High-Throughput DNA sequencing of ancient wood
Stefanie Wagner, Frédéric Lagane, Andaine Seguin-Orlando, Mikkel Schubert,
Thibault Leroy, Erwan Guichoux, Emilie Chancerel, Inger Bech-Hebelstrup,
Vincent Bernard, Cyrille Billard, et al.

To cite this version:
Stefanie Wagner, Frédéric Lagane, Andaine Seguin-Orlando, Mikkel Schubert, Thibault Leroy, et al..
High-Throughput DNA sequencing of ancient wood. Molecular Ecology, Wiley, 2018, 27 (5), pp.1138-1154. 10.1111/mec.14514. hal-01783724

HAL Id: hal-01783724
https://hal-univ-rennes1.archives-ouvertes.fr/hal-01783724
Submitted on 21 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
High-Throughput DNA sequencing of ancient wood

Stefanie Wagner,1,2,3* Frédéric Lagane,1 Andaine Seguin-Orlando,2 Mikkel Schubert,2 Thibault Leroy,1 Erwan Guichoux,1 Emilie Chancerel,1 Inger Bech-Hebelstrup,1 Vincent Bernard,5 Cyrille Billard,6 Yves Billaud,7,8 Matthias Bolliger,9 Christophe Croutsch,10,11 Katarina Ćufar,12 Frédérique Eynaud,13 Karl Uwe Heussner,14 Joachim Köninger,15 Fabien Langenegger,16 Frédéric Leroy,17 Christine Lima,17 Nicoletta Martinelli,18 Garry Mombre,19 André Billamboz,20 Oliver Nelle,20 Antoni Palomo,21 Raquel Piqué,21 Marianne Ramstein,9 Roswitha Schweichel,22 Harald Stäuble,23 Willy Tegel,24 Xavier Terradas,25 Florence Verdin,26 Christophe Plomion,1 Antoine Kremer,1* and Ludovic Orlando2,3†

1BIOGECO, INRA, University of Bordeaux, 33620 Cestas, France. 2Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, 1350 Copenhagen, Denmark. 3Laboratoire AMIS, CNRS UMR 5288, Université Paul Sabatier (UPS), 31000 Toulouse, France. 4Museum Lolland Falster, 4800 Nyköbing Falster, Denmark. 5Dendro-Archaeology, CNRS, UMR 6566 CReAAH, University of Rennes, 35042 Rennes cedex, France. 6DRAC Basse Normandie, 61000 Alençon, France. 7MCC/DRASSM, 13016 Marseille, France. 8Edytem, University of Savoie, 73376 Le Bourget-du-Lac, France. 9Siedlungs- und Unterwasserarchäologie, Amt für Kultur/Archäologischer Dienst, 3001 Bern, Switzerland. 10Archéologie Alsace 11, 67600 Sélestat, France. 11UMR 7044 ARCHIMEDE, 67083 Strasbourg Cedex, France. 12Biotechnical Faculty, Department of Wood Science and Technology, University of Ljubljana, SI-1000 Ljubljana, Slovenia. 13Laboratoire EPOC, UMR 5805, University of Bordeaux, 33615 Pessac, France. 14Deutsches Archäologisches Institut, Zentrale, Referat Naturwissenschaften/Dendrochronologie, 14195 Berlin, Germany. 15Janus Verlag Freiburg im Breisgau, 79100 Freiburg im Breisgau, Germany. 16Laténium, Museum of Archaeology, 2068 Hauterive, Switzerland. 17Département des Recherches Archéologiques Subaquatiques et Sous-Marines, 13016 Marseille, France. 18Dendrodata s.a.s., 37126 Verona, Italy. 19Maritime Archaeology Trust, National Oceanography Centre, Southampton SO14 3ZH, United Kingdom. 20Baden-Württemberg State Office for Cultural Heritage, Tree-ring Lab, 78343 Hemmenhofen, Germany. 21Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain. 22Amt fuer Archaeologie im Thurgau, 8510 Frauenfeld, Switzerland. 23Landesamt fuer Archaeologie, 01109 Dresden, Germany. 24Institute for Forest Growth, University of Freiburg, Germany. 25Spanish National Research Council. IMF - Archaeology of Social Dynamics. Barcelona, Spain. 26CNRS, UMR 5607 Ausonius, Maison de l'Archéologie, 33607 Pessac cedex, France.

Key words: Quercus petraea/robur, oak, ancient DNA, degradation, temperate trees, chloroplast DNA

† These authors are joint last authors
* These authors are the corresponding authors: stefanie.wagner@inra.fr; antoine.kremer@inra.fr
Abstract

Reconstructing the colonization and demographic dynamics that gave rise to extant forests is essential to forecasts of forest responses to environmental changes. Classical approaches to map how population of trees changed through space and time largely rely on pollen distribution patterns, with only a limited number of studies exploiting DNA molecules preserved in wooden tree archaeological and subfossil remains. Here, we advance such analyses by applying high throughput (HTS) DNA sequencing to wood archaeological and subfossil material for the first time, using a comprehensive sample of 167 European white oak waterlogged remains spanning a large temporal (from 550 to 9,800 years) and geographical range across Europe. The successful characterization of the endogenous DNA and exogenous microbial DNA of 140 (~83%) samples helped the identification of environmental conditions favoring long-term DNA preservation in wood remains, and started to unveil the first trends in the DNA decay process in wood material. Additionally, the maternally-inherited chloroplast haplotypes of 21 samples from three periods of forest human-induced use (Neolithic, Bronze Age and Middle Ages) were found to be consistent with those of modern populations growing in the same geographic areas. Our work paves the way for further studies aiming at using ancient DNA preserved in wood to reconstruct the micro-evolutionary response of trees to climate change and human forest management.

Introduction

Forests cover 31% of the land area worldwide and are among the most biologically-rich and genetically-diverse ecosystems on the planet (MacDicken 2015). Reconstructing the evolutionary dynamics that gave rise to the current distribution patterns of long-lived sessile organisms, such as trees, is challenging, but essential to understand how these ecosystems respond to environmental change (Aitken et al. 2008; Franks et al. 2014). Classically, the postglacial recolonization routes of trees from glacial refugia have been documented with large-scale pollen, macrofossil and modern genetic data (Petit et al. 2002b; Cheddadi et al. 2006; Magri 2008; Tollefsrud et al. 2008; Liepelt et al. 2009; Wagner et al. 2015b). In addition, environmental DNA has been used as a complementary
proxy and opened new possibilities for the reconstruction of past forest communities applying high-throughput DNA sequencing (HTS) on sedimentary time series (Jørgensen et al. 2012; Parducci et al. 2012; Giguet-Covex et al. 2014; Smith et al. 2015; Pedersen et al. 2016; Schmid et al. 2017). Ancient DNA (aDNA) from wood material has also been investigated, but to a much more limited extent, with only a handful of studies hitherto reported (Tani et al. 2003; Pollmann et al. 2005; Deguilloux et al. 2006; Liepelt et al. 2006; Speirs et al. 2009; Gómez-Zeledón et al. 2017; Lendvay et al. 2017). These have relied on a methodological framework including the amplification of short PCR fragments and/or metabarcodes, real-time PCR and sequencing, including, for the most recent study, amplicon-sequencing on Illumina platforms (Lendvay et al. 2017). They have not yet explored the full potential of aDNA methods based on shotgun HTS, which have now become almost routine for the study of the genome-scale variation present in ancient animals and pathogens (Orlando et al. 2015).

