDNA field. Still, this approach is in its infancy, and the full potential and limits of shotgun metagenomics for sedaDNA remain to be explored. Moving forward, shotgun metagenomics should rival metabarcoding for its taxonomic sensitivity, but with the additional benefit of characterizing complete ancient genomes to reveal new insights on past environment with organisms across all trophic layers (98).

5. Summary and perspectives

The first 35 years of ancient DNA research has established the basic expectations and limits of DNA preservation, utilized targeted strategies in hypothesis-driven studies across study areas, and embraced genomic methods to fully integrate diversity through time into molecular genetic frameworks. Plant-focused aDNA research has played key roles throughout this coming-of-age process for archaeogenomics and paleogenomics, and the genomic revolution has forced a fundamental re-calibration toward important evolutionary processes such as plant domestication. Moreover, aDNA methodologies have unlocked herbaria and museum collections as repositories of long-term genetic variation, giving insight into population dynamics and biodiversity through time, and making possible large-scale surveys of genomic variance and adaptation based on curated plant tissues. Ancient plant genomics leverages the power of diachronic sampling, a key strategy in archaeology and paleontology, combined with the tremendous information landscape of genomes. We are still in the beginning phases of deploying plant archaeogenomics to explore novel and long-standing research questions across fields. As within other focus areas of ancient DNA research, genomics allows us access to unprecedented complexity in the history of plants and people, and we continue to learn how aDNA rejects simplistic models of evolution. The admonition to “be sparing with Occam’s razor” (54) will describe the exciting next years of ancient plant genomics.

Acknowledgements
We thank Tom Gilbert for helpful discussions and feedback on the manuscript.

Literature Cited

1. Adams RP, Sharma LN. 2010. DNA from herbarium specimens: I. Correlation of DNA size with specimen age. *Phytologia*. 92:346–53
2. Allaby RG, Jones MK, Brown TA. 1994. DNA in charred wheat grains from the Iron Age hillfort at Danebury, England. *Antiquity*
3. Allaby RG, O’Donoghue K, Sallares R, Jones MK, Brown TA. 1997. Evidence for the survival of ancient DNA in charred wheat seeds from European archaeological sites. *Anc. Biomol.*
4. Allaby RG, Smith O, Kistler L. 2018. Archaeogenomics and Crop Adaptation. , pp. 189–203
5. Allaby RG, Ware RL, Kistler L. 2019. A re-evaluation of the domestication bottleneck from archaeogenomic evidence. *Evol. Appl.* 12(1):29–37
6. Allentoft ME, Collins M, Harker D, Haile J, Oskam CL, et al. 2012. The half-life of DNA in bone. *Meas. decay Kinet. 158 dated Foss.* 279(1748):4724–33
7. Alsos IG, Lammers Y, Yoccoz NG, Jørgensen T, Sjögren P, et al. 2018. Plant DNA

