New Phytologist

Research

*Methods*

# Improved recovery of ancient DNA from subfossil wood – application to the world's oldest Late Glacial pine forest

**Bertalan Lendvay[1], Martin Hartmann[1], Sabine Brodbeck[1], Daniel Nievergelt[1], Frederick Reinig[1], Stefan Zoller[2], Laura Parducci[3], Felix Gugerli[1], Ulf Büntgen[1,4,5] and Christoph Sperisen[1]**

[1]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland; [2]Genetic Diversity Centre, ETH Zurich, Universitätstrasse 16, CH-8092 Zurich, Switzerland; [3]Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, 75236 Uppsala, Sweden; [4]Department of Geography, University of Cambridge, Downing Place, Cambridge, CB2 3EN, UK; [5]Global Change Research Centre, Masaryk University, 613 00 Brno, Czech Republic

Author for correspondence:
*Christoph Sperisen*
*Tel: +41 44 739 2597*
*Email: christoph.sperisen@wsl.ch*

Received: *26 July 2017*
Accepted: *13 October 2017*

*New Phytologist* (2017)
**doi**: 10.1111/nph.14935

**Key words:** ancient DNA, DNA metabarcoding, experimental contamination, laser cleaning, post-mortem DNA damage, subfossil wood.

## Summary

- Ancient DNA from historical and subfossil wood has a great potential to provide new insights into the history of tree populations. However, its extraction and analysis have not become routine, mainly because contamination of the wood with modern plant material can complicate the verification of genetic information.
- Here, we used sapwood tissue from 22 subfossil pines that were growing *c.* 13 000 yr BP in Zurich, Switzerland. We developed and evaluated protocols to eliminate surface contamination, and we tested ancient DNA authenticity based on plastid DNA metabarcoding and the assessment of post-mortem DNA damage.
- A novel approach using laser irradiation coupled with bleaching and surface removal was most efficient in eliminating contaminating DNA. DNA metabarcoding confirmed which ancient DNA samples repeatedly amplified pine DNA and were free of exogenous plant taxa. Pine DNA sequences of these samples showed a high degree of cytosine to thymine mismatches, typical of post-mortem damage.
- Stringent decontamination of wood surfaces combined with DNA metabarcoding and assessment of post-mortem DNA damage allowed us to authenticate ancient DNA retrieved from the oldest Late Glacial pine forest. These techniques can be applied to any subfossil wood and are likely to improve the accessibility of relict wood for genome-scale ancient DNA studies.

## Introduction

Relict wood can be found in various settings, such as fluvial deposits, permafrost, peat bogs and archaeological sites. Extensive collections of well-preserved discs and cores from subfossil trunks now exist for many tree species and from numerous parts of the world. These archives represent an important source for developing annually resolved and absolutely dated tree-ring chronologies that have greatly contributed to the understanding of natural climate variability (Büntgen *et al.*, 2011), as well as associated causes and consequences (Büntgen *et al.*, 2016). Relict wood also offers ample opportunities for genetic studies. Genetic variation obtained from ancient wood samples can be linked to age information inferred from tree-ring chronologies, thus allowing direct insights into demographic processes (Parducci & Petit, 2004; Gugerli *et al.*, 2005). Furthermore, as tree-ring growth rates and stabile isotopic ratios measured in wood cellulose provide information about the environmental conditions in which a tree lived (Becker *et al.*, 1991), ancient wood offers the potential to reveal information on natural selection and its effect on genetic structures.

So far, only a few attempts have been made to extract and analyse ancient DNA (aDNA) from subfossil wood. Some studies have demonstrated the possibility of amplifying and sequencing aDNA from buried and waterlogged wood (Dumolin-Lapègue *et al.*, 1999; Tani *et al.*, 2003; Deguilloux *et al.*, 2006; Liepelt *et al.*, 2006; Gómez-Zeledón *et al.*, 2017). Although a few promising results have emerged, a multitude of data-related and methodological limitations have been encountered. In addition to the well-known fragmented nature of aDNA, wood chemical substances that inhibit enzymatic reactions (Rachmayanti *et al.*, 2009; Gómez-Zeledón *et al.*, 2017) and contamination with DNA from modern plant material appear to be critical shortcomings (Liepelt *et al.*, 2006).

