Conservation Biology

Contributed Paper

Use of Ancient Sedimentary DNA as a Novel Conservation Tool for High-Altitude Tropical Biodiversity

SANNE BOESSENKOOL,* ** GAYLE MCGLYNN,† LAURA S. EPP,* †† DAVID TAYLOR,‡ MANUEL PIMENTEL,* ‡‡ ABEL GIZAW,* §§ SILESHI NEMOMISSA,§ CHRISTIAN BROCHMANN,* AND MAGNUS POPP*

*National Centre for Biosystematics, Natural History Museum, University of Oslo, P.O. Box 1172, Blindern NO-0318, Oslo, Norway email sanneboessenkool@gmail.com

†School of Natural Sciences, Trinity College Dublin, Dublin 2, Ireland

‡Department of Geography, National University of Singapore, 1 Arts Link, Kent Ridge 117570, Singapore

§Department of Plant Biology & Biodiversity Management, College of Natural Sciences, Addis Ababa University, P.O. Box 3434, Addis Ababa, Ethiopia

Abstract: *Conservation of biodiversity may in the future increasingly depend upon the availability of scientific information to set suitable restoration targets. In traditional paleoecology, sediment-based pollen provides a means to define preanthropogenic impact conditions, but problems in establishing the exact provenance and ecologically meaningful levels of taxonomic resolution of the evidence are limiting. We explored the extent to which the use of sedimentary ancient DNA (sedaDNA) may complement pollen data in reconstructing past alpine environments in the tropics. We constructed a record of afro-alpine plants retrieved from DNA preserved in sediment cores from 2 volcanic crater sites in the Albertine Rift, eastern Africa. The record extended well beyond the onset of substantial anthropogenic effects on tropical mountains. To ensure high-quality taxonomic inference from the sedaDNA sequences, we built an extensive DNA reference library covering the majority of the afro-alpine flora, by sequencing DNA from taxonomically verified specimens. Comparisons with pollen records from the same sediment cores showed that plant diversity recovered with sedaDNA improved vegetation reconstructions based on pollen records by revealing both additional taxa and providing increased taxonomic resolution. Furthermore, combining the 2 measures assisted in distinguishing vegetation change at different geographic scales; sedaDNA almost exclusively reflects local vegetation, whereas pollen can potentially originate from a wide area that in highlands in particular can span several ecozones. Our results suggest that sedaDNA may provide information on restoration targets and the nature and magnitude of human-induced environmental changes, including in high conservation priority, biodiversity hotspots, where understanding of preanthropogenic impact (or reference) conditions is highly limited.*

Keywords: Albertine Rift, high-elevation biodiversity, lake sediment, metabarcoding, sedaDNA, vascular plants, Virunga volcanoes

**Current address: Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, P.O. Box 1066, Blindern, NO-0318, Oslo, Norway

††Current address: Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Research Unit Potsdam, Telegrafenberg A 43, 14473 Potsdam, Germany

‡‡Current address: Grupo de Investigación en Bioloxía Evolutiva (GIBE), Facultade de Ciencias, Universidade da Coruña, Campus da Zapateira sn, 15071 A Coruña, Galicia, Spain

§§Current address: Department of Plant Biology & Biodiversity Management, College of Natural Sciences, Addis Ababa University, P.O. Box 3434, Addis Ababa, Ethiopia

Paper submitted March 14, 2013; revised manuscript accepted July 17, 2013.

DOI: 10.1111/cobi.12195

Uso de ADN Sedimentario Antiguo como una Herramienta Novedosa de Conservación para la Biodiversidad Tropical de Grandes Altitudes

Resumen: *La conservación de la biodiversidad en el futuro puede depender cada vez más de la disponibilidad de la información científica para establecer objetivos de restauración adecuados. En la paleoecología tradicional el polen hallado en sedimentos proporciona un medio para definir las condiciones de impactos preantropogénicas, pero hay problemas en el momento de establecer su proveniencia exacta y los niveles ecológicamente significativos de su resolución taxonómica. Exploramos el grado en el que el uso de ADN sedimentario antiguo (ADN seda) puede complementar la información del polen en la reconstrucción de ambientes alpinos del pasado en los trópicos. Construimos un registro de plantas afro-alpinas tomadas de ADN preservado en núcleos sedimentarios de dos sitios de cráteres volcánicos en la Falla Albertina, al oriente de África. El registro se extendió más allá del comienzo de los efectos antropogénicos sustanciales en las montañas tropicales. Para asegurar la inferencia taxonómica de alta calidad de las secuencias de ADN seda, construimos una biblioteca extensa de referencias de ADN cubriendo la mayoría de la flora afro-alpina, secuenciado el ADN de especímenes taxonómicamente verificados. Las comparaciones con los registros de polen de los mismos núcleos sedimentarios mostraron que la diversidad de plantas recuperada con el ADN seda mejoró las reconstrucciones de vegetación con base en los registros de polen al revelar tanto taxones adicionales como proporcionando una resolución taxonómica incrementada. Además, el combinar las dos medidas ayudó a distinguir cambios en la vegetación en diferentes escalas geográficas; el ADN seda casi siempre refleja exclusivamente la vegetación local, mientras que el polen puede originarse potencialmente a partir de un área amplia que en las tierras altas en particular puede abarcar varias ecozonas. Nuestros resultados sugieren que el ADN seda puede proporcionar información sobre los objetivos de restauración y la naturaleza y magnitud de los cambios ambientales inducidos por el ser humano, incluyendo en prioridad de alta conservación, hotspots de biodiversidad, donde entender las condiciones del impacto preantropogénico (o su referencia) es difícil.*

Palabras Clave: ADN seda, biodiversidad de grandes altitudes, Falla Albertina, metabarcoding, plantas vasculares, sedimento lacustre, volcanes Virunga

Introduction

Global climate change is widely seen as a threat to the world's montane regions (Franzen & Molander 2012). Nowhere is this more apparent than in the tropics (La Sorte & Jetz 2010), where mountains provide a range of valuable ecosystem services, including exceptionally high levels of biodiversity (Dimitrov et al. 2012). Biodiversity conservation in these tropical highland ecosystems faces major challenges, not least of which is limited availability of scientific information (Buytaert et al. 2011). A paucity of scientific information limits understanding of natural variability and the nature, timing, and degree of human impacts, including those arising from recent climate changes (Colwell et al. 2008). Ecosystem restoration is increasingly regarded as an effective means of improving ecosystem services that have been lost or reduced as a result of anthropogenically induced effects (Bullock et al. 2011; Higgs & Roush 2011), but limited scientific information can also place severe restrictions on restoration efforts (Montoya et al. 2012). In particular, the definition of suitable restoration targets is made problematic by uncertainty over past conditions and the extent of natural variability (MacDougall et al. 2004). Paleoecology, generally the use of sediment-based information to reconstruct past communities of plants and animals and associated environmental conditions, can be of particular relevance here as a potential source of evidence of preanthropogenic impact or reference conditions (Willis & Birks 2006).