metabarcoding of lake sediments: How does it represent the contemporary vegetation. *PLoS One*
8. Ames M, Spooner DM. 2008. DNA from herbarium specimens settles a controversy about origins of the European potato. *Am. J. Bot.* 95(2):252–57
9. Anderson-Carpenter LL, McLachlan JS, Jackson ST, Kuch M, Lumibao CY, Poinar HN. 2011. Ancient DNA from lake sediments: Bridging the gap between paleoecology and genetics. *BMC Evol. Biol.*
10. Ávila-Arcos MCMC, Cappellini E, Romero-Navarro JA, Wales N, Moreno-Mayar JVV, et al. 2011. Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA. *Sci. Rep.* 1:
11. Bakker FT, Lei D, Yu J, Mohammadin S, Wei Z, et al. 2016. Herbarium genomics: Plastome sequence assembly from a range of herbarium specimens using an Iterative Organelle Genome Assembly pipeline. *Biol. J. Linn. Soc. Lond.* 117(1):33–43
12. Beaman RS, Cellinese N. 2012. Mass digitization of scientific collections: New opportunities to transform the use of biological specimens and underwrite biodiversity science. *Zookeys*, pp. 7–17
13. Besnard G, Christin P-A, Malé P-JG, Lhuillier E, Lauzeral C, et al. 2014. From museums to genomics: old herbarium specimens shed light on a C3 to C4 transition. *J. Exp. Bot.* 65(22):6711–21
14. Besnard G, Gaudeul M, Lavergne S, Muller S, Rouhan G, et al. 2018. Herbarium-based science in the twenty-first century. *Bot. Lett.* 165(3–4):323–27
15. Bieker VC, Martin MD. 2018. Implications and future prospects for evolutionary analyses of {DNA} in historical herbarium collections. *Bot. Lett.* 165(3–4):409–18
16. Bilinski P, Albert PS, Berg JJ, Birchler JA, Grote MN, et al. 2018. Parallel altitudinal clines reveal trends in adaptive evolution of genome size in Zea mays. *PLoS Genet.*
17. Boast AP, Weyrich LS, Wood JR, Metcalf JL, Knight R, Cooper A. 2018. Coprolites reveal ecological interactions lost with the extinction of New Zealand birds. *Proc. Natl. Acad. Sci.* 115(7):1546
18. Briggs AW, Stenzel U, Johnson PLF, Green RE, Kelso J, et al. 2007. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci.* 104(37):14616–21
19. Brown T a, Allaby RG, Sallares R, Jones G. 1998. Ancient DNA in charred wheats: Taxonomic identification of mixed and single grains
20. Brown TA, Cappellini E, Kistler L, Lister DL, Oliveira HR, et al. 2015. Recent advances in ancient DNA research and their implications for archaeobotany. *Veg. Hist. Archaeobot.* 24(1):207–14
21. Bunning SL, Jones G, Brown TA. 2012. Next generation sequencing of DNA in 3300-year-old charred cereal grains. *J. Archaeol. Sci.*
22. Campos PF, Craig OE, Turner-Walker G, Peacock E, Willerslev E, Gilbert MTP. 2012. DNA in ancient bone – Where is it located and how should we extract it? *Ann. Anat. - Anat. Anzeiger*. 194(1):7–16
23. Cappellini E, Gilbert MTP, Geuna F, Fiorentino G, Hall A, et al. 2010. A multidisciplinary study of archaeological grape seeds. *Naturwissenschaften*, pp. 205–17
24. Carøe C, Gopalakrishnan S, Vinner L, Mak SST, Sinding MHS, et al. 2018. Single-tube library preparation for degraded DNA. *Methods Ecol. Evol.* 9(2):410–19
25. Clarke CL, Edwards ME, Brown AG, Gielly L, Lammers Y, et al. 2019. Holocene floristic diversity and richness in northeast Norway revealed by sedimentary ancient DNA (sedaDNA) and pollen. *Boreas*
26. Cooper A, Poinar HN. 2000. Ancient DNA: Do it right or not at all. *Science (80-.).* 289(5482):1139
27. Cota-Sánchez JH, Remarchuk K, Ubayasena K. 2006. Ready-to-use DNA extracted with a CTAB method adapted for herbarium specimens and mucilaginous plant tissue. *Plant*

Mol. Biol. Report. 24(2):161
28. Cubero OF, Crespo A, Fatehi J, Bridge PD. 1999. DNA extraction and PCR amplification method suitable for fresh, herbarium-stored, lichenized, and other fungi. *Plant Syst. Evol.* 216(3):243–49
29. da Fonseca RR, Smith BD, Wales N, Cappellini E, Skoglund P, et al. 2015. The origin and evolution of maize in the American Southwest. *Nat. Plants*. 1:1–5
30. Dabney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, et al. 2013. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl. Acad. Sci. U. S. A.* 110(39):15758–63
31. Délye C, Deulvot C, Chauvel B. 2013. {DNA} analysis of herbarium Specimens of the grass weed Alopecurus myosuroides reveals herbicide resistance pre-dated herbicides. *PLoS One*. 8(10):e75117
32. Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19:11–15
33. Elbaum R, Melamed-bessudo C, Tuross N, Levy AA, Weiner S. 2009. New methods to isolate organic materials from silicified phytoliths reveal fragmented glycoproteins but no DNA. *Water*. 193:11–19
34. Erickson DL, Smith BD, Clarke AC, Sandweiss DH, Tuross N. 2005. An Asian origin for a 10,000-year-old domesticated plant in the Americas. *Proc. Natl. Acad. Sci. U. S. A.* 102(51):18315–20
35. Erkens RHJ, Cross H, Maas JW, Hoenselaar K, Chatrou LW. 2008. Assessment of age and greenness of herbarium specimens as predictors for successful extraction and amplification of DNA. *Blumea - Biodiversity, Evol. Biogeogr. Plants*. 53(2):407–28
36. Exposito-Alonso M, Becker C, Schuenemann VJ, Reiter E, Setzer C, et al. 2018. The rate and potential relevance of new mutations in a colonizing plant lineage. *PLoS Genet.* 14(2):e1007155
37. Fordyce SL, Ávila-Arcos MC, Rasmussen M, Cappellini E, Romero-Navarro JA, et al. 2013. Deep sequencing of RNA from ancient maize kernels. *PLoS One*. 8(1):e50961
38. Freitas FO, Bandel G, Allaby RG, Brown T a. 2003. DNA from primitive maize landraces and archaeological remains: implications for the domestication of maize and its expansion into South America. *J. Archaeol. Sci.* 30(7):901–8
39. Freitas FO, Bendel G, Allaby RG, Brown T a. 2003. DNA from primitive maize landraces and archaeological remains: implications for the domestication of maize and its expansion into South America. *J. Archaeol. Sci.* 30(7):901–8
40. Fricker EJ, Spigelman M, Fricker CR. 1997. The detection of Escherichia coli DNA in the ancient remains of Lindow Man using the polymerase chain reaction. *Lett. Appl. Microbiol.* 24(5):351–54
41. Fuller DQ, Denham T, Arroyo-Kalin M, Lucas L, Stevens CJ, et al. 2014. Convergent evolution and parallelism in plant domestication revealed by an expanding archaeological record. *Proc. Natl. Acad. Sci.* 111(17):6147–52
42. Fulton TL, Shapiro B. 2019. Setting up an ancient DNA laboratory. In *Methods in Molecular Biology*
43. Gansauge M-T, Gerber T, Glocke I, Korlević P, Lippik L, et al. 2017. Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic Acids Res.* 45(10):e79–e79
44. Gilbert MTP, Bandelt HJ, Hofreiter M, Barnes I. 2005. Assessing ancient DNA studies. *Trends Ecol. Evol.* 20(10):541–44
45. Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust EM, et al. 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotech*. 27(2):182–89
46. Golenberg EM, Giannasi DE, Clegg MT, Smiley CJ, Durbin M, et al. 1990. Chloroplast