Contamination with contemporary DNA can impose major biases on demographic and phylogenetic inferences from ancient material (Rawlence *et al.*, 2014; Brown & Barnes, 2015; Llamas *et al.*, 2017). Analysis of DNA from skeletal remains of ancient humans has provided strong evidence that contamination with modern DNA can occur during the excavation of samples, storage, anatomical examination and the final genetic analyses

(Sampietro *et al.*, 2006; Pilli *et al.*, 2013). Specific protocols have therefore been developed to identify and eliminate potential contamination, such as excavating the samples in protective clothing (Fortea *et al.*, 2008), bleaching and predigesting samples before DNA extraction (Barta *et al.*, 2013; Korlević *et al.*, 2015; Boessenkool *et al.*, 2017), using blocking primers (Gigli *et al.*, 2009) and separating endogenous aDNA molecules from modern DNA through *in vitro* (Gansauge & Meyer, 2014; Rohland *et al.*, 2015) or *in silico* approaches (Jónsson *et al.*, 2013; Skoglund *et al.*, 2014; Renaud *et al.*, 2015). The last two of these techniques rely on fixed DNA differences resulting from the deamination of cytosine to uracil, which accumulates post-mortem in aDNA and becomes apparent as cytosine to thymine changes in DNA sequences (Briggs *et al.*, 2007; Brotherton *et al.*, 2007; Sawyer *et al.*, 2012). For human or hominin remains, museum-stored samples are the most susceptible to contamination. Such samples have often been touched by researchers over decades, leaving DNA traces on their surface, which makes it crucial to completely eliminate the modern contamination or distinguish it from the authentic aDNA.

Samples of subfossil wood collections carry analogous contamination. Such material has been handled without awareness of avoiding contamination, and the wood surface may be polluted by pollen or cross-contaminated with wood powder from ancient or modern individuals during sample processing and/or sample storage (Deguilloux *et al.*, 2002). Most of the archived subfossil wood thus might be contaminated with a diversity of DNA from ancient and/or modern sources, from the same and/or other species. Further, limited effort has been made to systematically estimate the severity of external contamination of plant remains and to evaluate protocols for its elimination. The accurate separation of endogenous from contaminant DNA, however, is the most fundamental aspect of any genetic study, be it ancient or modern.

Here, we aimed to overcome some of the above-mentioned limitations. We took advantage of wood material from the world's oldest and largest Late Glacial pine forest. First, we developed an effective method for decontaminating wood surfaces. Second, we applied the developed technique to the subfossil pines and tested the DNA authenticity using DNA metabarcoding and post-mortem DNA damage. Finally, we used the authenticated DNA to taxonomically identify the Late Glacial pine trees.

## Materials and Methods

### Subfossil wood

A total of 253 subfossil pine (*Pinus* spp.) stumps were discovered at and excavated from a construction site in the Binz district of Zurich, Switzerland (Nievergelt *et al.*, 2014). This exceptional discovery near the foothills of the Uetliberg mountain (873 m above sea level (asl)) was made in 2013 by employees of the Swiss Federal Institute for Forest, Snow and Landscape Research WSL. Combined tree-ring and radiocarbon ($^{14}$C) dating indicate that the samples cover nearly 2000 yr between the Allerød and Preboreal at the transition from the Last Ice Age into the early Holocene (Nievergelt *et al.*, 2014). This discovery, the latest among several subfossil finds in the same region (Schaub *et al.*, 2005, 2008a,b; Kaiser *et al.*, 2012), provides not only the most material, but also the best-preserved wood. The age of the individual pine trees ranges from 41 to 506 yr, with all stumps containing sapwood and sometimes even bark (Fig. 1a,b). Wood anatomical analyses suggest their taxonomic identification as *Pinus* subgenus *Pinus*, which is represented in the Swiss flora by two tree-form species: Scots pine (*Pinus sylvestris* L.) and mountain pine (*Pinus uncinata* Ramond) (Lauber *et al.*, 2012).

