The untapped potential of macrofossils in ancient plant DNA research

Summary

The rapid development of ancient DNA analysis in the last decades has induced a paradigm shift in ecology and evolution. Driven by a combination of breakthroughs in DNA isolation techniques, high-throughput sequencing, and bioinformatics, ancient genome-scale data for a rapidly growing variety of taxa are now available, allowing researchers to directly observe demographic and evolutionary processes over time. However, the vast majority of paleogenomic studies still focus on human or animal remains. In this article, we make the case for a vast untapped resource of ancient plant material that is ideally suited for paleogenomic analyses: plant remains, such as needles, leaves, wood, seeds, or fruits, that are deposited in natural archives, such as lake sediments, permafrost, or even ice caves. Such plant remains are commonly found in large numbers and in stratigraphic sequence through time and have so far been used primarily to reconstruct past local species presences and abundances. However, they are also unique repositories of genetic information with the potential to revolutionize the fields of ecology and evolution by directly studying microevolutionary processes over time. Here, we give an overview of the current state-of-the-art, address important challenges, and highlight new research avenues to inspire future research.

Introduction

Over the last decades, the analysis of ancient DNA (aDNA) has evolved from the recovery of a few hundred base pairs (bp) of mitochondrial DNA from century-old historical samples (Higuchi et al., 1984) to the sequencing of whole genomes at high coverage (Meyer et al., 2012) and up to a million years old (van der Valk et al., 2021). Both the number of aDNA studies and the number of taxa for which ancient genomic information is available have increased exponentially (Orlando et al., 2021). This tremendous development has mainly been driven by the introduction of high-throughput sequencing (HTS), also referred to as next-generation sequencing (Goodwin et al., 2016), in combination with breakthroughs in DNA isolation techniques (Meyer et al., 2008; Dabney et al., 2013; Schmid et al., 2017; Lendvay et al., 2018b; Rohland et al., 2018). Paleogenomic data allow researchers to directly observe demographic and evolutionary processes over time. This includes population expansions and declines (Lorenzen et al., 2011), range shifts and migrations (Lipson et al., 2017; Moreno-Mayar et al., 2018), adaptation to environmental stressors (Marconi & Perry, 2017; Sandoval-Castellanos et al., 2017; Dehasque et al., 2020), domestication processes (da Fonseca et al., 2015; Scott et al., 2019; Librado et al., 2021), gene flow and hybridization (Der Sarkissian et al., 2013; Schaefer et al., 2016; van der Valk et al., 2021), species extinctions (Lorenzen et al., 2011; Dehasque et al., 2021), and speciation (van der Valk et al., 2021). The groundbreaking results of many aDNA studies have thus led to a paradigm shift in such different fields as archaeology, anthropology, ecology, and evolution.

aDNA studies are based on two different approaches: extracting DNA from either ancient, preserved tissues as starting material (i.e. referred to as aDNA), or from ancient source material, such as lake sediment, soil, permafrost, or ice (Box 1), which contains a mixture of tissue, cells, or extracellular DNA from a wide range of organisms (i.e. referred to as environmental DNA, or in the case of sediment as sedaDNA). Currently, the vast majority of paleogenomic studies still focus on human (Rasmussen et al., 2010; Meyer et al., 2012; Moreno-Mayar et al., 2018) or animal remains (Librado et al., 2021; van der Valk et al., 2021), where mineralized tissues, such as bones and teeth, may provide well-preserved aDNA. Those tissues mainly stem from archaeological sites or natural archives (e.g. permafrost; see Box 1) and allow an individual-based approach. By contrast, most studies focusing on plants so far have rather applied a metabarcoding approach from sedaDNA to reconstruct past floristic communities (Alsos et al., 2016; Parducci et al., 2017; Wang et al., 2021). SedaDNA studies can provide important information about the past presence, diversity, and possibly also abundance of species that cannot be resolved by paleoecological methods such as pollen analysis alone because of the latter’s lower taxonomic resolution (Parducci et al., 2017). A metabarcoding approach uses specific primers (e.g. targeting the chloroplast trnL P6 loop in plants; Taberlet et al., 2007) that amplify short DNA fragments known to harbor a high sequence variability among species (but low variation at the within-species level) that can then be sequenced using HTS and compared with a reference database (Alsos et al., 2016; Parducci et al., 2017; Wang et al., 2021). The advantage here is obviously that a multitude of taxa can be identified from a single sediment sample. Another option for the analysis of sedaDNA samples is to use shotgun metagenomics; that is, sequencing billions of reads from a single sediment sample and then using dedicated bioinformatic analyses to align them with publicly available sequence databases and/or newly established reference libraries to identify the species present (Pedersen et al., 2016; Parducci et al., 2017; Armbricht et al., 2021). However, the efficiency of the method is limited by the fraction of the reads that can be assigned to the target species due to the lack of reference
Box 1 Natural archives as a treasure trove of genetic information