DNA sequence from a miocene Magnolia species.
47. Goloubinoff P, Paabo S, Wilson AC. 1993. Evolution of maize inferred from sequence diversity of an Adh2 gene segment from archaeological specimens. *Proc. Natl. Acad. Sci.*
48. Graham CH, Ferrier S, Huettman F, Moritz C, Peterson AT. 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends Ecol. Evol.* 19(9):497–503
49. Grass RN, Heckel R, Puddu M, Paunescu D, Stark WJ. 2015. Robust Chemical Preservation of Digital Information on DNA in Silica with Error-Correcting Codes. *Angew. Chemie Int. Ed.* 54(8):2552–55
50. Green JE, Speller FC. 2017. Novel Substrates as Sources of Ancient DNA: Prospects and Hurdles. *Genes (Basel).* 8(7):
51. Gutaker RM, Burbano HA. 2017. Reinforcing plant evolutionary genomics using ancient DNA. *Curr. Opin. Plant Biol.* 36:38–45
52. Gutaker RM, Reiter E, Furtwängler A, Schuenemann VJ, Burbano HA. 2017. Extraction of ultrashort DNA molecules from herbarium specimens. *Biotechniques*. 62(2):76–79
53. Gutaker RM, Weiß CL, Ellis D, Anglin NL, Knapp S, et al. 2019. The origins and adaptation of European potatoes reconstructed from historical genomes. *Nat. Ecol. Evol.*
54. Haber M, Mezzavilla M, Xue Y, Tyler-Smith C. 2016. Ancient DNA and the rewriting of human history: be sparing with Occam's razor. *Genome Biol.* 17(1):1
55. Hansson MC, Foley BP. 2008. Ancient DNA fragments inside Classical Greek amphoras reveal cargo of 2400-year-old shipwreck. *J. Archaeol. Sci.* 35(5):1169–76
56. Hart ML, Forrest LL, Nicholls JA, Kidner CA. 2016. Retrieval of hundreds of nuclear loci from herbarium specimens. *Taxon*. 65(5):1081–92
57. Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC. 1984. DNA sequences from the quagga, an extinct member of the horse family. *Nature*
58. Iñiguez AM, Reinhard K, Carvalho Gonçalves ML, Ferreira LF, Araújo A, Paulo Vicente AC. 2006. SL1 RNA gene recovery from Enterobius vermicularis ancient DNA in pre-Columbian human coprolites. *Int. J. Parasitol.* 36(13):1419–25
59. Jaenicke-Despres V, Buckler ES, Smith BD, Gilbert MTP, Cooper A, et al. 2003. Early allelic selection in maize as revealed by ancient DNA. *Science (80-.).* 302(5648):1206–8
60. Johnson MG, Pokorny L, Dodsworth S, Botigué LR, Cowan RS, et al. 2019. A Universal Probe Set for Targeted Sequencing of 353 Nuclear Genes from Any Flowering Plant Designed Using k-Medoids Clustering. *Syst. Biol.* 68(4):594–606
61. Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*. 29(13):1682–84
62. Jørgensen T, Kjær KH, Haile J, Rasmussen M, Boessenkool S, et al. 2012. Islands in the ice: Detecting past vegetation on Greenlandic nunataks using historical records and sedimentary ancient DNA Meta-barcoding. *Mol. Ecol.*
63. Kapusta A, Suh A, Feschotte C. 2017. Dynamics of genome size evolution in birds and mammals. *Proc. Natl. Acad. Sci.* 114(8):E1460–69
64. Kennett DJ, Thakar HB, Vanderwarker AM, Webster DL, Culleton BJ, et al. 2017. High-precision chronology for Central American maize diversification from El Gigante rockshelter, Honduras. *Proc. Natl. Acad. Sci. U. S. A.* 114(34):
65. Kistler L, Maezumi SY, Gregorio de Souza J, Przelomska NAS, Malaquias Costa F, et al. 2018. Multiproxy evidence highlights a complex evolutionary legacy of maize in South America. *Science (80-.).* 362(6420):1309–13
66. Kistler L, Montenegro Á, Smith BD, Gifford JA, Green RE, et al. 2014. Transoceanic drift and the domestication of African bottle gourds in the Americas. *Proc. Natl. Acad. Sci. U. S. A.* 111(8):
67. Kistler L, Newsom LA, Ryan TM, Clarke AC, Smith BD, Perry GH. 2015. Gourds and