### Surface decontamination

**Experimental contamination and its detection** As starting material for DNA analyses, blocks with an edge length of *c.* 5 cm were excised from the outermost parts of wood discs prepared from the subfossil pine stumps. We performed three experiments to assess the efficiency of different procedures to decontaminate experimentally contaminated wood surfaces using a total of 27 wood blocks from 11 stumps (Table 1). Experimental contamination was performed by rubbing the wood surfaces with pollen or saw powder derived from two Norway spruce (*Picea abies* (L.) H. Karst.) trees from the WSL arboretum. Before applying the different decontamination protocols, a contaminated control sample was taken from the surface of each wood block to affirm that the sample was in fact contaminated. After extracting DNA (for the procedure, see section 'DNA extraction' below) from both contaminated controls and cleaned samples, we monitored the Norway spruce contamination by amplifying a 137-bp fragment of the plastid *trn*F-*trn*L intergenic spacer using primers specific for Norway spruce (see Supporting Information Methods S1). Each DNA extract, as well as extraction blank controls and nontemplate PCR controls, was amplified in duplicate (for the procedure, see section 'Amplification of the plastid *trn*L P6 loop for DNA metabarcoding'). Four randomly chosen PCR products of both contaminated controls and decontaminated samples that yielded PCR products were directly sequenced using Sanger methodology to verify the presence of Norway spruce DNA.

**Expt 1: decontamination with bleach solution** In this experiment, we tested the cleaning efficiency of bleach treatment, which is known to effectively destruct DNA (Champlot *et al.*, 2010). We tested bleach against contamination from saw powder and pollen using eight wood blocks, that is, four contaminated with saw powder and four with pollen. For decontamination, the wood pieces were first rinsed in molecular grade water. Subsequently, the wood blocks were placed in separate plastic bags, which were then completely filled with *c.* 200 ml of 12% sodium hypochlorite (bleach). The bags were sealed and the wood was incubated for 15 min with occasional agitation. The incubation typically caused a *c.* 0.3 cm-depth soak-through of the wood, and the wood pieces were then left to dry overnight. In the following step, the outer *c.* 0.5 cm surface of the wood pieces was shaved off with a Micromot motorized carver (Proxxon, Föhren, Germany). In order to bleach-treat the already carved wood surface, the surface was sprayed with 12% bleach. Using a new, clean blade (see Methods S2 for details of the cleaning of laboratory


**Fig. 1** The subfossil 'Binz' pines (*Pinus*) and cleaning of their wood surfaces before DNA extraction: (a) a pine tree stump in the clay sediment at the excavation site; (b) disc cross-section of a tree stump; (c) laser cleaning of the surface of a wood block; (d) collection of wood shavings for DNA extraction using a carving machine.

utensils), we carved off the newly wetted surface until we reached dry wood and finally collected our samples (Fig. 1d). All working steps were conducted in a laboratory in which no DNA work had been carried out before.

**Expt 2: decontamination with laser irradiation** We evaluated the performance of three laser cleaning systems for eliminating experimental contamination with spruce pollen (Fig. 1c): a 20-W (low-power, CL 20 cleanLASER; Clean-Lasersysteme, Herzogenrath, Germany), a 300-W (mid-power, CL 300) and a 1000-W (high-power, CL 1000) laser. Each laser system was tested with three experimentally contaminated wood blocks. Laser treatment was carried out in a dedicated room at Achermann Unternehmungen (Kloten, Switzerland). The laser pulse frequency was set to 40 Hz for the CL 20 and CL 300 systems and to 30 Hz for the CL 1000 system. The entire surface of the wood blocks was scanned twice with the laser beam at a speed of *c.* 1 cm $s^{-1}$, with special care being taken not to re-contaminate the laser-treated surfaces. The cleaned wood pieces were transported to the laboratory in separate sealed bags and subsequently rinsed with molecular grade water to immobilize the carbonized wood powder and prevent it from spreading. To retrieve samples for DNA extraction, the surface was removed and shavings were cut with the carving machine as described in the previous section. Similar laser cleaning systems are manufactured by numerous companies, particularly low- and mid-power laser cleaning systems. High-power laser cleaning systems are produced by at least seven companies with manufacturing centres located in Asia, Europe and North America. Most companies offer an in-house service for test programmes and short-term applications, and some also offer on-site services.