A weakness that characterizes much sediment-based evidence is uncertainty over its provenance. For pollen fixing the geographic location of parent plant(s) can be problematic, owing to the possibility of long-distance dispersal of many taxa (Seppä & Bennett 2003). Additionally, taxonomic resolution of pollen and spores is typically low. Finer taxonomic resolution is possible with some macrofossils, which are also local in origin (Birks & Birks 2000). Sediment-based macrofossils have recently been used to establish the invasive histories of a selection of plants on the Galapagos Islands (Coffey et al. 2011)—but macrofossils can suffer from problems of preservation and quantification, as well as identification.

Ancient DNA from sediment (sedaDNA) is a relatively new tool that may be used to address both problems of poor taxonomic resolution and geographic origin of the material. Identification with fine taxonomic resolution can be achieved for many plant groups (Taberlet et al. 2007; Sønstebø et al. 2010; Jørgensen et al. 2012*a*), and at least in permafrost, its provenance is suggested to be primarily local (Jørgensen et al. 2012*a*). Due to the temperature-related degradation of DNA (Lindahl 1993; Smith et al. 2001; Willerslev et al. 2004), paleoecological studies employing plant sedaDNA have up to now largely focused on relatively species-poor, high-latitude areas (e.g., Willerslev et al. 2003; Anderson-Carpenter

et al. 2011; Jørgensen et al. 2012*a*, 2012*b*; Parducci et al. 2012), although recent research presents the results of retrieval of mammal and plant DNA from herbivore middens in hot and arid locations in Australia and South Africa (Murray et al. 2012). Relatively little is known about the provenance and preservation of plant sedaDNA or the different sources of bias affecting the diversity that can be retrieved (Jørgensen et al. 2012*a*; Parducci et al. 2013). Comparisons of pollen and plant sedaDNA are rare, having only recently been carried out in high-latitude permafrost and lake sediments (Jørgensen et al. 2012*a*; Parducci et al. 2013) and in herbivore middens (Murray et al. 2012). The method has yet to be tested in other environments and in particular in the biodiversity-rich, humid tropics.

In very cold conditions, sedaDNA can potentially be preserved up to hundreds of thousands of years (e.g., Willerslev et al. 2003; Willerslev et al. 2007). From a conservation point of view, however, such long timescales may not always be necessary to provide baseline data of pristine states and detect biodiversity changes. This is because in many instances, the onset of the Anthropocene, some 200 or so years ago (Crutzen 2002), effectively marks the beginning of significant anthropogenic impact conditions. For planktonic organisms, successful retrieval of DNA from sediment cores has been carried out in the tropics on such timescales (Epp et al. 2010; Stoof-Leichsenring et al. 2012). In areas of high biodiversity in both temperate and tropical regions, sedaDNA thus also has the potential to provide a valuable tool for characterizing past changes in the composition of terrestrial biota.

Here, we present a record of plants retrieved from DNA preserved in sediments from 2 crater sites in the Virunga volcanoes of the Albertine Rift, eastern Africa (Fig. 1). The Albertine Rift is one of the world's biodiversity hotspots (Myers et al. 2000) most threatened by human pressures (Hartter et al. 2012). The area is therefore an important focus for conservation (Küper et al. 2004; Plumptre et al. 2007*a*). The Virunga volcanoes, with their marked altitudinal zonation in vegetation similar to other highlands in tropical Africa (Lind & Morrison 1974), straddle the borders between the Democratic Republic of the Congo, Rwanda, and Uganda (Fig. 1). Biodiversity management is centered upon 3 adjoining national parks, each located in a different country (Plumptre et al. 2007*b*). The area contains over half of the world's remaining mountain gorillas (*Gorilla beringei beringei*) and for several years has been associated with a high level of intervention-based approaches to conservation (Robbins et al. 2011).

The 2 sites upon which the current study is based, one a peat-forming swamp and the other a lake, are particularly interesting for sedaDNA because both are within the craters at the summit of neighboring volcanoes and therefore have extremely limited catchment areas, which enhances the likelihood that sedaDNA has a local source. To ensure high-quality taxonomic inference from the sedaDNA sequences obtained from the crater sites in the Virunga volcanoes, we constructed an extensive DNA reference library covering the majority of the afro-alpine flora by sequencing DNA in taxonomically verified specimens. We explored the extent to which the relatively novel method of sedaDNA may complement a more traditional form of paleoecological (pollen) data and can potentially contribute both to the reconstruction of past tropical environments and to conservation efforts in high biodiversity settings.

Methods

Coring of the crater swamp on Mt. Gahinga (1° 23.2′ S, 29° 38.7′ E, 3474 m.a.s.l.) and the crater lake on Mt. Muhavura (1° 23′ S, 29° 40.7′ E, 4127 m.a.s.l.) is described in McGlynn et al. (2013). We took extreme care to avoid contamination of the sediments during the sampling process. Cores were immediately double wrapped and sealed for transport to the lab following retrieval on site. Slicing and subsampling were performed using thoroughly cleaned equipment, and sediment samples for DNA analyses (Table 1) were collected from the interior part of the cores to minimize contamination by exogenous plant DNA. Samples were transferred to sterile falcon tubes and stored at 4 °C until processing in the laboratory. Development of age models for the sediment cores from each of the sites is described in McGlynn et al. (2013). The 9.74 m core extracted from the Mt. Gahinga crater swamp extended back to a date of circa 7800 calibrated years BP, whereas the 2.47 m core from the Mt. Muhavura crater lake extended back to circa 2800 calibrated years BP. Pollen preparation and enumeration followed standard procedures, and pollen zonations referred to in Figure 2 were defined numerically using a constrained incremental sum of squares (CONISS) cluster analysis based on total pollen percentage data (excluding unidentified and damaged grains; McGlynn et al. 2013).