Subfossil (non-petrified) wood logs and archaeological timber from temperate Europe are ubiquitous and, ancient DNA apart, have been extensively studied in a range of scientific areas, including paleoclimatology, dendrochronology and dendroarchaeology (Becker et al. 1991; Friedrich et al. 1999; Becker 2006; Büntgen et al. 2011; Billamboz 2014). Wood can be dated at a one-year resolution from tree-ring patterns, allowing the reconstruction of detailed history of forest cover changes, including natural vegetation succession and human exploitation (Čufar et al. 2010; Martinelli 2013; Billamboz 2014). Archaeological wood samples with exceptionally well-preserved morphological features have been excavated from waterlogged environments, such as groundwater-saturated wetlands and submerged archaeological layers within lakes and oceans (Čufar et al. 2010; Aguilera et al. 2011; Tegel et al. 2012; Martinelli 2013). In that regard, the prehistoric pile dwellings around the Alps, more than 800 of which have been classified as World Heritage sites by UNESCO (http://sites.palafittes.org/), constitute a unique archive for studying forest changes with large numbers of trees spanning multiple generations since the Neolithic period. Equally well-preserved subfossil logs corresponding to assemblies of submerged tree trunks have been reported from European foreshores and alluvial gravel pits (Friedrich et al. 2004; Nayling et al. 2007). However, after excavation, waterlogged wood decays very rapidly and little is known about the underlying microbial
taxa, which have mostly been characterized by microscopy and spectroscopy (Pedersen 2015), but extremely limited by DNA studies (Palla et al. 2014). The application of DNA-based molecular methods unveiling the microbiota associated with specific environmental requirements and/or degradation pathways may help facilitate both future aDNA analyses of waterlogged wood material and strategies for the ex situ and in situ conservation of waterlogged wooden cultural heritage items.

A number of studies have explored the potential of waterlogged environments for DNA preservation within plant remains of various types, including pips and seeds (Tani et al. 2003; Pollmann et al. 2005; Liepelt et al. 2006; Speirs et al. 2009; Schlumbaum et al. 2012; Wales et al. 2014, 2016; Gómez-Zeledón et al. 2017), and showed limited success on waterlogged wood (Tani et al. 2003; Pollmann et al. 2005; Deguilloux et al. 2006; Liepelt et al. 2006; Speirs et al. 2009). The methodology underlying all previous DNA studies of waterlogged wood relied on the PCR amplification of pre-selected DNA markers, and was not tailored to the molecular characteristics of ancient DNA, the vast majority of which consisting of ultra-short, degraded DNA molecules (Orlando et al. 2015). Additionally, dead cells are predominant in the heartwood (i.e. the inner part of the tree trunk), limiting the available amount of DNA molecules. Furthermore, wood material is notorious for containing large amounts of secondary metabolites and other inhibitory compounds hampering DNA experiments (Plomion et al. 2001; Deguilloux et al. 2002). It thus remains unclear whether the limited success of previous studies reflects methodological limitations or the poor potential for DNA preservation in waterlogged wood remains.

Here, we focused on two temperate trees, the pedunculate oak (Quercus robur) and the sessile oak (Quercus petraea). These two species are distributed across Europe and are represented by extensive paleontological and archaeological records, including wood remains compatible with dendrochronological dating (Becker et al. 1991; Friedrich et al. 2004; Becker 2006). The two species share multiple maternally-inherited chloroplast haplotypes (Petit et al. 2002a; Kremer et al. 2002), and species assignment requires appropriate diagnostic nuclear markers (Muir et al. 2000; Guichoux et al. 2011). The modern distribution of the chloroplast haplotypes is geographically structured and
has been studied based on more than 2600 modern populations (Petit et al. 2002a; b), providing a valuable reference for comparisons with ancient haplotypes. Oaks have long life expectancies (~500-800 years) and generation times (~30-80 years) (Lang 1994), and estimates of substitutions rates in chloroplast DNA within the Fagaceae family are low in comparison to other angiosperms, with substitution rates in the order of $10^{-10}$ substitutions per nucleotide per year (Frascaria et al. 1993). The presence in a given area of haplotypes otherwise characteristic of other geographical regions can thus be used with confidence to track recent migrations and/or anthropogenic translocations across different parts of the range (König et al. 2002; Lowe et al. 2004).

In this study, we applied HTS to aDNA extracts prepared from 167 subfossil and archaeological waterlogged samples of oak wood remains (*Quercus robur/petraea*), excavated from diverse environmental and taphonomical contexts across Europe and spanning the whole Holocene (Fig. 1, Table 1, Table S 1). We characterized aDNA preserved in waterlogged wood, identified potential environmental key factors favoring its preservation, uncovered DNA decay kinetics and tracked maternal haplotypes through time focusing on periods of human forest-use. This work paves the way for large-scale evolutionary studies on trees based on ancient organelle and nuclear DNA preserved in wood.

**Materials and Methods**

**Samples.** We collected 167 waterlogged wood samples from 26 sites spread across Spain, France, Switzerland, Italy, Slovenia, Germany, England and Denmark. These encompass a range of taphonomical and temporal (500 to 9,800 years) contexts (Fig. 1, Table 1) dated by dendrochronology and/or radiocarbon methods (Table S 1). Taxonomic resolution was limited to the *Quercus robur/petraea* level due to the close wood anatomical features present in both species (Schweingruber 1990). To increase the amount of initial DNA, we collected only samples with preserved sapwood (i.e. the outer, formerly functional part of the wood) (Plomion et al. 2001; Deguilloux et al. 2002). The wood remains from archaeological structures (e.g. pile dwellings, water
wells, fish weirs) and subfossil logs sampled were waterlogged and located in ground- or seawater-saturated wetlands or submerged in lakes or oceans. We gathered at least four samples from each site, except for Reims and Unteruhldingen site 1, for which only three samples could be studied. Most samples had been excavated within the last three years, but we also included 47 additional samples, which were part of earlier collections, in order to increase the temporal and spatial resolution. We note that the samples collected at Bouldnor Cliff consist of subfossil tree remains especially excavated from two new sampling localities, representing submerged peats rich in wooden macro-remains, and did not consist of sedimentary cores, in contrast to previous work (Smith et al. 2015; Weiß et al. 2015; Kistler et al. 2015). We reconstructed the storage conditions for all sites, including the archived samples (Table 1, Table S 1). At 14 sites, the samples were vacuum-packed and stored at +4°C immediately after their excavation, to minimize molecular degradation.