Natural archives, such as lake or mire sediments, fluvial and landslide deposits, ice caves, and permafrost can conserve biological remains for millennia, due to either anoxic conditions or low temperatures that prevent the decomposition of tissues (Birks, 2003; Jørgensen et al., 2012; Leunda et al., 2019, Fig. 1). Indeed, lake and mire archives have been regularly used in paleoecology for a century to reconstruct past species abundances and vegetation composition, mostly based on the analysis of pollen (Birks, 2019). Besides pollen that is ubiquitous in such archives, larger plant remains, such as leaves, needles, bud scales, wood, seeds, or fruits, also occur regularly and can often be determined to species level using morphological and anatomical characteristics. In contrast to pollen, which is readily dispersed by wind over large distances, such macrofossils commonly originate from the local vegetation around a site, thus more reliably indicating local species presence compared with pollen. The same processes that prevent the decomposition of these macrofossils should also minimize the degradation of the DNA within the plant cells, making macrofossils a viable source for past genetic information. There are also other advantages: plant remains are commonly deposited in stratigraphic order and, depending on size, can be directly dated using radiocarbon dating, instead of relying on the dating of surrounding organic material. In many cases, there are numerous remains available from different individuals within a certain time period, which allows the quantitative reconstruction of past local species abundance. This also makes the genetic analysis of entire populations possible and allows comparing the past genetic composition with present-day and/or nonlocal populations. However, even though the potential of macrofossils as a source for aDNA has been recognized for more than a decade (Parducci & Petit, 2004; Gugerli et al., 2005, 2013), only a handful of paleogenetic studies based on macrofossils exist so far. One of the earliest such studies is from the Carpathians, where the authors analyzed chloroplast aDNA from subfossil Picea abies (Norway spruce) seeds and cone scales, as well as pollen, and found the same genetic haplotypes as in extant populations, indicating strong demographic stasis over millennial timescales (Magyari et al., 2011). The authors could also show a decrease in genetic variability since the beginning of the Holocene, which could be associated with the repeated bottlenecks inferred from paleoecological data. The results were recently confirmed by a follow-up study at the same site based on macrofossils alone, including needles, concluding that such remains are an invaluable repository for information on past population genetic dynamics (Lendvay et al., 2018a). In the Southern Alps, genome-scale aDNA data extracted from subfossil Abies alba (silver fir) needles were used to infer changes in genetic variation between 7.2–5.8 ka cal. yr (calibrated years before present; Schmid et al., 2017), when anthropogenic disturbance led to a drastic decrease in population size (Tinner et al., 1999). The aDNA analysis revealed a lowered observed heterozygosity during the palynologically inferred population decrease, which confirms the paleoecological interpretation of population fragmentation in response to disturbance (Fig. 2). With a recovery of the estimated population size after 6.5 ka cal. yr, genetic variation also returned to predisturbance levels. The lack of genetic differentiation between the populations growing before and after the population decline indicates reexpansion of local trees in the study area (Schmid et al., 2017). Besides macrofossils preserved in lake sediments, plant aDNA has also been extracted from waterlogged, subfossil wood remains found in archaeological or sedimentological contexts (Lendvay et al., 2018b; Wagner et al., 2018). Such remains are widely used in dendroclimatology and chronology to reconstruct past climatic conditions or precisely date wood remains based on tree-ring patterns (Büntgen et al., 2011; Hafner et al., 2021). In the most comprehensive study to date, 167 waterlogged wood remains from European white oaks have been analyzed using high throughput sequencing (HTS) (Wagner et al., 2018). Even though endogenous DNA content was mostly low (< 1% of total DNA reads), the comparison of ancient and extant chloroplast haplotypes indicates a continuous presence of local populations with limited changes in haplotype composition over millennia. Another recent study was able to taxonomically identify 13 000-yr-old pine trunks buried in clay as Pinus sylvestris (Scots pine) using amplicon sequencing of chloroplast aDNA (Lendvay et al., 2018b). Similarly, a study from Lithuania, based on mitochondrial DNA and nuclear microsatellites, could link ancient haplotypes from 11 000-yr-old submerged Scots pine stumps in the Baltic Sea with extant populations and refugia in the Balkan Peninsula (Danusevičius et al., 2021).
DNA preserved in waterlogged sediments proceeds at slower rates than theoretically predicted or estimated for terrestrial environments (Corinaldesi et al., 2008).