The DNA extractions, polymerase chain reactions (PCRs) with the *trn*L *g* and *h* primers to obtain the *trn*L P6 loop (Taberlet et al. 2007), and sequencing with the Roche 454 platform were carried out as outlined in Supporting Information. Extractions and PCR setup were performed in facilities dedicated to ancient DNA analyses at the Centre for Geogenetics (University of Copenhagen) and the Natural History Museum (University of Oslo). We did 2 rounds of PCRs and sequencing (resulting in dataset 1 and dataset 2; Supporting Information). We did not consider the 2 obtained datasets replicates because different PCR and sequencing protocols were used for each. Instead, we treated them as complementary and merged the results to obtain the final taxon list.

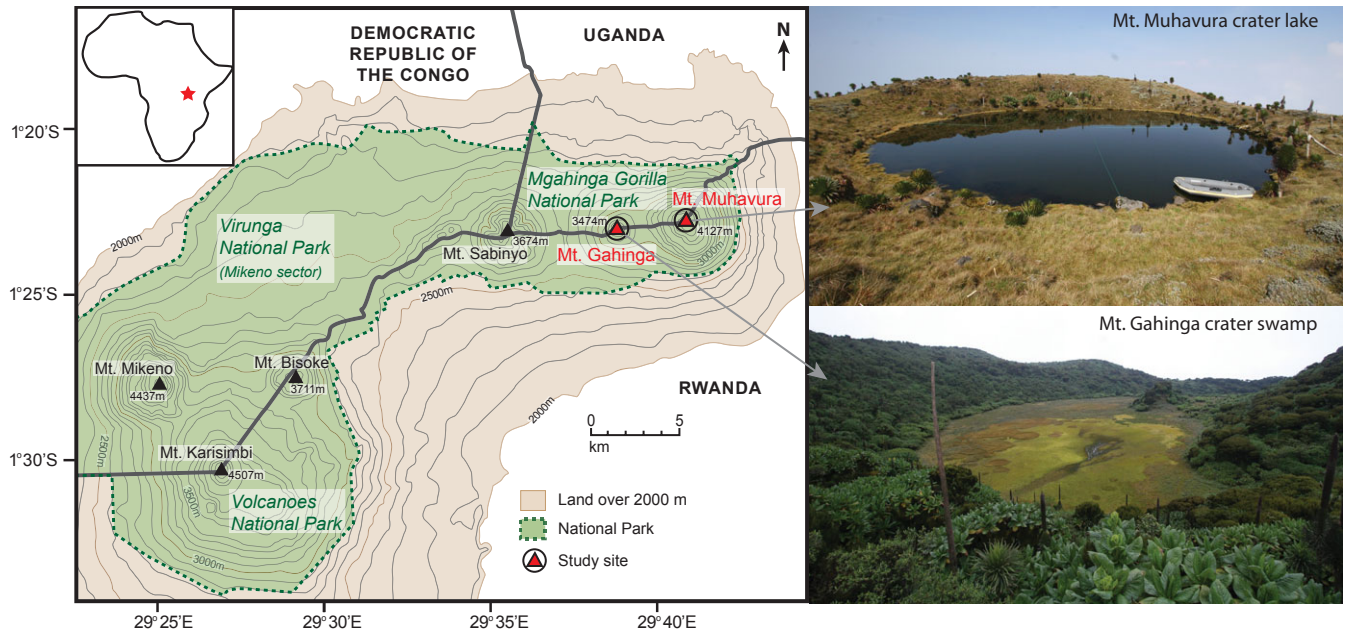

Figure 1. Location of the 2 study sites in the Virunga volcanoes of the Albertine Rift, eastern Africa, and photographs of the lake (Mt. Muhavura) and the swamp (Mt. Gahinga) sampled at the sites.

To minimize contamination, standard ancient DNA precautions were adhered to throughout the experiments (Willerslev et al. 2004; Willerslev & Cooper 2005). Nevertheless, laboratory reagents are often contaminated by DNA from cultivated plants, which can be easily amplified when using universal primers. For dataset 2, one of the PCR controls showed positive amplification, and all extraction and PCR controls were subsequently included in the sequencing. Post filtering (see below), only one of the PCR controls yielded a sequence for which taxonomic identity could be inferred (*Pinus resinosa*). This sequence was also retrieved in one of the samples, from which it was subsequently removed. Because clean control reactions do not necessarily mean that contamination is absent (Willerslev & Cooper 2005), we took the additional precautionary approach of removing all sequences (with taxonomic identity >0.95, see below) that may have resulted from contaminant DNA (including cultivated taxa or taxa not present in eastern Africa). This concerned the following taxa: Amygdaloideae, Anacardiaceae, *Hevea brasiliensis*, Juglandaceae, *Ligustrum* sp., Malvaceae, *Medicago polymorpha*, Micrandreae, Musaceae, Nolinoideae, Noteleae, Rutaceae, *Sicyos andreanus*, and Triticeae (Supporting Information).

We used the computing facilities provided by the Norwegian Metacenter for Computational Science (Notur) to filter and sort sequences with programs from the OBITools package (http://www.grenoble.prabi.fr/trac/OBITools). Filtering was performed as described in Boessenkool et al. (2012). Using the program Obiclean (part of the OBITools package; Bellemain et al. 2013), we conducted an additional cleaning step. Within each sample, only sequences represented by >10 reads were kept in the final dataset. For taxonomic assignment, we used the program ecoTag (see Boessenkool et al. 2012) and analyzed the sequences against 2 reference databases: a local library covering the afro-alpine flora and a reference database based on the embl standard sequences release 113. The reference database of the embl sequences was created by *in silico* PCR (with the program ecoPCR; Ficetola et al. 2010) on the embl standard sequences (release 113) with the *trn*L *g* and *trn*L *h* primers (5 mismatches allowed between primer and target sequence). Only sequences identified to at least family level with an identity over 0.95 were kept in the dataset.

Sampling for the afro-alpine taxonomic reference library was carried out during extensive fieldwork in 2007–2009 in Ethiopia, Kenya, Tanzania, and Uganda. Whenever possible, 5 individuals of all species found in 100 × 100 m plots in 5 different habitat types (grasslands, bogs, rocky outcrops, *Alchemilla* communities, and *Dendrosenecio* forests) were collected. Leaf samples from all 5 individuals from each species found in a plot were stored in silica for DNA preservation, and 3 individuals were pressed. Additional sampling was carried out outside plots to collect as many species as possible. The pressed material was deposited in the herbaria at the Natural History Museum (O), University of Oslo, Norway, and the National Herbarium of Ethiopia (ETH), Addis Ababa University, Ethiopia. The third collection was deposited in the country of collection (East African Herbarium [EA],

Table 1. Depth and uncalibrated ages of the sediment samples analyzed for sedaDNA, collected at the 2 study sites in the Virunga volcanoes of the Albertine Rift, eastern Africa.