**Expt 3: decontamination with high-power laser irradiation and bleaching combined** In this experiment, we tested 1000-W laser irradiation both alone and in combination with bleach treatment to decontaminate 10 wood blocks contaminated with pollen. First, samples for DNA extraction were taken directly after laser treatment as described for Expt 2. In a second round, the laser-treated wood blocks were bleach-cleaned and samples for DNA extraction were taken as described for Expt 1.

## Application of the improved decontamination procedure to 'Binz' pine wood and aDNA verification

**Sample preparation** We selected 12 wood discs originating from 11 pine stumps covering the period from *c.* 13 001 to 13 893 (mean tree ages) calendar years before present (cal. yr BP) (Table 2). From each of the 12 discs, two adjacent blocks were excised. From the first block, wood carvings were collected from the surface to estimate the level and nature of extraneous contamination. In the following text, these wood samples are named 'uncleaned wood surface'. The second wood block, named

**Table 1** Results of three decontamination treatments (Expts 1–3) applied to subfossil 'Binz' pine (*Pinus*) wood experimentally contaminated with wood powder and/or pollen of modern Norway spruce (*Picea abies*)

| Sample | Contaminant | PCR success *trn*F-*trn*L | |
|---|---|---|---|
| | | Contaminated wood | Bleach-cleaned wood |
| (a) Expt 1 | | | |
| BINZ082B | Wood powder | 2/2 | 0/2 |
| BINZ123Z | Wood powder | 2/2 | 2/2 |
| BINZ138C | Wood powder | 2/2 | 0/2 |
| BINZ175B | Wood powder | 2/2 | 1/2 |
| BINZ082B | Pollen | 2/2 | 2/2 |
| BINZ123Z | Pollen | 2/2 | 2/2 |
| BINZ138Z | Pollen | 2/2 | 2/2 |
| BINZ175B | Pollen | 2/2 | 2/2 |

| Sample | Contaminant | PCR success *trn*F-*trn*L | | | |
|---|---|---|---|---|---|
| | | Contaminated wood | Laser-cleaned wood | | |
| | | | 20 W | 300 W | 1000 W |
| (b) Expt 2 | | | | | |
| BINZ107C | Pollen | 2/2 | 2/2 | | |
| BINZ107C | Pollen | 2/2 | | 2/2 | |
| BINZ107C | Pollen | 2/2 | | | 0/2 |
| BINZ140B | Pollen | 2/2 | 2/2 | | |
| BINZ140B | Pollen | 2/2 | | 2/2 | |
| BINZ150B | Pollen | 2/2 | | | 0/2 |
| BINZ201C | Pollen | 2/2 | 2/2 | | |
| BINZ201C | Pollen | 2/2 | | 2/2 | |
| BINZ201C | Pollen | 2/2 | | | 2/2 |

| Sample | Contaminant | PCR success *trn*F-*trn*L | | |
|---|---|---|---|---|
| | | Contaminated wood | 1000-W laser-cleaned wood | |
| | | | − Bleach | + Bleach |
| (c) Expt 3 | | | | |
| BINZ146B | Pollen | 2/2 | 0/2 | 0/2 |
| BINZ146B | Pollen | 2/2 | 1/2 | 2/2 |
| BINZ146B | Pollen | 2/2 | 0/2 | 0/2 |
| BINZ146C | Pollen | 2/2 | 2/2 | 0/2 |
| BINZ146C | Pollen | 2/2 | 1/2 | 0/2 |
| BINZ146C | Pollen | 2/2 | 2/2 | 0/2 |
| BINZ152A | Pollen | 2/2 | 0/2 | 0/2 |
| BINZ152A | Pollen | 2/2 | 0/2 | 2/2 |
| BINZ158A | Pollen | 2/2 | 0/2 | 1/2 |
| BINZ158B | Pollen | 2/2 | 2/2 | 0/2 |

Removal of contamination was tested by PCR amplification of a 122-bp fragment of the plastid *trn*F-*trn*L intergenic region of Norway spruce. The success rate, as determined using gel electrophoresis, of two replicated PCR amplifications per DNA extract is given.