**DNA extraction, library construction and sequencing.** DNA extraction, library construction and PCR were performed at the dedicated ancient DNA facilities of the Centre for GeoGenetics, University of Copenhagen, Denmark. We focused on sapwood and combined previously published DNA extraction protocols (Cappellini et al. 2010; Wales et al. 2012; Gamba et al. 2016) to maximize DNA recovery. Our procedure consisted of the following seven steps: (1) removal of the external surface with sterile scalpels, (2) grinding of 1x1x0.2 cm$^3$ humid wood cuboids in sterile mortars, (3) digestion of the resulting wood paste at 37°C for 20 h in 3 mL lysis buffer consisting of 10 mM Tris-HCl, 10 mM NaCl, 2% w/v SDS, 5 mM CaCl$_2$, 2.5 mM EDTA, 40 mM DTT and 10% proteinase K, (4) extraction of DNA twice in phenol and then once in chloroform, (5) mixing of the DNA supernatant with 3 mL Tris-EDTA (1X), (6) concentration on Amicon Ultra 4 (30 kD) filters and (7) purification on MinElute PCR purification columns (QIAGEN, Hilden, Germany). For five samples, pairwise comparisons of the DNA extraction yields for sapwood and heartwood material were carried out. We expected higher DNA yields in sapwood, given the higher proportion of living cells present in this material relative to heartwood. We treated 13 DNA extracts with the USER enzyme mixture (uracil-DNA–glycosylase and endonuclease VIII), as described by Gamba et al. (2016), before DNA library construction, amplification and sequencing, with the aim of achieving a minimum of 8X
DNA libraries were constructed by blunt-end DNA ligation, with 16 μL of DNA solution, in a final volume of 25 μL, as described by Meyer and Kircher (2010) and Gamba et al. (2016). DNA library concentration was assessed by real-time PCR (Roche Light Cycler 480) and SYBR Green chemistry. Real-time PCR was performed in duplicate, with a 20 μL reaction volume consisting of 1 μL 1:20-diluted library, 0.1 U/μL TaqGold (Invitrogen, Life Technologies), 1X Gold Buffer, 4 mM MgCl₂, 1 mg/mL BSA, 1 mM dNTPs, 0.2 mM Illumina in PE1.0 primer (5’-AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC T) and 0.2 mM Multiplex Index Primer (5’- CAA GCA GAA GAC GGC ATA CGA GAT NNN NNN GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG), where NNN NNN is a unique, 6-mer index tag. Amplification conditions were set as follows: 10 minutes at 92°C (initial activation by heating), 40 cycles of 30 seconds at 92°C (denaturation), 30 seconds at 60°C (annealing) and 30 seconds at 72°C (elongation), and then 7 minutes at 72°C (final elongation). We determined Ct values with the second-derivative maximum algorithm implemented in LightCycler software. Indexing PCR was performed in the same conditions as real-time PCR, except that the final volume was 25 μL, the volume of undiluted library used was 4 μL, and the number of amplification cycles was set at the previously determined Ct value minus 5, to correct for dilution effects. PCR products were purified on MinElute columns (QIAGEN, Hilden, Germany), in a final volume of 25 μL and were diluted 1:10 for quantification on a 2200 TapeStation instrument (Agilent Technologies) with High-Sensitivity ScreenTape D1000 reagents (Agilent Technologies). Libraries were then pooled and shotgun-sequenced at the Danish National High-Throughput Sequencing Centre, on a HiSeq2500 instrument, with 100 single-end cycles.

Contamination control. Blank controls were included in all laboratory steps and were analyzed in parallel with the samples. They encompassed extraction blank controls (EBC), library blank controls (LBC), PCR controls (qPBC) and an indexing blank control (iPBC). For real-time PCR quantification, the Ct value obtained for blank controls was at least 10 cycles greater than that for the samples. Assuming a reaction efficiency of 100%, this limits the maximal contamination level to 0.1% (Table
Blank controls were also PCR-indexed with the maximum number of cycles observed across the experimental sample sets, but negative profiles were obtained on the TapeStation instrument.

**Endogenous and exogenous DNA content.** Raw reads were demultiplexed based on exact index match. We mapped the sequencing reads independently against a haploid version of the *Quercus robur* reference genome (740 Mb/C: Plomion *et al.* (2016) Leroy *et al.* (2017)), using PALEOMIX (Schubert *et al.* 2014), with the bwa-backtrack algorithm within the Burrows-Wheeler Aligner (Li *et al.* 2009). Seeding was disabled as per the recommendations from Schubert and colleagues (Schubert *et al.* 2012) and alignments with a mapping quality below 25, unmapped reads and PCR duplicates were discarded. The same procedure was repeated for identifying chloroplast DNA reads, using the chloroplast genome as a reference (160 kb: https://w3.pierroton.inra.fr/QuercusPortal/index.php?p=GENOMIC_SEQ). Independent alignment of DNA reads against the nuclear and organellar genomes has been previously recommended to avoid alternative hits in the nuclear genome due to the presence of nuclear organellar insertions (such as numts, for mitochondrial DNA). Such an alignment strategy can, however, result in the alignment against the organellar genome of divergent reads pertaining to nuclear organellar insertions and inflate heterozygosity levels. The limited heterozygosity levels observed on the haploid chloroplast genomes ruled out any significant impact of nuclear organellar insertions on the reconstructed chloroplast genomes (Fig. S 1). Endogenous DNA content was calculated as the fraction of high-quality nuclear and chloroplast sequences identified within the pool of sequences left after adapter trimming, quality filtering, and PCR duplicate removal. Fragment size distributions were approximated from the size distribution of high-quality read alignments. We note that the latter provides an under-estimate of the true size distribution of ancient DNA fragments, given that the full length of any template larger than 100 nucleotides cannot be determined from 100 cycle Single-End sequencing reactions. However, we note that extent of the size bias is extremely limited for the vast majority of the samples analysed, given that most aligned reads are extremely short, with the fraction of reads longer than 95 nucleotides generally representing ~0.6% (median, nuclear sequences) of the aligned reads (Fig. S 2).
Trimmed single-end reads were further processed against the MetaPhlAn2 database (Segata et al. 2012; Truong et al. 2015), including sequences from eukaryotes and viruses, with metaBIT (Louvel et al. 2016), to assess the exogenous microbial diversity captured within our samples. Reads were aligned with BOWTIE2 v2.1.0 (Langmead & Salzberg 2012), and the microbial composition of each sample was visualized with a range of graphical outputs provided by metaBIT, including Bray-Curtis distances (Truong et al. 2015; https://bitbucket.org/nsegata/hclust2/; Louvel et al. 2016). Shannon indices were calculated with the exclusion of taxa with an abundance <1% and fewer than 1,000 hits. Finally, the linear discriminant analysis effective size (LEfSe) method (Segata et al. 2011; Louvel et al. 2016) was run to identify and quantify the microbial taxa driving significant differences between predefined clusters identified in the principal coordinate analysis. The LEfSe analysis used default settings (non-parametric Kruskal Wallis rank-sum test, unpaired Wilcoxon test with p-values of 0.05, and threshold values of 2 for the logarithmic LDA score).