An increasing number of studies indicate that waterlogged plant remains represent a rich source of aDNA sequences. Waterlogged seeds (Kistler et al., 2015; Wales et al., 2016; Ramos-Madrigal et al., 2019), fruit fragments (Kistler et al., 2014), needles (Schmid et al., 2017), and wood (Wagner et al., 2018) have all been shown to yield aDNA suitable for chloroplast or nuclear genome-scale analyses (Table 1). As expected for aDNA, the recovered DNA was degraded to small average size (<95 bp) and, when analyzed, characterized by an increased occurrence of purines (adenine and guanosine residues) before strand breaks, putatively due to DNA depurination (Briggs et al., 2007). Moreover, an increased frequency of cytosine-to-thymine misincorporations close to the ends of the DNA fragments was observed, due to deamination of cytosine residues that occur primarily in the single-stranded DNA overhangs (Brotherton et al., 2007; Table 1). Such characteristic damage patterns can in turn also be used to authenticate aDNA (Hofreiter et al., 2001; Jönsson et al., 2013). The most detailed information about plant aDNA preservation is available for waterlogged wood of European white oaks (Wagner et al., 2018). Using wood retrieved from lake sediment, marine silt, clay, and peat, it was found that the DNA fragment size was linearly correlated with thermal age, a measure combining the age of the specimen with average temperatures since deposition (Smith et al., 2003). However, the data indicated that other factors in addition to depurination may contribute to DNA fragmentation (Wagner et al., 2018). Furthermore, it was observed that all millennia-old wood samples with moderate-to-high endogenous oak DNA contents (>1%–16.5% of total DNA reads), with the remaining DNA originating from microbes, were retrieved from wood samples embedded in calcareous lake sediments, suggesting that such sediments could represent particularly promising environments for the preservation of aDNA in wood.

**Challenges of plant ancient DNA analysis**

Despite indications about favorable environmental conditions for the preservation of aDNA in plant macrofossils, it is currently not possible to predict the suitability of samples for genome-scale aDNA analyses. In paleogenomic studies, it is common to initially...
Fig. 2  Past changes in the genetic diversity of Abies alba populations in response to the decline and subsequent recovery of population size around Lago di Origlio, southern Switzerland (Tinner et al., 1999; Schmid et al., 2017). (a) Anthropogenic disturbance caused a drastic decline in A. alba populations during the period 6.5–6.2 ka cal. BP, as reflected in pollen percentages. Individual A. alba needles from selected time periods (orange squares; n, number of needles) were used for ancient DNA analysis. (b) The genetic analysis revealed a significantly lower observed heterozygosity during the period of population decline at 6.5–6.2 ka cal. BP (letters above bars refer to statistically significant differences). However, the absence of significant changes in allelic richness after population recovery at 6.2–5.8 ka cal. BP indicates that genetic diversity was able to recover. (c) Since there was no genetic differentiation between the populations growing before (7.2–6.6 ka cal. BP) and after population recovery (6.2–5.8 ka cal. BP), this process was most likely driven by internal recruitment (H1) and not external recruitment (H2; Schmid et al., 2017).
| Study                | Species       | Tissue   | Age range (ka cal. yr) | No. of sites | Endogenous DNA content (%) | Average DNA fragment length (bp) | Method               | Target DNA damage                  | DNA damage |
|---------------------|---------------|----------|------------------------|--------------|-----------------------------|---------------------------------|-----------------------|-----------------------------------|------------|
| Kistler et al. (2014) | Lagenaria siceraria | Fruit fragment | 10.2–9.8               | 1            | na                          | 63                              | Target enrichment      | Large single-copy region of the plastid genome | na         |
| Wales et al. (2014)  | Vitis vinifera | Seed     | 1.35–1.25              | 1            | na                          | na                              | PCR                   | Plant rbCL marker (138 bp)           | na         |
| Kistler et al. (2015) | Cucurbita spp. | Seed     | Holocene               | 1            | 0.08–0.761                  | na                              | Target enrichment      | Plastid genome                     | Deamination |
| Wales et al. (2016)  | Vitis vinifera | Seed     | 3.2–0.45               | 21           | 1.42                        | 59.1–86.31                      | Target enrichment and shotgun | Plastid genome                     | Deamination |
| Lendvay et al. (2018b) | Pinus sylvestris | Wood     | 13.9–13.0              | 1            | na                          | na                              | PCR                   | Plastid trnL region (84 bp) and trnF region (109 bp) | Deamination |
| Schmid et al. (2017) | Abies alba    | Needle   | 7.2–5.8                | 1            | 0.01–0.33                   | na                              | Target enrichment      | Nuclear exome, complexity reduced   | Deamination |
| Wagner et al. (2018) | Quercus robur/ petraea | Wood     | 9.8–0.55               | 26           | 0–16.5                      | ≤95                             | Shotgun               | Part of plastid genome              | Deamination and depurination |
| Ramos-Madrigal et al. (2019) | Vitis vinifera | Seed     | 2.46–0.75              | 9            | 0–33.5                      | 58.1–77.31                      | Target enrichment and shotgun | Set of nuclear genes               | Deamination |

na, not analyzed.

1Values after target enrichment.

2Value for one sample.