squashes (Cucurbita spp.) adapted to megafaunal extinction and ecological anachronism through domestication. *Proc. Natl. Acad. Sci. U. S. A.* 112(49):15107–12
68. Kistler L, Shapiro B. 2011. Ancient DNA confirms a local origin of domesticated chenopod in eastern North America. *J. Archaeol. Sci.* 38(12):
69. Kistler L, Ware R, Smith O, Collins M, Allaby RG. 2017. A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Res.* 45(11):
70. Konrade L, Shaw J, Beck J. 2019. A rangewide herbarium-derived dataset indicates high levels of gene flow in black cherry (). *Ecol. Evol.* 9(3):975–85
71. Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*. 15(1):356
72. Korpelainen H, Pietiläinen M. 2019. The effects of sample age and taxonomic origin on the success rate of {DNA} barcoding when using herbarium material. *Plant Syst. Evol.* 305(4):319–24
73. Krause J, Unger T, Noçon A, Malaspinas A-S, Kolokotronis S-O, et al. 2008. Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene-Pliocene boundary. *BMC Evol. Biol.* 8:220
74. Kumagai M, Kanehara M, Shoda S, Fujita S, Onuki S, et al. 2016. Rice Varieties in Archaic East Asia: Reduction of Its Diversity from Past to Present Times. *Mol. Biol. Evol.*
75. Leitch I, Johnston E, Pellicer J, Hidalgo O, Bennett M. 2019. *Plant DNA C-values Database*
76. Leitch IJ, Bennett MD. 1997. Polyploidy in angiosperms. *Trends Plant Sci.* 2(12):470–76
77. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*
78. Lindahl T. 1993. Instability and decay of the primary structure of DNA. *Nature*. 362:709–15
79. Loreille O, Roumat E, Verneau O, Bouchet F, Hänni C. 2001. Ancient DNA from Ascaris: extraction amplification and sequences from eggs collected in coprolites. *Int. J. Parasitol.* 31(10):1101–6
80. Martin MD, Cappellini E, Samaniego J a, Zepeda ML, Campos PF, et al. 2013. Reconstructing genome evolution in historic samples of the Irish potato famine pathogen. *Nat. Commun.* 4:2172
81. Martin MD, Vieira FG, Ho SYW, Wales N, Schubert M, et al. 2016. Genomic characterization of a south American phytophthora hybrid mandates reassessment of the geographic origins of phytophthora infestans. *Mol. Biol. Evol.* 33(2):478–91
82. Mascher M, Schuenemann VJ, Davidovich U, Marom N, Himmelbach A, et al. 2016. Genomic analysis of 6,000-year-old cultivated grain illuminates the domestication history of barley. *Nat. Genet.* 48:1089
83. Matheson CD, Gurney C, Esau N, Lehto R. 2010. Assessing PCR inhibition from humic substances. *Open Enzym. Inhib. J.* 3(1):
84. Matsuoka Y, Vigouroux Y, Goodman MM, G JS, Buckler E, et al. 2002. A single domestication for maize shown by multilocus. *Sci. York*. 99(9):6080–84
85. Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Springs Harb. Protoc.*
86. Miller W, Drautz DI, Ratan A, Pusey B, Qi J, et al. 2008. Sequencing the nuclear genome of the extinct woolly mammoth. *Nature*. 456(7220):387–90
87. Muñoz-Rodr\'\iguez P, Carruthers T, Wood JRI, Williams BRM, Weitemier K, et al. 2018. Reconciling Conflicting Phylogenies in the Origin of Sweet Potato and Dispersal to Polynesia. *Curr. Biol.* 28(8):1246--1256.e12
88. Murray DC, Pearson SG, Fullagar R, Chase BM, Houston J, et al. 2012. High-throughput sequencing of ancient plant and mammal DNA preserved in herbivore middens. *Quat. Sci. Rev.*