**Authenticity.** We checked that the DNA obtained was authentic by assessing DNA fragmentation and nucleotide misincorporation patterns for nuclear and chloroplast DNA, with mapDamage2.0 (Jónsson et al. 2013; Schubert et al. 2014). For samples with at least 1,000 unique hits per target genome, we quantified deamination levels in DNA overhangs (δs), with mapDamage2.0 (Jónsson et al. 2013). This parameter was used as a proxy for post-mortem DNA degradation and, together with average fragment lengths, was plotted against thermal age. Thermal age values were calculated with the thermal age webtool (http://thermal-age.eu/) (Table S 1), which defines thermal age as the time necessary to produce a given degree of DNA degradation at a constant temperature of +10°C (Smith et al. 2003). The model inputs were chronological sample age, geographic coordinates, altitude, type and thickness of the burial layers, collection year, analysis year and storage conditions (Table S 1).

**Genotyping and chloroplast haplotype calling.** For the 13 samples for which we had USER-treated and non-USER-treated datasets, we rescaled individual base qualities for the non-USER-treated extracts, using mapDamage2.0 (Jónsson et al. 2013). Rescaled read alignments were subsequently merged together with those obtained for USER-treated data with SAMtools v0.1.18 (Li et al. 2009).
The PALEOMIX phylo pipeline was used to reconstruct chloroplast haplotypes from 34 SNP positions known to be polymorphic within a large reference panel of ~2,600 modern haplotypes (Petit et al. 2002a; Guichoux & Petit 2014). The minimum read depth for genotype calls was set to four, and the minimum quality score was set to 30. Finally, the ancient haplotypes were compared with the haplotypes found in extant populations within a radius of 100 km around the study sites (Table S3). The latter were extracted from a georeferenced database storing genetic data from >3000 European white oak populations from across Europe (https://gd2.pierroton.inra.fr/gd2/login/login).

Results

Endogenous and exogenous DNA content. We recovered shotgun sequence data from the DNA extracts of 140 waterlogged sapwood samples for 24 of the 26 sampling sites. The two sites where DNA levels of all samples were equivalent to that found in extraction blanks were the Amélie foreshore and the Lillemer peatlands, where the samples were gathered from piles exposed to tides and corresponded to relatively dry logs. For the 140 samples delivering sequence data, the levels of endogenous DNA present in each individual library were found to be low, <1% for the vast majority of samples (Fig. 2, Table S4). However, endogenous DNA content was higher for 14 samples showing chronological ages between 0.5 and 5.2 thousand years (thermal ages between 0.4 and 5.6 thousand years), with unique high-quality alignments making up to 16.5% of the sequences generated (Table S4). Interestingly, the two samples showing highest endogenous DNA content consisted of submerged piles inserted in limy lacustrine sediments from Lake Constance in southern Germany (Unteruhldingen/site 1, 3.7 thousand years, 4.8 thousand thermal years). The following samples showing high endogenous DNA levels were much younger (0.5-0.7 thousand years; 0.4-0.5 thousand thermal years) and were recovered from organic wetland sediments in northern Germany. All the other millennia-old samples with endogenous DNA contents >1% came from terrestrial and submerged limy contexts (Erstein, Eschenz-Orkopf and Sutz Lattrigen Ruette). All but one of these samples had been stored in controlled conditions (placed under vacuum and at +4°C immediately after excavation).
Read mapping against the MetaPhlAn2 database revealed the presence of diverse microbial communities, with Shannon indexes ranging from 0.05 to 3.01 (median value = 2.09; Fig. 3a, Fig. S 3, Table S 5) and 122 DNA libraries satisfying the threshold of at least 1,000 hits against the reference database. The sample with the highest endogenous DNA content displayed a predominance of the various members of the Methanomicrobia class (Fig. S 4). Principal Coordinate Analysis (PCoA) based on microbe abundances in all 122 samples (irrespective of the storage conditions) revealed no apparent structure reflecting similar environmental origins (Fig. S 5), except for samples of marine origins which projected on a closer PCoA space. However, PCoA restricted to samples stored in controlled conditions revealed clear clustering by site and/or taphonomical conditions, with the first three axes explaining more than 35% of the total variance for microbial abundance (Fig. 3b).

Although additional subclusters may be present (Fig. S 6), we restricted subsequent LEfSe analyses to the four most apparent such PCoA clusters corresponding to (1) organic wetland sediments (Greifswald, Hamburg), (2) limy lake sediments (Conjux, Eschenz Orkopf, Unteruhldingen/site 1 and Unteruhldingen/site 2), (3) marine sediments (Bouldnor Cliff, Etang de Thau and Champeaux) and (4) sands and loess on limy bedrock (Erstein 208). Samples from the coastal site at Le Gurp, which had a projected position intermediate between those of clusters of marine and terrestrial sediments, were disregarded. LEfSe revealed the presence of a number of taxa driving differences in the microbial composition of the four clusters (LDA scores >3) (Fig. 3c, Fig. S 7). Samples from organic wetland sediments were particularly rich in anaerobic Methanosetaeacteae, particularly *Methanosaeta concilii*, and Sphingomonadaceae, represented by *Sphingobium spec.*. Marine environments were characterized by a predominance of Gammaproteobacteria, including the Shewanellaceae (*Shewanella spec.*, *Shewanella frigidimarina*) and, typically, marine Oceanospirillaceae and Piscirickettsiaceae. The limy soil samples were characterized by a predominance of Betaproteobacteria, including the orders of Burkholderiales (Oxalobacteraceae, *Janthinobacterium*) and Pseudomonadales (*Pseudomonas* sp., *Pseudomonas mandelii*). Finally, samples from limy lacustrine environments included families not abundant in the other environments, such as Alphaproteobacteria, Caulobacteraceae (*Caulobacter* sp., *Caulobacter vibrioides*) and other phylotypes, such as Euryarchaeota (*Methanocella* sp.). Several of
these taxa can break down lignin (e.g. *Sphingobium* spec.), cellulose (e.g. *Caulobacter* spec.) and lignin-derived aromatic molecules (e.g. *Methanosarcina* spec. *Methanocella* spec. *Rhodopseudomonas* spec.) (Taylor 1983; Harwood & Gibson 1988; Song et al. 2013; Tsuji et al. 2015; Kato et al. 2015).