'cleaned wood', was used to perform surface cleaning with the 1000-W laser system, as described for Expt 2. The laser-treated wood pieces were transferred to the aDNA clean room laboratory available at WSL, where the most stringent rules in the research field were followed to avoid laboratory DNA contamination (for details, see Methods S2). Following laser treatment, sample

**Table 2** PCR amplification of the plastid *trn*L P6 loop from subfossil 'Binz' pine (*Pinus*) wood

| Sample | Mean age (cal. yr BP) | PCR success for *trn*L P6 loop | |
|---|---|---|---|
| | | Uncleaned wood surface | Cleaned wood |
| BINZ007A | 13 001 | 5/5 | 0/5 |
| BINZ039A | 13 893 | 5/5 | 5/5 |
| BINZ044B | 13 324 | 5/5 | 1/5 |
| BINZ046A | 13 015 | 5/5 | 5/5 |
| BINZ052A | 13 051 | 5/5 | 1/5 |
| BINZ056A | 13 553 | 5/5 | 1/5 |
| BINZ058A* | 13 813 | 5/5 | 5/5 |
| BINZ058B* | 13 813 | 5/5 | 5/5 |
| BINZ065A | 13 248 | 5/5 | 4/5 |
| BINZ131A | 13 336 | 5/5 | 0/5 |
| BINZ136B | 13 519 | 5/5 | 0/5 |
| BINZ152B | 13 345 | 5/5 | 2/5 |

The mean age of each pine sample and the success rate of five PCR attempts per DNA extract prepared before and after cleaning the wood surface are listed. Cleaning was performed using a combination of laser irradiation, bleach treatment and carving.
*Samples taken from different discs of the same tree trunk.

cleaning continued with two rounds of bleaching coupled with removal of their outer surfaces as described for Expt 1. The first round of bleach-cleaning took place in the vestibule of the aDNA laboratory (for details of the rooms, see Methods S2). The samples were then transferred to the sample preparation room, where the first round of carving, the second round of bleaching and the second round of carving took place for decontamination. The final carving step to collect material for DNA extraction was also performed there.

**DNA extraction** In the DNA extraction room of the aDNA laboratory, wood shavings obtained from the 12 'cleaned wood' blocks were ground to fine powder in 25-ml stainless steel grinding jars using two 10-mm-diameter grinding beads for 3 min at 30 Hz using a mixer mill (MM400; Retsch, Haan, Germany). From the recovered fine powder, *c.* 25 mg was collected in a 2-ml tube, and DNA was extracted using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). We modified the vendor's protocol by using 625 µl of lysis buffer containing 1 µg $ml^{-1}$ proteinase K (Qiagen). The amounts of wood powder and lysis buffer were empirically determined to match the maximum volume possible for the kit. The lysate was incubated under rotation for 12 h at room temperature and for an additional 3 h at 55°C. Following incubation, the lysate was supplemented with 200 µl of P3 buffer, kept at −20°C for 5 min, and centrifuged at 16 000 ***g*** for 5 min. The supernatant was centrifuged through the QIAshredder spin column at 16 000 ***g*** for 4 min and mixed with 975 µl of AW1 buffer. The remaining steps followed the Qiagen protocol. The DNA was eluted in 80 µl of elution buffer. Three extraction blank controls were processed alongside both the 12 'clean wood' and the 12 'uncleaned wood surface' samples. DNA from the 'uncleaned wood surface' samples was extracted as described for the 'cleaned wood' samples but was carried out in a modern

DNA laboratory. In addition, in order to estimate the rate of sequencing errors, we prepared a reference sample that contained the lowest possible post-mortem DNA damage: DNA was extracted from fresh needles of a Scots pine tree from the WSL arboretum (sample named 'modern needle control' in the following text).