89. Myles S, Boyko AR, Owens CL, Brown PJ, Grassi F, et al. 2011. Genetic structure and domestication history of the grape. *Proc. Natl. Acad. Sci.*
90. Nielsen R, Paul JS, Albrechtsen A, Song YS. 2011. Genotype and SNP calling from next-generation sequencing data
91. Nistelberger HM, Smith O, Wales N, Star B, Boessenkool S. 2016. The efficacy of high-throughput sequencing and target enrichment on charred archaeobotanical remains. *Sci. Rep.*
92. Olofsson JK, Bianconi M, Besnard G, Dunning LT, Lundgren MR, et al. 2016. Genome biogeography reveals the intraspecific spread of adaptive mutations for a complex trait. *Mol. Ecol.* 25(24):6107–23
93. Olofsson JK, Cantera I, de Paer C, Hong-Wa C, Zedane L, et al. 2019. Phylogenomics using low-depth whole genome sequencing: A case study with the olive tribe. *Mol. Ecol. Resour.* 66:786
94. Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, et al. 2013. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*. 499(7456):74–78
95. Paabo S. 1989. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proc. Natl. Acad. Sci.* 86(6):1939–43
96. Palmer S a, Smith O, Allaby RG. 2012. The blossoming of plant archaeogenetics. *Ann. Anat.* 194(1):146–56
97. Palmer SA, Clapham AJ, Rose P, Freitas FO, Owen BD, et al. 2012. Archaeogenomic evidence of punctuated genome evolution in gossypium. *Mol. Biol. Evol.* 29(8):2031–38
98. Parducci L, Bennett KD, Ficetola GF, Alsos IG, Suyama Y, et al. 2017. Ancient plant DNA in lake sediments
99. Parducci L, Nota K, Wood J. 2018. Reconstructing Past Vegetation Communities Using Ancient DNA from Lake Sediments
100. Parducci L, Suyama Y, Lascoux M, Bennett KD. 2005. Ancient DNA from pollen: A genetic record of population history in Scots pine. *Mol. Ecol.*
101. Pearsall DM. 2000. *Paleoethnobotany : a handbook of procedures*. San Diego: Academic Press. xxxii, 700 p. pp. 2nd ed.
102. Pedersen MW, Ginolhac A, Orlando L, Olsen J, Andersen K, et al. 2013. A comparative study of ancient environmental DNA to pollen and macrofossils from lake sediments reveals taxonomic overlap and additional plant taxa. *Quat. Sci. Rev.*
103. Pedersen MW, Ruter A, Schweger C, Friebe H, Staff RA, et al. 2016. Postglacial viability and colonization in North America's ice-free corridor. *Nature*
104. Piperno DR. 2016. Phytolith radiocarbon dating in archaeological and paleoecological research: a case study of phytoliths from modern Neotropical plants and a review of the previous dating evidence. *J. Archaeol. Sci.* 68:54–61
105. Piperno DR, Flannery K V, Cave N. 2001. The earliest archaeological maize (Zea mays L .) from highland Mexico : New accelerator mass spectrometry. *Archaeology*. 98(4):2101–3
106. Piperno DR, Ranere AJ, Holst I, Iriarte J, Dickau R. 2009. Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. *Proc. Natl. Acad. Sci.* 106(13):5019–24
107. Poinar HN, Hofreiter M, Spaulding WG, Martin PS, Stankiewicz BA, et al. 1998. Molecular Coproscopy: Dung and Diet of the Extinct Ground Sloth Nothrotheriops shastensis. *Science (80-.).* 281(5375):402
108. Poinar HN, Schwarz C, Qi J, Shapiro B, MacPhee RDE, et al. 2006. Metagenomics to Paleogenomics: Large-Scale Sequencing of Mammoth DNA. *Science (80-.).* 311(5759):392–94
109. Pont C, Wagner S, Kremer A, Orlando L, Plomion C, Salse J. 2019. Paleogenomics:

Reconstruction of plant evolutionary trajectories from modern and ancient DNA
110. Pyke GH, Ehrlich PR. 2010. Biological collections and ecological/environmental research: A review, some observations and a look to the future. *Biol. Rev. Camb. Philos. Soc.* 85(2):247–66
111. Ramos-Madrigal J, Runge AKW, Bouby L, Lacombe T, Samaniego Castruita JA, et al. 2019. Palaeogenomic insights into the origins of French grapevine diversity. *Nat. Plants*
112. Ramos-Madrigal J, Smith BD, Moreno-Mayar JV, Gopalakrishnan S, Ross-Ibarra J, et al. 2016. Genome Sequence of a 5,310-Year-Old Maize Cob Provides Insights into the Early Stages of Maize Domestication. *Curr. Biol.* 26(23):3195–3201
113. Rasmussen M, Cummings LS, Gilbert MTP, Bryant V, Smith C, et al. 2009. Response to Comment by Goldberg et al. on "DNA from Pre-Clovis Human Coprolites in Oregon, North America." *Science (80-.).* 325(5937):148
114. Reinhard KJ, Bryant VM. 1992. Coprolite Analysis: A Biological Perspective on Archaeology. *Archaeol. Method Theory*. 4:245–88
115. Rogers SO, Bendich AJ. 1985. Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol. Biol.*
116. Rohland N, Glocke I, Aximu-Petri A, Meyer M. 2018. Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nat. Protoc.* 13(11):2447–61
117. Rohland N, Harney E, Mallick S, Nordenfelt S, Reich D. 2014. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Philos. Trans. R. Soc. B Biol. Sci.* 370(1660):20130624–20130624
118. Rohland N, Hofreiter M. 2007. Comparison and optimization of ancient DNA extraction. *Biotechniques*. 42(3):343–52
119. Rollo F. 1985. Characterisation by molecular hybridization of RNA fragments isolated from ancient (1400 B.C.) seeds. *Theor. Appl. Genet.* 71(2):330–33
120. Rollo F, Amici A, Salvi R, Garbuglia A. 1988. Short but faithful pieces of ancient DNA
121. 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(20):12594
122. Roullier C, Benoit L, McKey DB, Lebot V. 2013. Historical collections reveal patterns of diffusion of sweet potato in Oceania obscured by modern plant movements and recombination. *Proc. Natl. Acad. Sci. U. S. A.*, pp. 1–6
123. Sablok G, Amiryousefi A, He X, Hyvönen J, Poczai P. 2019. Sequencing the plastid genome of giant ragweed (Ambrosia trifida, Asteraceae) from a herbarium specimen. *Front. Plant Sci.*
124. Sánchez Barreiro F, Vieira FG, Martin MD, Haile J, Gilbert MTP, Wales N. 2016. Characterizing restriction enzyme-associated loci in historic ragweed (Ambrosia artemisiifolia) voucher specimens using custom-designed {RNA} probes. *Mol. Ecol. Resour.* 17:209–20
125. Santiago-Rodriguez TM, Fornaciari G, Luciani S, Dowd SE, Toranzos GA, et al. 2015. Gut Microbiome of an 11th Century A.D. Pre-Columbian Andean Mummy. *PLoS One*. 10(9):e0138135
126. Särkinen T, Staats M, Richardson JE, Cowan RS, Bakker FT. 2012. How to Open the Treasure Chest? Optimising DNA Extraction from Herbarium Specimens. *PLoS One*. 7(8):e43808
127. Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S. 2012. Temporal Patterns of Nucleotide Misincorporations and DNA Fragmentation in Ancient DNA. *PLoS One*. 7(3):e34131
128. Schubert M, Ginolhac A, Lindgreen S, Thompson JF, AL-Rasheid KAS, et al. 2012. Improving ancient DNA read mapping against modern reference genomes. *BMC*