Sample	*Average depth (cm)*	*Age*[a]
MUH4_10–11 cm	10.5	1980 AD
MUH4_18–19 cm	18.5	1960 AD
MUH4_26–27 cm	26.5	1940 AD
MUH2–1_75–79 cm	77	1190 AD
MUH2–1_130–134 cm	132	300 AD
MUH2–2_160–163 cm	161.5	510 AD
MUH2–3_220–223 cm	221.5	170 BC
MUH2–4_245–248 cm	246.5	350 BC
GAH1_11–15 cm	63	1920 AD
GAH1_35–38 cm	86.5	1810 AD
GAH2_110–113 cm	111.5	1700 AD
GAH2_158–161 cm	159.5	1480 AD
GAH3_210–212 cm	211	1250 AD
GAH3_280–282 cm	281	950 AD
GAH7_640–642 cm	641	2790 BC

[a]*Ages calculated according to the age models described in McGlynn et al. (2013) based on 9 radiocarbon dates from the Mt. Gahinga (GAH) sediment sequence and 7 radiocarbon dates from the Mt. Muhavura (MUH) sediment sequence.*

Kenya; Sokoine University of Agriculture [SUA], Tanzania; Makerere University Herbarium [MHU], Uganda). In addition to careful identification work carried out in O and ETH, taxonomic experts were consulted to identify several taxonomically difficult groups. For most species, DNA extraction was performed on 2 specimens from different parts of the species distribution area. Details on DNA extractions, PCRs, sequencing, and formatting of the final database are described in the Supporting Information. All sequences were compared to published sequences with NCBI/BLAST, and phylogenetic analyses included sequences from closely related taxa to verify the taxonomic identity.

We compared the taxon list obtained from the sedaDNA analysis with the pollen taxa from McGlynn et al. (2013) that attained an abundance of >1% in at least 1 sample. We evaluated both the taxa retrieved within each of the pollen zones recognized and the overall taxonomic resolution reached by the 2 methods. Furthermore, we contrasted the 2 sites (Mt. Gahinga and Mt. Muhavura), comparing the taxa retrieved for each of the 2 methods in each of the sites for samples dating to within the last 1000 years (comprising the most recent pollen zones).

Results

For the afro-alpine taxonomic reference library, sequences covering the complete P6 loop were obtained from 564 specimens representing 42 families, 131 genera, and 271 species (31 specimens were identified to genus only; see Supporting Information for a complete taxon list). With this library, 100% of the taxa can be identified to family, 64.1% to genus, and 24.4% to species.

DNA sequences were obtained from all but one of the sediment samples (GAH2_110–113cm). A total of 51,157 and 14,5511 raw reads could be assigned to samples for datasets 1 and 2, respectively. After cleaning with obiclean, 29,179 reads (129 unique) for dataset 1 and 70,799 reads (389 unique) for dataset 2 remained. The number of reads remaining after each filtering step is provided in Supporting Information. Sequences with identifications <1.00 (but over 0.95) were checked manually against the closest matching reference sequences. Mismatches were in most cases due to variation in homopolymers, a well-known sequencing error in 454 sequencing (Margulies et al. 2005; Huse et al. 2007). We consulted family-based regional flora (the Flora of Tropical East Africa) and unpublished local species checklists to confirm occurrence of taxa in our study area. For genera for which not all the relevant species were included in the library, we adjusted any identification to species level up to genus level. Similarly, we were able to narrow down some identifications based on knowledge of the local flora.

From the DNA analyses, 14 and 22 taxa were identified from Mt. Gahinga and Mt. Muhavura, respectively (Fig. 2 & Supporting Information). The level of identification varied from species to family, and individual taxa were found in 1–6 of the samples. At both sites, the number of detected taxa declined with sample age—a trend most pronounced at Mt. Gahinga, where 9 taxa were recovered from the youngest sample compared with just 1 or 2 taxa from the older samples. In addition to vascular plants, for which the *trn*L *g* and *h* primers are designed, we also recovered several bryophyte taxa (e.g., *Andreaea nitida*, *Breutelia* sp., *Sphagnum* sp.) and 1 algae (*Closterium* sp.).

Strong differences could be observed between the taxonomic diversity inferred from the DNA and from the pollen data (Fig. 2). Overall, more taxa were recovered from pollen and these include many montane forest species, such as *Podocarpus* sp., *Celtis* sp., and *Olea* sp. However, the taxa inferred from the DNA are not a subset of those recovered from pollen: several taxa were detected on the basis of the DNA analyses only, for example, *Callitriche* sp., *Veronica* sp., Gnaphalieae, Iridaceae, and *Ranunculus* sp. (Table 2). Furthermore, for some families (particularly Cyperaceae, Ericaceae, and Poaceae), the DNA results reached a higher taxonomic resolution than was possible for the pollen (Table 2).

The taxa inferred from DNA reflected mostly local vegetation growing proximate to the lake or the swamp (personal observation G.M., M.P.). Given the altitude of the sites, only 3 taxa (Fabaceae, *Morella* sp., and *Podocarpus* sp.) were unlikely to be from local-growing parent plants. By contrast, the potential source area for much of the pollen was large, as indicated by the contributions to