**DNA authenticity and degradation.** A total number of 77 samples originating from 16 sites across Europe yielded at least 1,000 unique high-quality hits against the pedunculate oak nuclear reference genome. The thermal ages of these samples were between 0.4 and 8.7 thousand years (chronological ages between 0.5 and 5.6 thousand years). DNA fragments were of limited size, even in samples younger than 1 thousand thermal years, in which the vast majority of the endogenous DNA templates identified were shorter than 80 bp long. We found that the average size of endogenous DNA fragments tended to decrease linearly with increasing thermal age (Fig. 4a, \( f(x)=63.57-0.001x, \ r^2=0.202, \ p=4.05e-5 \)). However, the magnitude of the slope of the linear model was limited and the linear model poorly explained the data observed in multiple samples, which appeared as clear outliers to the confidence range predicted by the model. Additionally, there were large size differences between samples of similar age. Altogether, this suggests that thermal age as well as additional factors, possibly related to unmodelled micro-environmental conditions, impacted on DNA fragmentation. We quantified post-mortem deamination in single-strand overhangs (\( \delta_s \)) and found a significant correlation with thermal age until approximately 5 thousand thermal years ago (Pearson coefficient = 0.746, \( p<2e-5 \), Fig. 4b), when \( \delta_s \) values reached a plateau. However, deamination was variable across samples of similar age. Nuclear and chloroplast DNA displayed the same cytosine deamination kinetics after death, as the correlation between nuclear \( \delta_s \) and chloroplast \( \delta_s \) values was found to be highly significant \( (r^2=0.992, \ p<2e-16, \) Fig. S 8), with a slope close to 1. In absence of significant difference between nuclear and chloroplast damage kinetics, we repeated the damage analyses, applying the threshold of 1,000 unique high-quality alignments to the sum of nuclear and chloroplast hits. This resulted in the inclusion of four additional data points (for a total of 81 samples, Fig. S 8), two of which representing samples from sites already represented in the analyses above (Unteruhldingen site 2, Conjux) and two from an additional site, the Bouldnor Cliff (Sample 523 and...
This analysis confirmed the findings above. Given that the authenticity of plant DNA sequences identified at the Bouldnor Cliff site has been challenged (Smith et al. 2015; Weiβ et al. 2015; Kistler et al. 2015), it is noteworthy that the oak DNA templates identified from two of the samples from a first locality at Bouldnor Cliff (samples 523 and 524) are characterized by very small fragment sizes (<40bp), strong purine enrichment in the genomic positions preceding read start (in support of post-mortem DNA fragmentation through depurination; Briggs et al. 2007) and nearly maximal cytosine deamination rates within overhangs (δs=0.97). However, no evidence of cytosine deamination or depurination was found in the third sample, that was characterized by extremely small fragment sizes (<30bp). This sample comes from a second locality at Bouldnor Cliff (sample 522). In absence of clear post-mortem DNA damage evidence, the data pertaining to this third sample should be taken with caution, especially given that the alignment false positive rate is likely significantly inflated for DNA fragments of such extremely limited size. We, however, note that previous work have reported limited DNA depurination levels in marine sediments (Corinaldesi et al. 2008) and that the differences in DNA degradation levels observed might reflect different micro-environmental conditions (and contamination rates) prevailing in the three samples and/or within the two localities investigated at Bouldnor Cliff. Some of the other archaeological sites where we analyzed multiple samples have indeed shown significant differences in DNA damage levels (Fig. 4). Further work will be needed to investigate the specific DNA degradation conditions at Bouldnor Cliff. Additional information and discussion on the Bouldnor Cliff samples compared to previous DNA results obtained for sedimentary samples (Smith et al. 2015) are provided as Supplementary Material (Figs S 9–S 13, Table S 6 and Text S 1).

The base composition at the genomic positions preceding read starts and following read ends were typical of ancient DNA fragmentation by depurination (Figs 4c and 4d). Typical misincorporation patterns were also observed, with higher rates of C-to-T substitution close to sequence starts, and higher G-to-A substitution rates at sequence ends (Figs 4c and 4d). This pattern indicates preferential post-mortem deamination within overhanging ends (Briggs et al. 2007). Finally, we compared the endogenous DNA content of libraries constructed from sapwood and heartwood extracts from five
pairs of samples. Although they were all characterized by extremely low endogenous DNA content (<0.13%, Fig. S 7), sapwood showed higher endogenous DNA levels than the heartwood in four of the five pairs of samples. This situation reflects the known, larger amounts of DNA in fresh sapwood than in heartwood (Deguilloux et al. 2002) and suggests that the oak DNA detected is truly endogenous, as environmental contamination would be expected to result in similar amounts in the sapwood and heartwood. This comparison adds on the evidence supporting the authenticity of our results, especially in situations where environmental conditions can hardly be controlled (e.g. sampling of wood during pollen season, uncontrolled archive conditions).

**Inference of chloroplast haplotypes.** The sequencing of USER- and non-USER-treated libraries led to the characterization of 21 chloroplast genomes with a mean depth-of-coverage exceeding 5X, and 11 chloroplast genomes with a mean coverage between 1X and 5X (Table 2). The nuclear coverage achieved was generally limited (<1%), except for two samples, both from Unteruhldingen/site 1 (3.7 thousand years, 4.8 thousand thermal years), where 14% nuclear coverage could be achieved using 13.5 and 10 million sequencing reads, respectively. The characterization of the complete nuclear sequence of ancient oak genomes in such samples is thus possible at a relatively reasonable cost, even in the absence of DNA library enrichment.

For the 21 samples showing chloroplast coverage >5X (0.5-5.6 thousand years, 0.4-8.7 thousand thermal years), we identified five different ancient chloroplast haplotypes from three different major lineages (Petit et al. 2002a): haplotypes 10 and 11 belonging to lineage B (currently spread from Spain to the British Isles), haplotype 1 belonging to lineage C (from Italy to Scandinavia) and haplotype 7 belonging to lineage A (from the Balkans to the Baltic Sea and southern France). At all sites and for all time periods considered (Neolithic, 7 haplotypes; Bronze Age, 7 haplotypes; Middle Ages, 7 haplotypes), the ancient haplotypes matched the most frequent haplotypes found in nearby extant populations (Fig. 5).
Discussion

Analyses of ancient DNA from trees can provide invaluable temporal series to investigate the response of forest ecosystems in the face of major past environmental changes. However, previous studies on ancient tree DNA have been mostly restricted to the analysis of environmental records and pollen grains (Parducci et al. 2005, 2012; Giguet-Covex et al. 2014; Willerslev et al. 2014; Smith et al. 2015; Pedersen et al. 2016; Schmid et al. 2017). The limited number of studies retrieving aDNA from wood archaeological and subfossil material were mostly based on the PCR amplification of a few candidate markers (Tani et al. 2003; Deguilloux et al. 2006; Liepelt et al. 2006; Speirs et al. 2009; Gómez-Zeledón et al. 2017) and most recently on amplicon-sequencing on Illumina platforms (Lendvay et al. 2017). They have not yet explored the full potential of aDNA methods based on shotgun HTS. In this study, we applied such a methodology for the first time to a comprehensive collection of 167 subfossil and archaeological samples of European oak (Quercus robur/petraea) wood remains. Our collection spans 26 sites across the whole western European range of this species assemblage (42-55°N/-2°W-15°E), and encompasses various taphonomical (terrestrial/lacustrine/marine) and temporal (0.5-9.8 thousand years) contexts.