Genomics
129. Seersholm FV, Pedersen MW, Søe MJ, Shokry H, Mak SST, et al. 2016. DNA evidence of bowhead whale exploitation by Greenlandic Paleo-Inuit 4,000 years ago. *Nat. Commun.*
130. Seguin-Orlando A, Schubert M, Clary J, Stagegaard J, Alberdi MT, et al. 2013. Ligation bias in illumina next-generation DNA libraries: implications for sequencing ancient genomes. *PLoS One*. 8(10):e78575
131. Shapiro B, Hofreiter M. 2012. Preface. In *Ancient DNA: Methods and Protocols*, ed B Shapiro, M Hofreiter. v–vii. New York: Springer
132. Shepherd L, Perrie L. 2014. Genetic analyses of herbarium material: Is more care required? *Taxon*. 63(5):972–73
133. Skoglund P, Mathieson I. 2018. Ancient Genomics of Modern Humans: The First Decade. *Annu. Rev. Genomics Hum. Genet.*
134. Skoglund P, Northoff BH, Shunkov M V, Derevianko AP, Pääbo S, et al. 2014. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc. Natl. Acad. Sci. U. S. A.* 111(6):2229–34
135. Slon V, Hopfe C, Weiß CL, Mafessoni F, De La Rasilla M, et al. 2017. Neandertal and Denisovan DNA from Pleistocene sediments. *Science (80-.).*
136. Smith O, Clapham A, Rose P, Liu Y, Wang J, Allaby RG. 2014. A complete ancient RNA genome: Identification, reconstruction and evolutionary history of archaeological Barley Stripe Mosaic Virus. *Sci. Rep.*
137. Smith O, Clapham AJ, Rose P, Liu Y, Wang J, Allaby RG. 2014. Genomic methylation patterns in archaeological barley show de-methylation as a time-dependent diagenetic process. *Sci. Rep.* 4:
138. Smith O, Momber G, Bates R, Garwood P, Fitch S, et al. 2015. Sedimentary DNA from a submerged site reveals wheat in the British Isles 8000 years ago. *Science (80-.).* 347(6225):998–1001
139. Smith O, Nicholson W V., Kistler L, Mace E, Clapham A, et al. 2019. A domestication history of dynamic adaptation and genomic deterioration in Sorghum. *Nat. Plants*
140. Smith O, Palmer SA, Clapham AJ, Rose P, Liu Y, et al. 2017. Small RNA Activity in Archeological Barley Shows Novel Germination Inhibition in Response to Environment. *Mol. Biol. Evol.* 34(10):2555–62
141. Staats M, Cuenca A, Richardson JE, Vrielink-van Ginkel R, Petersen G, et al. 2011. DNA Damage in Plant Herbarium Tissue. *PLoS One*. 6(12):e28448
142. Stahlschmidt MC, Collin TC, Fernandes DM, Bar-Oz G, Belfer-Cohen A, et al. 2019. Ancient Mammalian and Plant DNA from Late Quaternary Stalagmite Layers at Solkota Cave, Georgia. *Sci. Rep.*
143. Stankiewicz BA, Poinar HN, Briggs DEG, Evershed RP, Poinar J. 1998. Chemical preservation of plants and insects in natural resins. *Proc. R. Soc. B Biol. Sci.*
144. Sutton M, Malik M, Ogram A. 1996. Experiments on the Determination of Gender from Coprolites by DNA Analysis. *J. Archaeol. Sci.* 23(2):263–67
145. Swarts K, Gutaker RM, Benz B, Blake M, Bukowski R, et al. 2017. Genomic estimation of complex traits reveals ancient maize adaptation to temperate North America. *Science*. 357(6350):512–15
146. Tulig M, Tarnowsky N, Bevans M, Anthony Kirchgessner, Thiers BM. 2012. Increasing the efficiency of digitization workflows for herbarium specimens. *Zookeys*, pp. 103–13
147. Vallebueno-Estrada M, Rodríguez-Arévalo I, Rougon-Cardoso A, Martínez González J, García Cook A, et al. 2016. The earliest maize from San Marcos Tehuacán is a partial domesticate with genomic evidence of inbreeding. *Proc. Natl. Acad. Sci.* 113(49):14151–56
148. Van de Paer C, Hong-Wa C, Jeziorski C, Besnard G. 2016. Mitogenomics of