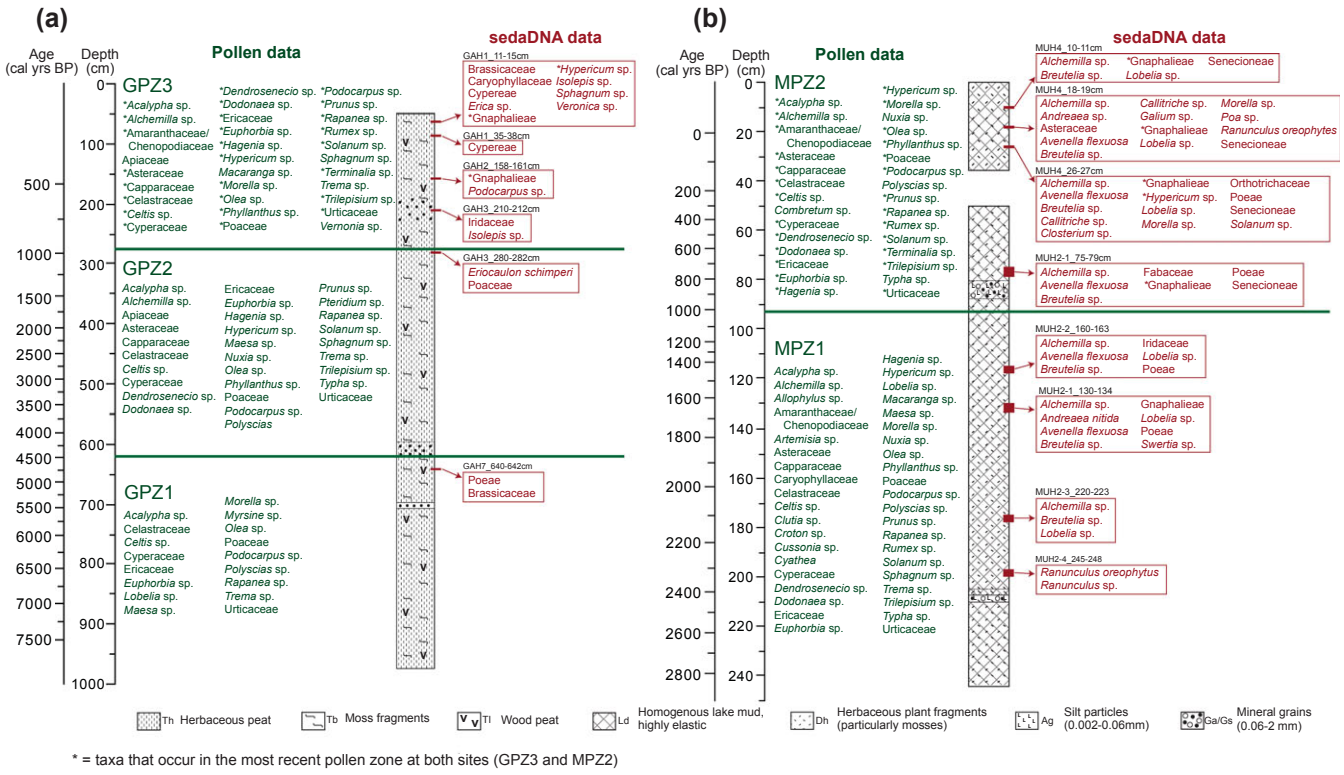

Figure 2. Sediment stratigraphy (based on Troels-Smith's (1955) classification, as described in Aaby and Berglund (1986)), pollen taxa with abundance over 1% (in green), and DNA taxa (in red) found in cores from (a) the Mt. Gahinga crater swamp and (b) the Mt. Muhavura crater lake. For the DNA samples, the location on the core is indicated by the solid red rectangles, where the height of the rectangle reflects the number of centimeters incorporated in the samples (GPZ, Mt. Gahinga pollen zone; MPZ, Mt. Muhavura pollen zone [McGlynn et al. 2013]).

enumerated pollen from species known to be present in forest at lower altitudes than the sample sites. This difference in size of the potential contributing areas between DNA and pollen was also apparent when comparing the diversity recovered from the past 1000 years between the 2 sites (Fig. 2). For the DNA, only 2 out of 26 taxa were detected in the sediments from both Mt. Gahinga and Mt. Muhavura, whereas for pollen 26 out of 34 taxa were recovered from both sites. Finally, DNA and pollen differed in taxonomic resolution; 71% and 22% of taxa were identified to species or genus level for the DNA and pollen analyses, respectively, whereas the remainder were identified to family level.

Discussion

The successful retrieval of authentic sedaDNA from sediment cores from 2 crater sites associated with the Albertine Rift testifies to the preservation of DNA in such samples up to, but not limited to, almost 5000 years old. Such long-term preservation is at least partly due to the maintenance of relatively cool temperatures at high altitudes in the humid tropics. The diversity recovered from sediment samples tended to decline with sample age, thereby conforming to the known postmortem, temporal degradation of DNA. For the Mt. Muhavura lake sediments, however, this drop is seen after circa 1700 years, whereas in the deposits from the Mt. Gahinga swamp, all but the youngest samples had 1 or 2 taxa only. This difference may be due to the more acidic sedimentary environments of peat-forming swamps compared with lakes, which can accelerate DNA decay (Lindahl 1993). The potential for sedaDNA to infer past plant diversity in more acidic wetland environments close to the equator thus seems limited to <100 years, whereas in high-altitude lake systems, this extends to a few thousand years. Even under acidic conditions, however, reconstructions of past variations in wetland vegetation based on sedaDNA data would appear to have the potential to extend far beyond monitoring-based records, which in many parts of central Africa at best date to only the last decade or so.

Relatively detailed reconstructions of past plant community composition are an essential component of meaningful restoration targets (Sanchez et al. 2013). Plant diversity recovered through sedaDNA may complement vegetation reconstruction from pollen records by revealing additional taxa such as *Callitriche* sp.,

Table 2. Taxonomic resolution of taxa identified by DNA and pollen analysis from the sediment samples collected at the 2 study sites in the Virunga volcanoes of the Albertine Rift, eastern Africa.[a]

Family	*sedaDNA*	*Pollen*
Asteraceae	Asteraceae	Asteraceae
	Gnaphalieae[b]	*Artemisia* sp.
	Senecioneae[b]	*Dendrosenecio* sp.
		Vernonia sp.
Brassicaceae	Brassicaceae[b]	
Campanulaceae	*Lobelia* sp.	*Lobelia* sp.
Caryophyllaceae	Caryophyllaceae	Caryophyllaceae
Cyperaceae	Cypereae[b]	Cyperaceae
	Isolepis sp.[b]	
Ericaceae	*Erica* sp.[b]	Ericaceae
Eriocaulaceae	*Eriocaulon schimperi*[b]	
Fabaceae	Fabaceae[b]	
Gentianaceae	*Swertia* sp.[b]	
Hypericaceae	*Hypericum* sp.	*Hypericum* sp.
Iridaceae	Iridaceae[b]	
Myricaceae	*Morella* sp.	*Morella* sp.
Plantaginaceae	*Callitriche* sp.[b]	
	Veronica sp.[b]	
Poaceae	Poaceae	Poaceae
	Avenella flexuosa[b]	
	Poeae∗	
Podocarpaceae	*Podocarpus* sp.	*Podocarpus* sp.
Ranunculaceae	*Ranunculus oreophytus*[b]	
Rosaceae	*Alchemilla* sp.	*Alchemilla* sp.
		Hagenia sp.
		Prunus sp.
Rubiaceae	*Galium* sp.[b]	
Solanaceae	*Solanum* sp.	*Solanum* sp.