We focused on waterlogged environments as these preserve wood material in an exceptional morphological state. By relying on sapwood, which consists of the cell-rich outer, formerly functional part of the wood, we maximized the available amount of DNA. With the molecular tools used in this study, we successfully recovered shotgun sequence data from 140 samples from 24 sites. Only two taphonomical contexts, corresponding to aerobic and/or acidic environments known to accelerate DNA decay (Collins et al. 2002; Smith et al. 2003; Nielsen-Marsh et al. 2007), were not amenable to DNA sequencing. Most of the sequenced samples were dominated by environmental microbial DNA, as is also generally observed in animal calcified tissues, such as bones and teeth, with the exception of petrosal bones where higher DNA preservation can be observed (Pinhasi et al. 2015). However, 14 samples (0.5-5.2 thousand years) had endogenous DNA contents of up to 16.5%, a value compatible with relatively cost-effective genome sequencing by shotgun sequencing.
The oak DNA samples showing the highest levels of endogenous DNA were all associated with large fractions of strictly anaerobic, methane-reducing Methanomicrobia. This suggests that anoxic conditions providing environmental conditions compatible with the survival and multiplication of such microbiota may favor the preservation of ancient DNA in waterlogged environments. We also note that all the millennia-old samples with moderate-to-high oak DNA contents (1-16.5%) originated from archaeological structures (pile dwellings, water wells) embedded in limy contexts. This suggests that such calcareous sediments could represent particularly promising environments for the preservation of ancient DNA in wooden tissues. Similar endogenous DNA contents (3-9.8%) have been obtained for archaeological samples from organic wetland deposits of neutral pH, as indicated by the molecular identification of the bacterium *Methanosaeta concilii* (pH values of 7.1-7.5). However, these samples were relatively young (<1 thousand years) and it remains unclear whether such environments are favorable for DNA preservation over longer periods of time.

Our microbial abundance results showed that storage in warm aerobic conditions greatly altered the composition of the microbial communities since clear environmental clustering of microbial profiles could only be observed using sequence data from to samples stored in controlled conditions. It is possible that such storage conditions also accelerated the decay of ancient wood remains as various microbial taxa involved in the degradation of lignin, cellulose or lignin-derived aromatic compounds could be identified (Taylor 1983; Harwood & Gibson 1988; Song *et al.* 2013; Tsuji *et al.* 2015; Kato *et al.* 2015). Further work is, however, required before the exact impact of storage conditions on DNA preservation levels can be assessed, and biomarkers of molecular DNA degradation based on taxonomic composition of microbiota can be developed.

We provide the first clear molecular evidence for the authenticity of aDNA extracted from wood at the nucleotide level reflected in typical post-mortem damage patterns, small fragment sizes and progressive DNA decay over a period of 8.7 thousand thermal years. Post-mortem cytosine deamination rates within single-stranded overhangs tended to increase with age and were found
saturated from approximately 5 thousand thermal years ago. We found that the DNA fragment size was linearly correlated with thermal age but that other factors can be expected to impact the rate of depurination-driven DNA fragmentation in waterlogged environments (Corinaldesi et al. 2008). We uncovered similar decay kinetics for organellar and nuclear DNA which is in agreement with findings made for herbarium specimen (Weiβ et al. 2016) but contrasts with results on bone material (Schwarz et al. 2009; Allentoft et al. 2012; Kistler et al. 2015, 2017). Consistent with the presence of heterogeneous microenvironmental conditions driving DNA decay as formerly reported for animals (Allentoft et al. 2012), we found a lot of variation between samples of similar age.

In 21 oak samples with a chloroplast DNA coverage >5X, we identified ancient chloroplast haplotypes from Neolithic sites in Switzerland (Bieler See, Eschenz Orkopf) and Slovenia (Strojanova voda, Spica), Bronze Age sites in Germany (Unteruhldingen) and France (Erstein), and Medieval sites from northern Germany and northern France (Hamburg, Greifswald, Champeaux). All of these haplotypes matched the most frequent haplotypes present in nearby modern populations. This pattern is compatible with a continuity of the maternally-transmitted lineages within these areas. In the absence of sufficient nuclear DNA data, this pattern is, however, not sufficient to establish continuity and does not allow us to rule out alternative scenarios, such as cross-regional movements within the range of the respective haplotype distributions. Given the limited evolutionary rate of the chloroplast genome, it is highly unlikely that even a single of the 34 nucleotide sites used for defining haplogroup can be affected by a mutational event. Hence, populations that split a few centuries and/or millennia ago can show identical chloroplast haplotypes, thereby limiting our capacity to disentangle local and non-local origins from chloroplast DNA information. Additionally, the geographic and temporal range for which chloroplast DNA data could be collected is admittedly limited, and not sufficient to test whether the genetic structure of modern population was already established during the Early Holocene (Petit et al. 2002b). Future large-scale screenings of both nuclear and chloroplast aDNA should aim at charting ancient haplotype distributions through space and time, in order to clarify geographical origin and timing of human-induced translocations (König et al. 2002; Lowe et al. 2004; Wagner et al. 2015a).

Future work should undertake the characterize sufficient amount of nuclear DNA variation before
formal continuity tests such as those developed by Rasmussen and colleagues (Rasmussen et al. 2014) can be implemented. For now, we note that pollen, charcoal and/or tree-rings data available in the western Lake Constance area, where two of our study sites are located (Unteruhldingen, 3.7 thousand years, and Eschenz-Orkopf, 5.2 thousand years), suggest that oak populations expanded into Corylus-dominated woodland seven to eight thousand years ago (Fig. 6). Early Neolithic forest clearance (cutting, burning) then promoted the expansion of beech and a decrease in the size of the oak population. Future characterization of nuclear data will help determine whether the ancient wood remains collected from Eschenz Orkopf and Unteruhldingen presumably originated from remaining sources of oaks following human-driven coppicing and clearings during the Neolithic and Bronze Age (Fig. 6, (Billamboz 2014; Billamboz & Martinelli 2015)), or by the short-distance movements of acorns from nearby stands.

In conclusion, our study highlights the potential value of waterlogged wood as a source of ancient tree DNA. The methodology described paves the way for large-scale paleogenomic investigations of temperate trees and the genetic consequences of environmental changes during the Holocene. These results open up new possibilities for discovering ancient phylogeographic patterns and micro-evolutionary trajectories in trees, and for improving conservation strategies for cultural heritage items composed of wood.

Acknowledgments
This research was supported by the European Research Council, through the Advanced Grant Project TREEPEACE (#FP7-339728) and the Consolidator PEGASUS (#H2020-681605). Sequencing was performed at the Danish National High-Throughput Sequencing Centre, Copenhagen, Denmark. We thank Nathan Wales, Enrico Cappellini, Cristina Gamba, Charleen Gaunitz, Antoine Fages, Kristian Hanghøj and Clio Der Sarkissian for support and assistance during extraction, library construction and data analysis; Matthew Collins, Rémy Petit, Isabelle Lesur, Francois Ehrenmann, Lucie Kremer, Laura Truffaut, Michael Neumann, Manfred Rösch, Pernille Selmer Olsen and Tina Blumensaad.
Brandt for discussions and assistance. Site permissions were obtained for Lucone from the Ministero dei beni e delle attività culturali e del turismo (Prot. N 993 Cl. 34.07.19,f.2) and for Viverone from the Direzione regionale per i beni culturali e paesaggistici del Piemonte (Prot. 773 cl. 34.16.1919.2).