Hesperelaea, an extinct genus of Oleaceae. *Gene*
149. Vandepitte K, Meyer TIMDE, Helsen K, Acker KVAN, Roldán-Ruiz I, et al. 2014. Rapid genetic adaptation precedes the spread of an exotic plant species. *Mol. Ecol.* 23:2157–64
150. VanDerwarker AM, Bardolph DN, Hoppa KM, Thakar HB, Martin LS, et al. 2016. New World Paleoethnobotany in the New Millennium (2000–2013). *J. Archaeol. Res.* 24(2):125–77
151. Wagner S, Lagane F, Seguin-Orlando A, Schubert M, Leroy T, et al. 2018. High-Throughput DNA sequencing of ancient wood. *Mol. Ecol.* 27(5):1138–54
152. Wales N, Akman M, Watson RHB, Sánchez Barreiro F, Smith BD, et al. 2019. Ancient DNA reveals the timing and persistence of organellar genetic bottlenecks over 3,000 years of sunflower domestication and improvement. *Evol. Appl.*
153. Wales N, Andersen K, Cappellini E, Ávila-Arcos MC, Gilbert MTP. 2014. Optimization of DNA recovery and amplification from non-carbonized archaeobotanical remains. *PLoS One*. 9(1):
154. Wales N, Kistler L. 2019. Extraction of Ancient DNA from Plant Remains. In *Ancient DNA: Methods and Protocols*, ed B Shapiro, A Barlow, PD Heintzman, M Hofreiter, JLA Paijmans, AER Soares, pp. 45–55. New York, NY: Springer New York
155. Wales N, Ramos Madrigal J, Cappellini E, Carmona Baez A, Samaniego Castruita JA, et al. 2016. The limits and potential of paleogenomic techniques for reconstructing grapevine domestication. *J. Archaeol. Sci.*
156. Webber WB, Ernest LJ, Vangapandu S. 2011. Mercury exposures in university herbarium collections. *J. Chem. Heal. Saf.* 18(2):9–12
157. Weiß CL, Schuenemann VJ, Devos J, Shirsekar G, Reiter E, et al. 2016. Temporal patterns of damage and decay kinetics of DNA retrieved from plant herbarium specimens. *R. Soc. Open Sci.* 3(6):160239
158. Willerslev E, Cappellini E, Boomsma W, Nielsen R, Hebsgaard MB, et al. 2007. Ancient Biomolecules from Deep Ice Cores Reveal a Forested Southern Greenland. *Science (80-.).* 317(5834):111–14
159. Willerslev E, Davison J, Moora M, Zobel M, Coissac E, et al. 2014. Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature*. 506(7486):47–51
160. Willerslev E, Hansen AJ, Binladen J, Tina B, Gilbert MTP, et al. 2003. Diverse Plant and Animal Genetic Records from Holocene and Pleistocene Sediments. *Science (80-.).* 300(May):791–95
161. Winkel T, Aguirre MG, Arizio CM, Aschero CA, del Pilar Babot M, et al. 2018. Discontinuities in quinoa biodiversity in the dry andes: An 18-century perspective based on allelic genotyping. *PLoS One*
162. Wood JR, Crown A, Cole TL, Wilmshurst JM. 2016. Microscopic and ancient DNA profiling of Polynesian dog (kurī) coprolites from northern New Zealand. *J. Archaeol. Sci. Reports*. 6:496–505
163. Woodward, Weyand N, Bunnell M. 1994. DNA sequence from Cretaceous period bone fragments. *Science (80-.).* 266(5188):1229–32
164. Yoshida K, Sasaki E, Kamoun S. 2015. Computational analyses of ancient pathogen {DNA} from herbarium samples: challenges and prospects. *Front. Plant Sci.* 6(September):771
165. Yoshida K, Schuenemann VJ, Cano LM, Pais M, Mishra B, et al. 2013. The rise and fall of the Phytophthora infestans lineage that triggered the Irish potato famine. *Elife*. 2:
166. Zhou Y, Minio A, Massonnet M, Solares E, Lv Y, et al. 2019. Structural variants, hemizygosity and clonal propagation in grapevines. *bioRxiv*
167. Ziesemer KA, Mann AE, Sankaranarayanan K, Schroeder H, Ozga AT, et al. 2015. Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. *Sci. Rep.* 5:16498

168. Zohary D, Hopf M, Weiss E. 2012. *Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin*. Oxford: Oxford University Press. 264 pp. 4th ed.

Tables

Table 1 – Step-by-step comparison of processes and assumptions in modern genomics and ancient plant DNA research.

Experimental Stage	**Fresh plant DNA**	**Degraded plant DNA (ancient/historic)**
Sample type	• Freshly collected leaf tissue, or other tissue with low concentrations of compounds that inhibit enzyme activity.	• Archaeological samples: desiccated or waterlogged seeds or vegetative structures. • Herbarium vouchers: leaves and tissues which have been treated by curators, potentially with heat or toxic preservatives like mercuric chloride.
Processing	• Performed in standard molecular biology laboratory.	• All pre-PCR steps performed in dedicated degraded DNA facility. • Researchers wear protective coveralls, face masks, and multiple pairs of gloves. • Strict workflow protocols to prevent contamination.
Extraction	• Techniques generally focused on recovery of high molecular weight DNA. • Methods often designed for removing carbohydrates, polyphenols, proteins, and other plant compounds.	• Methods based on plant DNA extractions, but with optimizations for ultrashort DNA (<50bp). • Special attention paid to avoid pigments derived from tissue decomposition and sediment. • Extraction blanks essential to monitor contamination in reagents and laboratory.
DNA manipulation	• Generally none, as most protocols designed for high molecular weight DNA.	• Extracted DNA optionally treated with uracil-DNA glycosylase (UDG) to remove uracil residues resulting from DNA damage (117).
HTS library preparation	• DNA fragmented mechanically or enzymatically. • Ligation of adapters with T-A overhangs most common.	• DNA not sheared, as fragmentation occurs peri- and post-mortem. • Ligation of blunt-end adapters advantageous. • Purification steps optimized to retain ultrashort DNA. • Samples with low endogenous content may require targeted enrichment.
Sequencing	• Paired-end (PE) sequencing	• Single read sequencing (80-100bp) often

	usually preferable to infer the full length of DNA inserts.	sufficient to recover full length of ancient molecules, although PE sequencing improves accurate base calling.
Bioinformatics	• High performance computing cluster generally required, using standard genomic software and custom tools.	• Requires additional steps for authentication of DNA damage and recovering short fragments with minimal mapping bias. • Metagenomic analyses often necessary to infer DNA origins. • Organellar DNA may need to be mapped against microbial panels to exclude environmental taxa.

FigureS

Figure 1 – Timeline of key milestones in ancient plant DNA from 1985 to the present, with reference numbers in Literature Cited.