[a]*Mosses are not included in this table because they do not produce pollen.*
[b]*Taxa only identified by DNA analysis.*

Gnaphalieae, Iridaceae, and *Ranunculus* sp. The absence of these taxa (even at family level) in the pollen records may reflect taphonomic processes because the current presence of these taxa at the site was noticed during fieldwork. Incorporating the analysis of sedaDNA may also improve vegetation records based on tropical lake sediments by increasing taxonomic resolution, as has been the case for polar-based research (Sønstebø et al. 2010; Jørgensen et al. 2012*a*; Parducci et al. 2013). This is particularly valuable for large and diverse families such as Poaceae and Ericaceae, with high representation in alpine environments. However, for certain families, the taxonomic resolution is limited when using short DNA markers such as the P6 loop in this study (e.g., Asteraceae, Cyperaceae), and additional family-specific markers are needed.

Combining pollen and sedaDNA data may help distinguish vegetation change at different geographic scales. In treeless alpine environments, the pollen signal is often dominated by distal wind-dispersed pollen, whereas local pollen production is low (Birks & Birks 2000). Distinguishing local versus regional pollen is a particularly pertinent issue in relation to pollen types that cannot be identified to a low taxonomic level. For example, Poaceae, Ericaceae, and Cyperaceae are all represented by afro-alpine and ericaceous vegetation types growing close to the Virunga crater sites (Fig. 2), but they could also (at least in part) be derived from a larger region. By contrast, sedaDNA is thought to be primarily local in origin: DNA from taxa located long distances from the site of deposition was not found in sedaDNA analyses of permafrost (Jørgensen et al. 2012*a*), and it has proved extremely difficult to extract chloroplast DNA directly from pollen (Parducci et al. 2005). In particular, the lack of DNA from, for example, *Olea* and *Celtis* in samples from parts of the cores that were dominated by pollen from these taxa (see figs. 6 and 7 in McGlynn et al. 2013) supports the assumed local origin of plant sedaDNA. However, pollen cannot be entirely excluded as a source for sedaDNA, as is proven by our observation of *Podocarpus* sp., *Morella* sp., and Fabaceae DNA, all of which are likely to be from forest trees growing outside the craters and at some distance below the rims.

The pollen records from the 2 sites were highly congruent, particularly for the last 1000 years (GPZ3 and MPZ2, Fig. 2): vegetational changes mostly reflected variation in montane forest vegetation, whereas little change is observed in the afro-alpine vegetation (McGlynn et al. 2013). In contrast, there was much less overlap in taxa represented at the 2 sites based on sedaDNA data, which would appear to reflect more closely the differences in conditions in the 2 craters. Thus, whereas the longer and steeper walls of the crater at Mt. Gahinga are characterized by dense stands of afro-alpine vegetation (particularly *Dendrosenecio* and *Lobelia*), more open, herbaceous vegetation is present on the relatively narrow, exposed rim of the crater at Mt. Muhavura (Fig. 1).

The reference DNA library we developed was essential for achieving reliable, high taxonomic resolution from sedaDNA. Large public databases such as GenBank contain errors in both sequences and taxonomy (e.g., Harris 2003), which can lead to spurious taxonomic inferences of the DNA data. Nevertheless, such large databases are needed to identify taxa missing in the local library. Local libraries allow stringent control to minimize errors, such as sequencing multiple specimens and ensuring appropriate expertise for taxonomic identification. Furthermore, local libraries restrict the number of plant species that are potentially represented in the data and thus increase taxonomic resolution of short DNA sequences (Taberlet et al. 2007; Sønstebø et al. 2010). Even though generating and curating local libraries is time-consuming and financially costly, such libraries are extremely valuable and open up research opportunities for many years to come.

We conducted the first plant sedaDNA analysis from an alpine region in the tropics and found that there

is a promising role for sedaDNA to complement vegetation reconstructions from traditional paleoecological data sources in high conservation priority areas and biodiversity hotspots. This is especially the case where records of ecologically meaningful long-term variations in the composition of biological communities and associated environmental information are difficult to acquire or simply unavailable, as is the case for many parts of the humid tropics in Africa, including the most biodiversity-rich and vulnerable areas. The successful retrieval of DNA from low-pH conditions that characterize Mt. Gahinga is particularly promising from a conservation point of view (especially given the higher number of taxa identified in the most recent samples) because many wetlands in central Africa have undergone extensive anthropogenic modification over the last 6 decades (Chapman et al. 2001), and the pace of change has quickened over the last circa 10 years (Maclean et al. 2011). Detailed inventories of preimpact ecological conditions are only very rarely available (Junk 2002). One consequence of increased exploitation is a wave of introduced and invasive taxa rolling through wetlands of many African countries (Howard & Chege 2007), while a dearth of detailed baseline ecological information renders establishing and then mitigating the impacts virtually impossible.

The field of sedaDNA is still in its infancy, and further studies are needed to explore some of the limitations of the method for conservation applications. Specifically, it is still unclear how different sources of bias may affect the taxa recovered. Such sources include variation in DNA preservation, divergence in the primer-binding site, sequence lengths differing between species (PCR tends to preferentially amplify shorter DNA regions), and the effect of the composition of the extract (including contaminant DNA; Boessenkool et al. 2012; Jørgensen et al. 2012*a*; Parducci et al. 2013). Until we have a better understanding of the different sources of bias affecting sedaDNA, these vegetation assemblages should be reconstructed conservatively and regarded as revealing a subset of the diversity only; sedaDNA may be most powerful as a tool for environmental reconstruction when used in combination with other sediment-based remains. Due to the biases, it is also currently not possible to analyze DNA sequence data quantitatively for inferring past plant abundance, as can be done with pollen records (but see Yoccoz et al. 2012 for modern plant sedDNA). Furthermore, the temporal precision of sedaDNA in lake sediments needs to be addressed: DNA leaching is thought to be unlikely in lake sediments (Anderson-Carpenter et al. 2011), but it has not been tested experimentally. With new studies addressing these current uncertainties, and the ongoing improvements in next-generation sequencing, we expect the range of environments and research questions where sedaDNA is an appropriate tool to continue to expand.