**Enzyme inhibition of DNA extracts** To assess potential PCR inhibition in the DNA extracts, we monitored the effect of each DNA extract on a known amount of nonendogenous DNA that was amplified as an internal positive control (IPC) using quantitative polymerase chain reaction (qPCR). We applied the assay designed by King *et al.* (2009), targeting a 85-bp sequence derived from mRNA of the human β-2-microglobulin gene (*B2M*) with the PCR primers flanking two intron regions. We carried out the following three types of reactions: standard reaction containing $10^5$ copies of the IPC; reactions including the IPC and DNA extracts either undiluted or 1/10 diluted at 10% of the final PCR volume; and nontemplate controls. The three reaction types were performed in 20 μl with 1× PowerUp SYBR Green Master Mix (Thermo Fisher Scientific, Waltham, MA, USA) and 0.5 μM of each primer. Each type of reaction was run in three replicates. Cycling was carried out in an ABI 7500 Fast instrument (Thermo Fisher Scientific) as follows: 2 min of uracil-DNA glycosylase activation at 50°C, 2 min of polymerase activation at 95°C, followed by 45 cycles of 30 s of denaturation at 95°C, 30 s of annealing at 58°C and 30 s of extension at 72°C. Two measurements of PCR inhibition were obtained using the qPCR data: (1) the PCR efficiency ($E$) of each sample reaction, derived from the Hill slope of the amplification plot, was expressed as a percentage of the average PCR efficiency of three standard reactions; (2) shifts in the quantification cycle ($C_q$) values were used to evaluate the effects of extract dilution (see King *et al.*, 2009).

**Amplification of the plastid *trn*L P6 loop for DNA metabarcoding** To identify and compare the taxonomic identity of the DNA from the 'uncleaned wood surface' and 'cleaned wood' samples, we used the plastid *trn*L (UAA) P6 loop as a metabarcoding marker. Amplicons were generated with primers modified from Liepelt *et al.* (2006; see Methods S1); these primers are similar to the widely used *g* and *h* primers (Taberlet *et al.*, 2007) but amplify one base less at the 3′ end and were observed to produce substantially fewer primer-dimers (data not shown). The 5′ ends of the primers were tagged with the common sequences CS1 and CS2 (Fluidigm, South San Francisco, CA, USA) for multiplexing samples. The PCRs of the 'cleaned wood' samples were set up in the library-setup room of the aDNA laboratory with five technical replicates for each sample, three extraction blank controls, and three nontemplate PCR controls. PCR reactions were performed in 40 μl containing 1× Gold Buffer (Thermo Fisher Scientific), 25 mM $MgCl_2$, 50 mg $ml^{-1}$ UltraPure BSA (Thermo Fisher Scientific), 10 mM dNTPs, 10 μM of each primer, 1.6 U AmpliTaq Gold 360 DNA polymerase (Thermo Fisher Scientific) and 4 μl of DNA extract. PCRs were run in a Veriti thermocycler (Thermo Fisher Scientific) with the following cycling profile: 6 min of denaturation at 95°C followed by 45 cycles of 20 s of denaturation at 95°C, 35 s of annealing at 60°C and 45 s of elongation at 72°C. Final elongation took place for 10 min at 72°C. PCR products were visualized on ethidium bromide-stained 2% agarose gels. PCRs of the 'uncleaned wood surface' and 'modern needle control' samples were performed as described for the 'cleaned wood' samples but were run in the modern DNA laboratory.

**Illumina sequencing of amplicons** All samples that produced the slightest amplification bands on gels were selected for the downstream measurements. In addition, five apparently clear nontemplate PCR controls were selected, each of which was prepared alongside one of the five PCR replicates. Amplicons were sequenced at the Génome Québec Innovation Center at McGill University (Montréal, Canada) using the Access Array System (Fluidigm) for barcoding. Incorporation of the barcodes was performed in 15 cycles using the FastStart High Fidelity PCR System (Sigma-Aldrich, St Louis, MO, USA). For sequencing, sample volumes were adjusted according to the amount of PCR product, except for the nontemplate PCR controls, from which a fixed volume was loaded. Paired-end sequencing was performed on a MiSeq platform (Illumina, San Diego, CA, USA), with three blank samples run in parallel, using Reagent Kit v2, generating 150-bp reads per end. The samples were spiked with PhiX Control DNA (Illumina) to create a more diverse set of clusters.

**Analysis of the amplicon sequence data** Quality filtering of raw sequence reads and assignment to *trn*L reference sequences were performed using USEARCH v.9 (Edgar, 2010) as follows. Paired-end reads were merged using the USEARCH *fastq_mergepairs* algorithm (Edgar & Flyvbjerg, 2015), allowing no mismatch in the alignment. PhiX sequences were removed using the USEARCH *filter_phix* command. PCR primers were detected and trimmed using CUTADAPT (Martin, 2011), allowing no mismatch. Trimmed reads were quality-filtered using the USEARCH *fastq_filter* function with a maximum expected error threshold of 0.1 (Edgar & Flyvbjerg, 2015). All singleton sequences (sequence types represented by a single read in the entire data set) were omitted during subsequent analyses.