Author contributions
Conceived and designed the experiments: SW, AK and LO. Performed the experiments: SW and ASO, with input from LO. Analyzed the data: SW and LO. Contributed materials/experimental/analysis tools: SW, FL, MS, TL, EG, EC, IBH, VB, AB, CB, YB, MB, CC, KC, FE, UH, JK, FL, FL, CL, NM, GM, ON, AP, RP, MR, RS, HS, WT, XT, FV, AK and LO. Wrote the paper: SW, CP, AK and LO.

References
Billamboz A, Martinelli N (2015) Dendrochronology and Bronze Age pile-dwellings on both sides of


Kistler L, Smith O, Ware R et al. (2015) Thermal age, cytosine deamination and the veracity of 8,000 year old wheat DNA from sediments. bioRxiv.


Lendvay B, Hartmann M, Brodbeck S et al. (2017) Improved recovery of ancient DNA from subfossil wood - application to the world's oldest Late Glacial pine forest. New Phytologist. doi: 10.1111/nph.14935


Louvel G, Der Sarkissian C, Hanghøj K, Orlando L (2016) metaBIT, an integrative and automated metagenomic pipeline for analysing microbial profiles from high-throughput sequencing shotgun


Data Accessibility

All the sequencing data are available for download from the European Nucleotide Archive under accession number PRJEB21188.

Supplementary Material

Supplementary Figures
“SI_AncientOak.docx”
Fig. S1 | Observed chloroplast DNA heterozygosities.
Fig. S2 | Fragment lengths distribution across all samples.
Fig. S3 | Heatmap of microbe species found across all samples.
Fig. S4 | Heatmap of microbe families in samples with highest endogenous DNA content.
Fig. S5 | Principal Coordinate Analysis charts including all samples.
Fig. S6 | Eigenvalues of PCoA for 48 samples stored in controlled conditions.
Fig. S7 | LDA scores obtained for the four main clusters at the genus and species level.
Fig. S8 | DNA degradation patterns.
Fig. S9 | DNA damage plots for nuclear sequences of sample #524 from the Bouldnor Cliff.
Fig. S10 | DNA damage plots for chloroplast sequences of sample #524 from the Bouldnor Cliff.
Fig. S11 | DNA damage plots for nuclear sequences of sample #524 from the Bouldnor Cliff.
Fig. S12 | DNA damage plots for chloroplast sequences of sample #524 from the Bouldnor Cliff.
Fig. S13 | DNA damage plots for nuclear sequences of sample #524 from the Bouldnor Cliff.
Fig. S14 | Comparison of endogenous DNA content of ancient sapwood and heartwood.

Supplementary Text
Text S 1 BouldnorCliffDiscussion (docx)

Supplementary Tables
Table S 1 | Geographic data and ages (xlsx)
Table S 2 | PCR results (xlsx)
Table S 3 | Modern haplotypes (xlsx)
Table S 4 | Endogenous DNA content (xlsx)
Table S 5 | Shannon Indexes (xlsx)
Table S 6 | BouldnorCliff (xlsx)
Table 1 | Summary of ancient wood samples analyzed in this study (ka = kilo-annus). See Supplementary Table 1 for geographical coordinates, dating, storage history and thermal age calculations.