Acknowledgments

This work was supported by the Research Council of Norway (grant no. 191627/V40 to C.B.), the Norwegian Programme for Development, Research and Higher Education (NUFU) project 2007/10058 (AFROALP-II—Afroalpine sky islands: genetic versus taxonomic biodiversity, climate change, and conservation) to S.N. and C.B., and the Irish Research Council for Science, Engineering and Technology (to G.M.), with support also gratefully acknowledged from the National Geographic Society (NGS research grant "Climate change, fire, and ecosystem response on the Virunga volcanoes" to G.M. and D.T.) and a Natural History Museum, University of Oslo Småforsk grant to C.B. Permission to carry out collection of sediment cores from the Virunga volcanoes was provided to D.T. and G.M. by the Uganda Wildlife Authority (UWA/TDO/33/02) and the Ugandan National Council for Science and Technology (NS196). T. Wondimu, F. M. Tusiime, C. A. Masao, A. A. Abdi, P. Munishi, G. Mwachala, and G. Eilu helped collect the material for the taxonomic reference library, whereas J. Lejju, M. Murungi, S. Taylor, and Y. Okoth helped collect the sediment cores. The following people contributed to taxonomic identifications of specimens for the reference library: M. Thulin, K. Lye, and M. Escudero. We thank V. Mirré, A. Schrøder-Nielsen, S. Rucina, V. Muiruri, and S. McCarron for their assistance with laboratory-based analyses and E. Willerslev for providing laboratory space to extract the sediment samples.

Supporting Information

Supplementary methods, list of species included in the Afro-Alpine reference library version 1.0, number of reads remaining after each filtering step, assigned taxonomic identities of sequences per sample for both dataset 1 and dataset 2, and taxonomic identities of sequences removed from the datasets (Appendix S1) are available on-line. The authors are solely responsible for the content and functionality of these materials. Queries (other than absence of the material) should be directed to the corresponding author.

A fasta file of the taxonomic reference library version 1.0 and all sequences with their taxonomic identity (output from the program ecoTag) have been deposited in the Dryad Digital Repository: http://doi.org/10.5061/dryad.vg7b1.

Literature Cited

Aaby, B., and B. E. Berglund. 1986. Characterization of peat and lake deposits. Pages 231–246 in B. E. Berglund, editor. Handbook of Holocene palaeoecology and paleohydrology. Wiley, New York.

Anderson-Carpenter, L., J. McLachlan, S. Jackson, M. Kuch, C. Lumibao, and H. Poinar. 2011. Ancient DNA from lake sediments: bridging the

gap between paleoecology and genetics. BMC Evolutionary Biology **11:**30.

Bellemain, E., et al. 2013. Fungal paleodiversity revealed using high-throughput metabarcoding of ancient DNA from Arctic permafrost. Environmental Microbiology **15:**1176–1189. DOI: 10.1111/1462-2920.12020.

Birks, H. H., and H. J. B. Birks. 2000. Future uses of pollen analysis must include plant macrofossils. Journal of Biogeography **27:**31–35.

Boessenkool, S., L. S. Epp, J. Haile, E. V. A. Bellemain, M. Edwards, E. Coissac, E. Willerslev, and C. Brochmann. 2012. Blocking human contaminant DNA during PCR allows amplification of rare mammal species from sedimentary ancient DNA. Molecular Ecology **21:**1806–1815.

Bullock, J. M., J. Aronson, A. C. Newton, R. F. Pywell, and J. M. Rey-Benayas. 2011. Restoration of ecosystem services and biodiversity: conflicts and opportunities. Trends in Ecology & Evolution **26:**541–549.

Buytaert, W., F. Cuesta-Camacho, and C. Tobón. 2011. Potential impacts of climate change on the environmental services of humid tropical alpine regions. Global Ecology and Biogeography **20:**19–33.

Chapman, L. J., J. Balirwa, F. W. B. Bugenyi, C. Chapman, and T. L. Crisman. 2001. Wetlands of East Africa: biodiversity, exploitation and policy perspectives. Pages 101–131 in B. Gopal, W. J. Junk, and J. A. Davies, editors. Biodiversity in wetlands: assessment, function and conservation. Backhuys Publishers, Lieden, The Netherlands.

Coffey, E. E. D., C. A. Froyd, and K. J. Willis. 2011. When is an invasive not an invasive? Macrofossil evidence of doubtful native plant species in the Galapagos Islands. Ecology **92:**805–812.

Colwell, R. K., G. Brehm, C. L. Cardelús, A. C. Gilman, and J. T. Longino. 2008. Global warming, elevational range shifts, and lowland biotic attrition in the wet tropics. Science **322:**258–261.

Crutzen, P. J. 2002. Geology of mankind. Nature **415:**23.

Dimitrov, D., D. Nogués-Bravo, and N. Scharff. 2012. Why do tropical mountains support exceptionally high biodiversity? The Eastern Arc mountains and the drivers of Saintpaulia diversity. Public Library of Science ONE **7:**e48908. DOI:10.1371/journal.pone.0048908.

Epp, L. S., K. R. Stoof, M. H. Trauth, and R. Tiedemann. 2010. Historical genetics on a sediment core from a Kenyan lake: intraspecific genotype turnover in a tropical rotifer is related to past environmental changes. Journal of Paleolimnology **43:**939–954.

Ficetola, G. F., E. Coissac, S. Zundel, T. Riaz, W. Shehzad, J. Bessière, P. Taberlet, and F. Pompanon. 2010. An in silico approach for the evaluation of DNA barcodes. BMC Genomics **11:**434–443.

Franzen, M., and M. Molander. 2012. How threatened are alpine environments? A cross taxonomic study. Biodiversity and Conservation **21:**517–526.

Harris, J. D. 2003. Can you bank on GenBank? Trends in Ecology & Evolution **18:**317–319.

Hartter, J., M. D. Stampone, S. J. Ryan, K. Kirner, C. A. Chapman, and A. Goldman. 2012. Patterns and perceptions of climate change in a biodiversity conservation hotspot. Public Library of Science ONE **7:**e32408. DOI:10.1371/journal.pone.0032408.

Higgs, E. S., and W. M. Roush. 2011. Restoring remote ecosystems. Restoration Ecology **19:**553–558.

Howard, G. W., and F. W. Chege. 2007. Invasions by plants in the inland waters and wetlands of Africa. Pages 193–208 in F. Gherardi, editor. Biological invaders in inland waters: profiles, distribution, and threats. Invading Nature: Springer Series in Invasion Ecology, Volume 2. Springer, Dordrecht, The Netherlands.