Therefore, pilot studies are often required to identify the best practice for a given set of samples (Kistler et al., 2020). Most plant tissues are rich in polysaccharides and polyphenols, and waterlogged plant remains can also contain humic acids derived from sediment (Kistler et al., 2020). All these molecules tend to coextract with DNA and can act as inhibitors for downstream enzymatic reactions (Kistler, 2012). Several methods have been developed for extracting DNA from modern plant material, aiming to maximize the DNA yield and simultaneously reduce inhibitors. Most of these protocols include either sodium dodecyl sulfate (SDS) or cetyltrimethyl ammonium bromide (CTAB) as detergents in the extraction buffer. The anionic SDS is used to precipitate polysaccharides, whereas the DNA yield and simultaneously reduce inhibitors. Most of these protocols include either sodium dodecyl sulfate (SDS) or cetyltrimethyl ammonium bromide (CTAB) as detergents in the extraction buffer. The anionic SDS is used to precipitate polysaccharides (Doyle & Doyle, 1987). In a comparative study, it was shown that extraction with SDS yields higher DNA amounts from ancient and historical plant remains than extraction with CTAB does (Wales et al., 2014). Indeed, the majority of plant aDNA studies conducted so far have used SDS-based extraction buffers (see Pont et al., 2019, table 2), in some cases (e.g. Gutaker et al., 2017; Schmid et al., 2017) in combination with N-phenoacythiazolium bromide (PTB), an agent that cleaves glucose-derived protein crosslinks and thus can help to release DNA from protein–DNA complexes (Poinar et al., 1998). Extraction of aDNA from herbarium specimens with PTB and SDS was found to decrease the average DNA fragment length when compared with CTAB (Gutaker et al., 2017). Additionally, silica-based DNA purification techniques allow the efficient recovery of short DNA fragments (Rohland et al., 2017).
and by adjusting chaotropic salt concentrations of the binding buffer, fragments as short as 35 bp (Dabney et al., 2013) or even shorter (≥ 25 bp; Glocke & Meyer, 2017) can be retained. After extraction and purification, aDNA molecules must be converted into sequencing libraries, which requires the addition of individual barcodes and platform-specific sequencing adapters to each DNA molecule (Goodwin et al., 2016). Library preparation should also be optimized for degraded DNA. In comparison with double-stranded, single-stranded library preparation techniques minimize the loss of DNA molecules with single-strand breaks on both strands and/or DNA molecules with end modifications located on one of the two strands (Gansauge & Meyer, 2013), thereby increasing the number of library molecules that can be retrieved from highly degraded DNA. Recent studies show fast and inexpensive, single-stranded library preparation methods (Troll et al., 2019) even optimized for aDNA (Tin et al., 2014; Kapp et al., 2021).

Ancient DNA libraries often contain <1% endogenous DNA, with the majority of sequencing capacity taken up by DNA from other sources, such as microorganisms. A way to overcome this limitation is to enrich the libraries using hybridization probes before sequencing (Carpenter et al., 2013). These methods use either commercially synthesized probes (e.g. Ali et al., 2016; Ramos-Madrigal et al., 2019), which can be costly, or benchtop-produced hybridization probes (Suchan et al., 2016; Schmid et al., 2017). Plant genomes are generally complex, containing 10–80% noncoding repeated elements (Metcalfe & Casane, 2013), and can thus be very large, such as in conifers (Nystedt et al., 2013; Mosca et al., 2019). Libraries can be enriched for chloroplast genomes (Meucci et al., 2021; Schulte et al., 2021) or exome (protein-coding) sequences, using probes generated from messenger RNA (Schmid et al., 2017; Toussaint et al., 2021), which may significantly reduce sequencing costs. Last but not least, bioinformatic suites have made it possible to apply the whole set of postsequencing analytical steps in a glance (Schubert et al., 2014; Fellows Yates et al., 2021), including corrections for postmortem damage (Jönsson et al., 2013) and incorporating uncertainty in the genotype calling for low-coverage sequence data (Nielsen et al., 2011). For a whole review on the downstream aDNA bioinformatic analyses, see Orlando et al. (2021).

Applications of ancient DNA analyses based on plant macrofossils

Reconstructing postglacial range shifts using a multisite approach

Ongoing and future climate change is expected to lead to widespread range shifts of plant species that are tracking their current climatic niche (Parmesan & Yohe, 2003; Steinbauer et al., 2018). The unprecedented rate of change is raising the question of whether the dispersal capacity of plants is sufficient to keep up with the rising temperatures. Some scientists have even argued that species might need ‘assisted migration’ to prevent their local extinction (McLachlan et al., 2007; Aitken & Bemmels, 2016; Dauphin et al., 2021). Species migration rates have either been inferred by estimating the species’ dispersal capacity (e.g. by directly or indirectly determining seed dispersal distances) or by tracking the first establishment of a species at different sites in response to past climatic changes using pollen and macrofossil analyses (Pearson, 2006; Feurdean et al., 2013; Birks, 2019). Paleoecological techniques have also been applied to estimate expansion pathways from refugial locations during the last Ice Age, sometimes in combination with ecological niche modeling and phylogenetic data that provide information about past geographic isolation and location of refugia (Gavin et al., 2014). However, these approaches commonly rely on the location of source populations from the main refugia and may ignore secondary or cryptic refugia, which could significantly alter effective species dispersal rates (Birks, 2019). Additionally, paleoecological approaches alone are not able to resolve population-level dynamics due to intrinsic taxonomic constraints that do not allow the identification of within-species lineages. This makes it impossible to track species range shifts in detail. Phylogeographic approaches, on the other hand, rely on present-day genetic variation only, and, therefore, are not able to identify cryptic lineages that became extinct in the past. It is, however, possible to infer past demographic changes and migration patterns from extant populations using demographic inference (Marchi et al., 2021).