<table>
<thead>
<tr>
<th>Site</th>
<th>Age [ka]</th>
<th>Thermal Age [ka@10°C]</th>
<th>Collection Year</th>
<th>Taphonomical context</th>
<th>Type of wood</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amélie</td>
<td>4.5</td>
<td>7.9</td>
<td>2014*</td>
<td>Marine intertidal silts, foreshore</td>
<td>Piles exposed to tides</td>
<td>4</td>
</tr>
<tr>
<td>Bouldnor Cliff</td>
<td>8.0</td>
<td>7.0</td>
<td>2015*</td>
<td>Marine submerged peat</td>
<td>Logs, submerged forest</td>
<td>4</td>
</tr>
<tr>
<td>Champeaux</td>
<td>1.4</td>
<td>2.3</td>
<td>2015*</td>
<td>Marine intertidal silts, foreshore</td>
<td>Buried piles, fish weir</td>
<td>8</td>
</tr>
<tr>
<td>Conjux</td>
<td>2.8</td>
<td>5.2</td>
<td>2014*</td>
<td>Limy lake sediments</td>
<td>Piles, dwellings</td>
<td>5</td>
</tr>
<tr>
<td>Droßdorf</td>
<td>7.1</td>
<td>7.8</td>
<td>2014</td>
<td>Terrestrial glacio-fluvial sands</td>
<td>Planks, water well</td>
<td>12</td>
</tr>
<tr>
<td>Erstein 204</td>
<td>3.0</td>
<td>4.0</td>
<td>2014</td>
<td>Terrestrial sands and loess on limy bedrock</td>
<td>Planks, water well</td>
<td>5</td>
</tr>
<tr>
<td>Erstein 218</td>
<td>3.7</td>
<td>5.0</td>
<td>2014*</td>
<td>Terrestrial sands and loess on limy bedrock</td>
<td>Planks, water well</td>
<td>5</td>
</tr>
<tr>
<td>Eschenz Orkopf</td>
<td>5.2</td>
<td>5.4</td>
<td>2015*</td>
<td>Limy lake sediments</td>
<td>Piles, dwellings</td>
<td>9</td>
</tr>
<tr>
<td>Etang de Thau</td>
<td>2.8</td>
<td>6.3</td>
<td>2015*</td>
<td>Marine lagoonal sands</td>
<td>Piles</td>
<td>4</td>
</tr>
<tr>
<td>Greifswald</td>
<td>0.7</td>
<td>0.5</td>
<td>2014*</td>
<td>Organic wetland sediment</td>
<td>Planks</td>
<td>4</td>
</tr>
<tr>
<td>Hamburg</td>
<td>0.5</td>
<td>0.4</td>
<td>2015*</td>
<td>Organic wetland sediment</td>
<td>Planks</td>
<td>8</td>
</tr>
<tr>
<td>La Draga</td>
<td>7.1</td>
<td>10.0</td>
<td>2014</td>
<td>Peat and limy lake sediments</td>
<td>Piles, dwellings</td>
<td>9</td>
</tr>
<tr>
<td>Le Gurp</td>
<td>9.8</td>
<td>14.0</td>
<td>2014*</td>
<td>Coastal peat and sands</td>
<td>Logs, buried forest</td>
<td>4</td>
</tr>
<tr>
<td>Lillemer</td>
<td>5.8</td>
<td>13.8</td>
<td>2015*</td>
<td>Black peat</td>
<td>Logs, buried forest</td>
<td>4</td>
</tr>
<tr>
<td>Lucone</td>
<td>4.0</td>
<td>7.0</td>
<td>2014</td>
<td>Peat (gyttja)</td>
<td>Piles, dwellings</td>
<td>9</td>
</tr>
<tr>
<td>Ludwigshafen Holzplatz</td>
<td>5.2</td>
<td>5.4</td>
<td>2009-2014</td>
<td>Limy lake sediments</td>
<td>Piles, dwellings</td>
<td>12</td>
</tr>
<tr>
<td>Reims</td>
<td>1.7</td>
<td>1.5</td>
<td>2008</td>
<td>Clay</td>
<td>Piles</td>
<td>3</td>
</tr>
<tr>
<td>Rødby</td>
<td>6.9</td>
<td>5.9</td>
<td>2015*</td>
<td>Peat (gyttja)</td>
<td>Logs, buried forest</td>
<td>4</td>
</tr>
<tr>
<td>Saint Loup de Varennnes</td>
<td>1.9</td>
<td>2.5</td>
<td>2014</td>
<td>Clay</td>
<td>Planks</td>
<td>3</td>
</tr>
<tr>
<td>Spica</td>
<td>4.5</td>
<td>7.8</td>
<td>2010</td>
<td>Marsh and limy lake sediments</td>
<td>Piles, dwellings</td>
<td>7</td>
</tr>
<tr>
<td>Strojanove vode</td>
<td>5.6</td>
<td>8.7</td>
<td>2012</td>
<td>Marsh and limy lake sediments</td>
<td>Piles, dwellings</td>
<td>9</td>
</tr>
</tbody>
</table>
Table 2: Samples with highest final coverages achieved in this study. U: USER treated libraries.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Site</th>
<th>Retained sequences</th>
<th>Chloroplast coverage</th>
<th>Nuclear coverage [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>546</td>
<td>Unteruhldingen site 1</td>
<td>1.35E+07</td>
<td>47.98</td>
<td>14.34</td>
</tr>
<tr>
<td>382 U</td>
<td>Hamburg</td>
<td>1.36E+07</td>
<td>36.45</td>
<td>10.32</td>
</tr>
<tr>
<td>384 U</td>
<td>Hamburg</td>
<td>3.23E+07</td>
<td>18.64</td>
<td>10.13</td>
</tr>
<tr>
<td>049</td>
<td>Erstein 218</td>
<td>1.09E+07</td>
<td>17.03</td>
<td>1.47</td>
</tr>
<tr>
<td>423 U</td>
<td>Greifswald</td>
<td>1.53E+07</td>
<td>16.41</td>
<td>8.89</td>
</tr>
<tr>
<td>442 U</td>
<td>Eschenz Orkopf</td>
<td>4.71E+08</td>
<td>16.22</td>
<td>1.41</td>
</tr>
<tr>
<td>044 U</td>
<td>Erstein 218</td>
<td>2.57E+07</td>
<td>15.98</td>
<td>1.55</td>
</tr>
<tr>
<td>235 U</td>
<td>Sutz Lattringen Ruette</td>
<td>9.56E+07</td>
<td>15.47</td>
<td>6.05</td>
</tr>
<tr>
<td>045</td>
<td>Erstein 218</td>
<td>4.63E+07</td>
<td>15.17</td>
<td>2.85</td>
</tr>
<tr>
<td>233 U</td>
<td>Sutz Lattringen Ruette</td>
<td>7.16E+07</td>
<td>14.99</td>
<td>1.08</td>
</tr>
<tr>
<td>447 U</td>
<td>Eschenz Orkopf</td>
<td>2.10E+08</td>
<td>14.92</td>
<td>4.31</td>
</tr>
<tr>
<td>046 U</td>
<td>Erstein 218</td>
<td>2.34E+07</td>
<td>14.76</td>
<td>3.52</td>
</tr>
<tr>
<td>379</td>
<td>Hamburg</td>
<td>7.35E+06</td>
<td>12.79</td>
<td>5.74</td>
</tr>
<tr>
<td>321 U</td>
<td>Spica</td>
<td>2.31E+08</td>
<td>12.12</td>
<td>1.06</td>
</tr>
<tr>
<td>354 U</td>
<td>Strojanove vode</td>
<td>6.07E+07</td>
<td>12.01</td>
<td>1.14</td>
</tr>
<tr>
<td>548</td>
<td>Unteruhldingen site 1</td>
<td>1.01E+07</td>
<td>10.41</td>
<td>13.53</td>
</tr>
<tr>
<td>466 U</td>
<td>Champeaux</td>
<td>4.37E+07</td>
<td>10.34</td>
<td>2.64</td>
</tr>
<tr>
<td>469</td>
<td>Champeaux</td>
<td>1.85E+07</td>
<td>9.47</td>
<td>0.46</td>
</tr>
<tr>
<td>542</td>
<td>Erstein 204</td>
<td>2.55E+07</td>
<td>8.57</td>
<td>1.05</td>
</tr>
<tr>
<td>016</td>
<td>Eschenz Orkopf</td>
<td>5.48E+07</td>
<td>6.72</td>
<td>3.17</td>
</tr>
<tr>
<td>435</td>
<td>Hamburg</td>
<td>7.59E+06</td>
<td>5.86</td>
<td>2.72</td>
</tr>
<tr>
<td>523 U</td>
<td>Bouldner Cliff 2</td>
<td>3.08E+08</td>
<td>3.36</td>
<td>0.03</td>
</tr>
<tr>
<td>438</td>
<td>Hamburg</td>
<td>6.30E+06</td>
<td>2.37</td>
<td>2.41</td>
</tr>
</tbody>
</table>

* controlled storage: vacuum-packed and at 4°C.
<table>
<thead>
<tr>
<th>Sample</th>
<th>Location</th>
<th>Kilo-annus</th>
<th>Coefficient</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>048</td>
<td>Erstein 218</td>
<td>1.27E+07</td>
<td>2.34</td>
<td>0.34</td>
</tr>
<tr>
<td>543</td>
<td>Erstein 204</td>
<td>2.17E+07</td>
<td>2.28</td>
<td>0.97</td>
</tr>
<tr>
<td>552</td>
<td>Unteruhldingen site 2</td>
<td>9.15E+06</td>
<td>1.76</td>
<td>0.26</td>
</tr>
<tr>
<td>462</td>
<td>Champeaux</td>
<td>6.22E+06</td>
<td>1.53</td>
<td>0.09</td>
</tr>
<tr>
<td>381</td>
<td>Hamburg</td>
<td>6.68E+06</td>
<td>1.52</td>
<td>0.59</td>
</tr>
<tr>
<td>401</td>
<td>Ludwigshafen Holzplatz</td>
<td>1.48E+07</td>
<td>1.11</td>
<td>0.13</td>
</tr>
<tr>
<td>496</td>
<td>Etang de Thau</td>
<td>1.32E+07</td>
<td>1.09</td>
<td>0.08</td>
</tr>
</tbody>
</table>