Huse, S., J. Huber, H. Morrison, M. Sogin, and D. Welch. 2007. Accuracy and quality of massively parallel DNA pyrosequencing. Genome Biology **8:**R143. DOI:10.1186/gb-2007-8-7-r143.

Jørgensen, T., et al. 2012*a*. A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. Molecular Ecology **21:**1989–2003.

Jørgensen, T., et al. 2012*b*. Islands in the ice: detecting past vegetation on Greenlandic nunataks using historical records and sedimentary ancient DNA meta-barcoding. Molecular Ecology **21:**1980–1988.

Junk, W. J. 2002. Long-term environmental trends and the future of tropical wetlands. Environmental Conservation **29:**414–435.

Küper, W., et al. 2004. Africa's hotspots of biodiversity redefined. Annals of the Missouri Botanical Garden **91:**525–535.

La Sorte, F. A., and W. Jetz. 2010. Projected range contractions of montane biodiversity under global warming. Proceedings of the Royal Society B: Biological Sciences **277:**3401–3410.

Lind, E. M., and M. E. S. Morrison. 1974. East African vegetation. Longman, London.

Lindahl, T. 1993. Instability and decay of the primary structure of DNA. Nature **362:**709–715.

MacDougall, A. S., B. R. Beckwith, and C. Y. Maslovat. 2004. Defining conservation strategies with historical perspectives: a case study from a degraded oak grassland ecosystem. Conservation Biology **18:**455–465.

Maclean, I. M. D., R. R. Boar, and C. Lugo. 2011. A review of the relative merits of conserving, using, or draining papyrus swamps. Environmental Management **47:**218–229.

Margulies, M., et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380.

McGlynn, G., S. Mooney, and D. Taylor. 2013. Palaeoecological evidence for Holocene environmental change from the Virunga volcanoes in the Albertine Rift, central Africa. Quaternary Science Reviews **61:**32–46.

Montoya, D., L. Rogers, and J. Memmott. 2012. Emerging perspectives in the restoration of biodiversity-based ecosystem services. Trends in Ecology & Evolution **27:**666–672.

Murray, D. C., et al. 2012. High-throughput sequencing of ancient plant and mammal DNA preserved in herbivore middens. Quaternary Science Reviews **58:**135–145.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature **403:**853–858.

Parducci, L., Y. Suyama, M. Lascoux, and K. D. Bennett. 2005. Ancient DNA from pollen: a genetic record of population history in Scots pine. Molecular Ecology **14:**2873–2882.

Parducci, L., et al. 2012. Glacial survival of boreal trees in northern Scandinavia. Science **335:**1083–1086.

Parducci, L., I. Matetovici, S. L. Fontana, K. D. Bennett, Y. Suyama, J. Haile, K. H. Kjær, N. K. Larsen, A. D. Drouzas, and E. Willerslev. 2013. Molecular- and pollen-based vegetation analyses in lake sediments from central Scandinavia. Molecular Ecology **22:**3511–3524. DOI:10.1111/mec.12298.

Plumptre, A. J., et al. 2007*a*. The biodiversity of the Albertine rift. Biological Conservation **134:**178–194.

Plumptre, A. J., D. Kujirakwinja, A. Treves, I. Owiunji, and H. Rainer. 2007*b*. Transboundary conservation in the Greater Virunga Landscape: its importance for landscape species. Biological Conservation **134:**279–287.

Robbins, M. M., et al. 2011. Extreme conservation leads to recovery of the Virunga mountain gorillas. Public Library of Science ONE **6:**e19788. DOI:10.1371/journal.pone.0019788.

Sanchez, C., E. E. Gaiser, C. J. Saunders, A. H. Wachnicka, N. Oehm, and C. Craft. 2013. Challenges in using siliceous subfossils as a tool for inferring past water level and hydroperiod in Everglades marshes. Journal of Paleolimnology **49:**45–66.

Seppä, H., and K. D. Bennett. 2003. Quaternary pollen analysis: recent progress in palaeoecology and palaeoclimatology. Progress in Physical Geography **27:**548–579.

Smith, C. I., A. T. Chamberlain, M. S. Riley, A. Cooper, C. B. Stringer, and M. J. Collins. 2001. Neanderthal DNA: Not just old but old and cold? Nature **410:**771–772.

Sønstebø, J. H., et al. 2010. Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. Molecular Ecology Resources **10:**1009–1018.

Stoof-Leichsenring, K. R., L. S. Epp, M. H. Trauth, and R. Tiedemann. 2012. Hidden diversity in diatoms of Kenyan Lake Naivasha: a genetic approach detects temporal variation. Molecular Ecology **21:**1918–1930.

Taberlet, P., E. Coissac, F. Pompanon, L. Gielly, C. Miquel, A. Valentini, T. Vermat, G. Corthier, C. Brochmann, and E. Willerslev. 2007. Power and limitations of the chloroplast *trn*L (UAA) intron for plant DNA barcoding. Nucleic Acids Research **35:**e14. DOI:10.1093/nar/gkl938.

Troels-Smith, J. 1955. Karakterisering af løse jordarter (characterization of unconsolidated sediments). Danmarks Geologiske Underøgelse Series IV. **3**(10):73 pp.

Willerslev, E., A. J. Hansen, J. Binladen, T. B. Brand, M. T. P. Gilbert, B. Shapiro, M. Bunce, C. Wiuf, D. A. Gilichinsky, and A. Cooper. 2003. Diverse plant and animal genetic records from Holocene and Pleistocene sediments. Science **300:**791–795.

Willerslev, E., A. J. Hansen, and H. N. Poinar. 2004. Isolation of nucleic acids and cultures from fossil ice and permafrost. Trends in Ecology & Evolution **19:**141–147.

Willerslev, E., and A. Cooper. 2005. Ancient DNA. Proceedings of the Royal Society B: Biological Sciences **272:**3–16.

Willerslev, E., et al. 2007. Ancient biomolecules from deap ice cores reveal a forested southern Greenland. Science **317:**111–114.

Willis, K. J., and H. J. B. Birks. 2006. What is natural? The need for a long-term perspective in biodiversity conservation. Science **314:**1261–1265.

Yoccoz, N. G., et al. 2012. DNA from soil mirrors plant taxonomic and growth form diversity. Molecular Ecology **21:**3647–3655.