Macrofossils deposited in natural archives not only allow determining and dating the local population establishment (e.g. in response to past climate warming), but the genetic information preserved within such remains also provides crucial information about the relationship among populations. By analyzing aDNA from the first populations that established around a network of sites and inferring the degree of relationship among them, the expansion of a population can be tracked with unprecedented detail (Fig. 3a). The analysis of aDNA also allows identifying populations that become extinct during the Holocene and/or might have originated from previously unknown (‘cryptic’) refugia. The identification of refugial populations is important to calculate expansion rates more precisely and to understand the processes involved in species survival under adverse climatic conditions. Indeed, a study identifying sedaDNA of Scots pine and Norway spruce in the lake sediment of an ice-free potential refugium from northern Scandinavia during the Last Glacial Maximum, as well as the Early Holocene presence of a rare mitochondrial haplotype, point to the persistence of trees in northern Scandinavia during the last glacial period (Parducci et al., 2012), even though this interpretation immediately aroused criticism (Birks et al., 2012). Moreover, a more recent study also based on sedaDNA could not fully confirm nor reject the findings, because the low presence of spruce and pine DNA was not distinguishable from background contamination (Allos et al., 2020). Unfortunately, the metabarcoding approach used here erases the signature of deamination patterns potentially present in the original DNA template, which could have represented evidence for ancient DNA.

Tracking changes in genetic diversity through time

Genetic diversity is one of the fundamental components of biodiversity and an important prerequisite for adaptation to
changing environmental conditions. It is therefore crucial for preserving species and maintaining ecosystem resilience. The effects of demographic processes, such as range shifts or population declines, on genetic diversity have been intensely investigated theoretically (Pauls et al., 2013; Dauphin et al., 2021), but empirical studies, especially with long-lived organisms such as trees, are rare and are unable to resolve the impacts of climate change over several generations (Pluess, 2011; Lesser et al., 2013; Elleouet &
Aitken, 2019). Recently, there has been a lot of concern about the effect of population declines on the genetic diversity of many species. However, there is virtually no baseline to compare present vs ancient levels of genetic diversity (but see Leigh et al., 2019; Gauthier et al., 2020).

Plant aDNA studies based on macrofossils would allow the reconstruction of changes in the genetic diversity of a species over extended time periods as required for long-lived organisms such as trees (Fig. 3b). In contrast to herbarium collections, which are also used as an important resource of past genetic diversity (Bieker & Martin, 2018; Lopez et al., 2020), natural archives go beyond the historical period of human-driven impacts on ecosystems, thereby providing information from truly natural populations. In long-lived organisms such as trees, extant populations can also be used to study allele frequency changes over a few generations (Dauphin et al., 2021). By using plant remains deposited in natural archives, such analyses can be extended over much longer time periods (e.g. Schmid et al., 2017; Fig. 3b). Neutral population genetic processes can be tracked by using aDNA, given that allelic frequencies at the population scale are directly impacted by demographic events such as population expansions and declines, gene flow from neighboring populations, or random loss of certain alleles due to genetic drift. A better understanding of the effects of demographic processes on genetic diversity would help us to make more accurate predictions about future changes in genetic diversity.

Testing the adaptive potential of plants to climate change

It is still an open question whether (or to what extent and at what speed) plants can genetically adapt to rapid climatic changes (Birks, 2019). However, this knowledge is crucial in assessing the impact of future climate change on the vegetation. It is clear that species can adapt to local environmental conditions through natural selection, resulting in distinct phenotypes, but the pace at which such processes occur is still debated. The novel research field of landscape genomics aims to identify genes that are associated with certain environmental conditions and result in the expression of respective phenotypes that convey higher fitness (Sork et al., 2013). Adaptive loci are either identified by genome-wide association studies (GWAS) that link adaptive genes to associated phenotypic traits (Bragg et al., 2015) or by environmental association analysis (EAA; also termed genotype–environment associations), which is based on correlations between genetic variants with environmental conditions (Rellstab et al., 2015). In a recent article by Napier et al. (2020), the authors argue that both GWAS and EAA can also be applied to aDNA, thereby testing if plants were able to adapt to past climatic changes. Similarly, a recent study linking genomic information of adult and juvenile Swiss stone pine (Pinus cembra) age cohorts in the Alps with environmental data indicates that environment-driven allele frequency changes over centuries are small, which suggests that such long-lived species may not be able to adapt fast enough, potentially resulting in epigenetically mediated acclimation rather than adaptation, or to local extinction (Dauphin et al., 2021).

Expanding the temporal scale of plant species adaptation to changing environments, one may compare the genetic information preserved within macrofossils from time periods with marked climatic changes (Fig. 3c). For example, the transition from the Younger Dryas cold period to the current Holocene interglacial c. 11 700 yr ago in Europe is considered a close analogue to the current climate warming regarding the rate of climate change, with temperatures rising 2–3°C within less than a century (Heiri et al., 2014). By comparing the allele frequencies of putatively adaptive loci between populations growing before and after the climate warming at the same site and comparing this change with the situation in extant populations, it would be possible to better estimate the adaptive potential of a species.

Conclusions and outlook

Climate change will have profound impacts on plant distribution and abundance, as well as associated ecosystem services and functioning. Analyzing plant aDNA from macrofossils deposited in natural archives has the potential to assess the effects of past rapid climate change on plant species at the genetic level. This will ultimately allow better predictions about the effects of future climate change on the abundance, distribution, adaptive potential, and genetic diversity of plants.

With ever more ancient genetic information available, it will also be possible to test and validate population genetic models. Such paleo-validated models can then be used in turn to make detailed predictions about future changes in genetic diversity. This approach of comparing model output with paleo-data is standard procedure for climate and vegetation models, but is not very common for population genetic models, at least for long-lived organisms over long time scales. Paleogenetic information from individual species (aDNA) could also be combined with data from multixtaxon approaches based on sedaDNA (Dussex et al., 2021).

Overall, we believe that the proposed framework has the potential to fundamentally improve our understanding of population genetic processes, by opening a window into the past and allowing us to retrospectively track genetic changes over time.

Acknowledgements

This project was supported by the SwissForestLab (research grant SFL-19 P4) and by the Federal Office for the Environment FOEN, as well as the Swiss National Science Foundation within the project HOLOGENE (grant no. SNF- 200021_188472). We would like to thank three anonymous reviewers for valuable comments on an earlier version of this article.

Author contributions

C Schwörer, NA, FG and C Sperisen designed the research; C Schwörer and ML wrote a first draft and all authors contributed to the final manuscript.

ORCID

Nadir Alvarez https://orcid.org/0000-0002-0729-166X
Felix Gugerli https://orcid.org/0000-0003-3878-1845
Data availability

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

References

Aitken SN, Emmels JB. 2016. Time to get moving: assisted gene flow of forest trees. Evolutionary Applications 9: 271–290.

Ali OA, O’Rourke SM, Amish SJ, Meek MH, Luikart G, Jeffres C, Miller MR. 2007. Patterns of damage in genomic base modification events as the sole cause of post mortem miscoding lesions. Nucleic Acids Research 35: 5717–5728.

Büntgen U, Tegel W, Nicolussi K, McCormick M, Frank D, Trouet V, Kaplan JO, Hergt F, Heussner K-U, Wanner H et al. 2011. 2500 Years of European climate variability and human susceptibility. Science 331: 578–582.

Carpenter M, Buenrostro J, Valdivieso C, Schroeder H, Allenfost M, Sikora M, Rasmussen M, Gravel S, Guillon S, Nekhrizov G et al. 2013. Pulling out the 1% whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. The American Journal of Human Genetics 93: 852–864.

Corinaldesi C, Beolchini F, Dell’anno A. 2008. Damage and degradation rates of extracellular DNA in marine sediments: implications for the preservation of gene sequences. Molecular Ecology 17: 3939–3951.

Dahney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, Nickl B, Valdivieso C, García N, Pábo S, Arruaga J-L et al. 2013. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National Academy of Sciences, USA 110: 15758–15763.

Danusevicius D, Buchovska J, Zulkuš V, Daugnora L, Gribnikas A. 2021. DNA markers reveal genetic associations among 11,000-year-old Scots pine (Pinus sylvestris L.) found in the Baltic Sea with the present-day gene pools in Lithuania. Forests 12: 317.

Dauphin B, Rollast C, Schmid M, Zoller S, Karger DN, Broodbek S, Guillaume F, Gugerli F. 2021. Genomic variability to rapid climate warming in a tree species with a long generation time. Global Change Biology 27: 1181–1195.

Dehaes M, Ávila-Arcos MC, Diez-del-Molino D, Fumagalli M, Guschanski K, Lorenzen ED, Malaspina A-S, Marques-Bonet T, Martin MD, Murray GGR et al. 2020. Inference of natural selection from ancient DNA. Evolution Letters 4: 94–108.

Dehaes M, Pecenerová P, Muller H, Tikhonov A, Nikolíková P, Tsiganovka VI, Danilov GK, Diez-del-Molino D, Vartanyan S, Dälen L et al. 2021. Combining Bayesian age models and genetics to investigate population dynamics and extinction of the last mammoths in northern Siberia. Quaternary Science Reviews 259: e106913.

Dellaporta SL, Wood J, Hicks JB. 1983. A plant DNA minipreparation: version II. Plant Molecular Biology Reporter 1: 19–21.

Der Sarkissian C, Balanovsky O, Brandt G, Khartanovich V, Buzhilova A, Koshel E, Der Sarkissian C, Balanovsky O, Brandt G, Khartanovich V, Buzhilova A, Koshel E, Dabney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, Nickl B, Valdivieso C et al. 2015. Ancient DNA reveals prehistoric gene-flow from Siberia in the complex human population history of north east Europe. PLoS Genetics 9: e1003296.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19: 11–15.

Dusse N, Bergfeldt N, de Anca Prado V, Dehaes M, Diez-del-Molino D, Eermark E, Kanellidou F, Larsson P, Leme C, Nedsingh B, Pegg D, Calafat A et al. 2021. Integrating multi-taxa palaeogenomes and sedimentary ancient DNA to study past ecosystem dynamics. Proceedings of the Royal Society B: Biological Sciences 288: e20211252.

Eloueit JS, Aitken SN. 2019. Long-distance pollen dispersal during recent colonization favors a rapid but partial recovery of genetic diversity in Picea sitchensis. New Phytologist 222: 1088–1100.

Estrada O, Breen J, Richards SM, Cooper A. 2018. Ancient plant DNA in the genomic era. Nature Plants 4: 394–396.

Fellows Yates JA, Lamnidiis TC, Borry M, Valteuña AA, Fagerini Z, Clayton S, GM, Neuquam J, Peltzer A. 2021. Reproducible, portable, and efficient ancient genome reconstruction with NF-CORE/EAGER. PeerJ 9: e10947.

Feurdean A, Bhagwat SA, Willis KJ, Birks HJB, Lischke H, Hickler T. 2013. Combining Bayesian age models and genetics to investigate population dynamics and extinction of the last mammoths in northern Siberia. Quaternary Science Reviews 259: e106913.

Gansauge M-T, Meyer M. 2013. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nature Protocols 8: 737–748.

Gauthier J, Pajkovic M, Neuenwander S, Kaila L, Schmid S, Orlando L, Alvarez N. 2020. Museomics identifies genetic erosion in two butterfly species across the 20th century in Finland. Molecular Ecology Resources 20: 1191–1205.
Napier JD, de Lafontaine G, Chipman ML. 2020. The evolution of paleoecology. *Trends in Ecology & Evolution* 35: 295–297.

Nielsen R, Paul JS, Albrechtsen A, Song YS. 2011. Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics* 12: 443–451.

Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin Y-C, Scofield DG, Zevi F, Delhomme N, Giacomello S, Alexeyenko A et al. 2013. The Norway spruce genome sequence and conifer genome evolution. *Nature* 497: 579–584.

Orlando L, Allaby R, Skoglund P, Der Sarkissian C, Stockhammer PW, Avila-Arcos MC, Fu Q, Krause J, Willerslev E, Stone AC et al. 2021. Ancient DNA analysis. *Nature Reviews Methods Primers* 1: e14.

Parducci L, Bennett KD, Ficetola GF, Alos SG, Suyama Y, Wood JR, Pedersen MW. 2017. Ancient plant DNA in lake sediments. *New Phytologist* 214: 924–942.

Parducci L, Jorgensen T, Tollefsrud MM, Elverland E, Alm T, Fontana SL, Bennett KD, Haile J, Matetovici I, Suyama Y et al. 2012. Glacial survival of boreal trees in northern Scandinavia. *Science* 335: 1083–1086.

Parducci L, Petri RJ. 2004. Ancient DNA – unlocking plants’ fossils secrets. *New Phytologist* 161: 335–339.

Parducci L, Suyama Y, Lascoux M, Bennett KD. 2005. Ancient DNA from pollen: a record of population history in Scots pine. *Molecular Ecology* 14: 2875–2882.

Parmesan C, Yohe G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.

Pauls SU, Nowak C, Balint M, Pfenninger M. 2013. The impact of global climate change on genetic diversity within populations and species. *Molecular Ecology* 22: 925–946.

Pearson RG. 2006. Climate change and the migration capacity of species. *Trends in Ecology & Evolution* 21: 111–113.

Pedersen MW, Ruter A, Schweger C, Friebe H, Staff RA, Kjeldsen KK, Mendoza Ramos-Madrigal J, Runge AKW, Bouby L, Lacombe T, Samaniego Castruita JA, Pont C, Wagner S, Lagane F, Seguin-Orlando A, Schubert M, Leroy T, Guichoux E, Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brochmann C, Willerslev E. 2007. Power and limitations of the chloroplast trnl (UAA) intron for plant DNA barcoding. *Nucleic Acids Research* 35: e14.

Tinner W, Hubschmid P, Wehrli M, Ammann B, Conedera M. 1999. Optimization of DNA recovery and amplification from non-carbonized ancient DNA. *Genome Research* 20: 2873–2882.

Schulte L, Bernhardt N, Stoofo-Leichsenring K, Zimmermann HH, Pestryakova LA, Epp LS, Herzschuh U. 2021. Hybridization capture of larch (*Larix Mill.*) chloroplast genomes from sedimentary ancient DNA reveals past changes of Siberian forest. *Molecular Ecology Resources* 21: 801–815.

Scott MF, Botigué LR, Brace S, Stevens CJ, Mullin VE, Stevenson A, Thomas MG, Fuller DQ, Mott R. 2019. A 3,000-year-old Egyptian emmer wheat genome reveals dispersal and domestication history. *Nature Plants* 5: 1120–1128.

Smith CJ, Chamberlain AT, Riley MS, Stringer C, Collins MJ. 2003. The thermal history of human fossils and the likelihood of successful DNA amplification. *Journal of Human Evolution* 45: 203–217.

Sork VL, Atkin SN, Dyer RJ, Eckert AJ, Legendre P, Neale DB. 2013. Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes* 9: 901–911.

Steinbauer MJ, Grytnes J-A, Juriesen G, Külonen A, Leninor J, Paoli H, Rixen C, Winkler M, Bardy-Durchhalter M, Barni E et al. 2018. Accelerated increase in plant species richness on mountain summits is linked to warming. *Nature* 556: 231–234.

Suchan T, Pitteloud C, Gerasimova NS, Kostikova A, Schmid S, Arrigo N, Pajkovic M, Ronikier M, Alvarez N. 2016. Hybridization capture using RAD probes (hyRAD), a new tool for performing genomic analyses on collection specimens. *PLoS ONE* 11: e0151651.

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brochmann C, Willerslev E. 2007. Power and limitations of the chloroplast trnl (UAA) intron for plant DNA barcoding. *Nucleic Acids Research* 35: e14.

Tin MM-Y, Economou EP, Mikheyev AS. 2014. Sequencing degraded DNA from non-destructively sampled museum specimens for RAD-tagging and low-coverage shotgun phylogenomics. *PLoS ONE* 9: e96793.

Tinner W, Hubschmid P, Wehrli M, Ammann B, Conedera M. 1999. Long-term forest fire ecology and dynamics in southern Switzerland. *Journal of Ecology* 87: 273–289.

Toussaint EFA, Gauthier J, Bilat J, Gillett CPDT, Gough HM, Lundkvist H, Blan C, Muñoz-Ramírez CP, Alvarez N. 2021. HyRAD-X exome capture museomics unravels giant bean beetle evolution. *Genome Biology and Evolution* 13: eva112.

Troll CJ, Kapp J, Rao V, Harkins KM, Cole C, Naughton C, Morgan JM, Shapiro B, Green RE. 2019. A ligation-based single-stranded library preparation method to analyze cell-free DNA and synthetic oligos. *BMC Genomics* 20: e1023.

van der Valk T, Pečenović P, Diez-del-Molino D, Bergström A, Oppenheimer J, Hartmann S, Xenikoudakis G, Thomas JA, Dehasque M, Saglcan E et al. 2021. Million-year-old DNA sheds light on the genomic history of mammoths. *Nature* 591: 265–269.

Wagner S, Lagane F, Seguin-Orlando A, Schubert M, Leroy T, Guichoux E, Chancreel E, Bech-Hebelstrup I, Bernard V, Billard C et al. 2018. High-throughput DNA sequencing of ancient wood. *Molecular Ecology* 27: 1138–1154.

Wales N, Andersen K, Cappellini E, Avila-Arcos MC, Gilbert MTP. 2014. Optimization of DNA recovery and amplification from non-carbonized archaeobotanical remains. *PLoS ONE* 9: e86827.

Wales N, Ramos-Madrigal J, Cappellini E, Carmona Baeza A, Samaniego Castruita JA, Romero-Narváro JA, Carne C, Avila-Arcos MC, Penaloya F, Moreno-Mayar JV et al. 2016. The limits and potential of paleogenomic techniques for reconstructing grapevine domestication. *Journal of Archaeological Science* 72: 57–70.

Wales N, Romero-Narváro JA, Cappellini E, Gilbert MTP. 2012. Choosing the best plant for the job: a cost-effective assay to prescreen ancient plant remains destined for shotgun sequencing. *PLoS ONE* 7: e45644.

Wang Y, Pedersen MW, Ålós IG, De Sanctis B, Racimo F, Prohaska A, Coissac E, Owens HL, Merkel MKF, Fernandez-Guerra A et al. 2021. Late Quaternary dynamics of Arctic biota from ancient environmental genomics. *Nature* 590: 86–92.

Key words: genetic diversity, Holocene, lake sediment, paleoecology, paleogenomics, range shifts.

Received, 9 December 2021; accepted, 7 March 2022.