For the taxonomic assignment, the PlantAligDB on-line reference library (http://plantaligdb.portugene.com; C. Santos *et al.*, unpublished) was used. This library is a regularly updated, curated resource of barcoding DNA sequences based on data obtained from public databases. We retrieved the sequence data for the *trn*L intron, and prepared a broad taxonomic reference library after further curation of the data. The complete *c-d* and *g-h* *trn*L sequence alignments were collapsed and trimmed to match the P6 loop barcoding sequence flanked by our primers. We performed an additional data quality filtering step by eliminating all data entries that had ambiguous species identity (i.e. entries identified only to the genus level; environmental samples; unidentified hybrids; entries with *confer*, *species affinis* and *vidimus* remarks). The complete final data set contained 32 449 plant species representing 190 families (Table S1). Each entry in the reference library was assigned its taxonomic classification

(phylum, class, order, family, genus, species and subspecies) retrieved from the Taxonomy Database of the National Center for Biotechnology Information (NCBI) on 17.02.2017. Quality-filtered reads were aligned to the reference sequences using the USEARCH *usearch_global* algorithm (*maxaccepts 0, maxrejects 0, top_hits_only*) by iteratively increasing the number of allowed mismatches from 0 to 20. In this process, each query sequence was assigned to the most similar sequence(s) of the reference library. In cases of multiple hits, the consensus taxonomy of the reference sequences assigned to a particular query was calculated using the *classify.otu* algorithm implemented in MOTHUR with the threshold set to 100% (Schloss *et al.*, 2009). Raw sequences were deposited in the European Nucleotide Archive (ENA) under the accession number PRJEB21801.

In authentic aDNA, cytosine deamination (characteristic of post-mortem DNA damage) is expected to be prevalent, while in modern contamination it is presumed to be low or not present. We compared the rates of C>T and complementary G>A nucleotide misincorporations present in five groups of sequences: *Pinus* sequences of the 'modern needle control', 'uncleaned wood surface' and 'cleaned wood', and *Picea* (the most common contaminant taxon observed) sequences of the 'uncleaned wood surface' and 'cleaned wood'. In the *Pinus* 'cleaned wood' group, we considered only sequences of DNA extracts that amplified predominantly pine DNA to avoid bias introduced by contaminating modern pine DNA. By contrast, for all other groups, all pertaining PCR samples were considered, where *Pinus* or *Picea* sequences were represented by at least 1000 sequence reads, to retain the largest possible group sizes with meaningfully high sequence numbers. Each *Pinus*/*Picea* sequence was aligned with the *Pinus sylvestris* and *Picea abies* sequence in the reference data set, using the software NGS-EVAL (May *et al.*, 2015). We calculated the nucleotide C>T plus G>A misincorporation rate ($m$) according to Olivieri *et al.* (2010) as $m = M/Ln$, where $M$ is the sum of the C>T and G>A misincorporations, $L$ is the length of the sequence, and $n$ is the number of sequences. The misincorporation rate was multiplied by the T+A/C+G ratio to correct for the under-representation of the G and C bases of the *Pinus sylvestris* and *Picea abies* reference sequences. To ensure that the among-sample differences in C>T plus G>A misincorporation rates were not merely a product of the overall difference in the misincorporation rates, we additionally performed the calculations for the sum of misincorporation types other than C>T or G>A (hereafter referred to as noise). The differences in C>T plus G>A misincorporation rates and in noise among the five groups were compared with nonparametric Kruskal–Wallis tests, assuming that PCR replicates were independent. Dunn method post hoc tests were used with the Benjamini–Hochberg false discovery rate correction at a $P = 0.05$ significance limit.

### Species identification of 'Binz' pines

To distinguish between the two potential species, Scots pine and mountain pine, we took advantage of a species-specific single nucleotide polymorphism (SNP), located at position 150 of the plastid *trn*F-*trn*L intergenic spacer (Wachowiak *et al.*, 2006). This SNP was found to be diagnostic in more than 30 populations of Scots pine (carrying a cytosine) and species of the *Pinus mugo* Turra complex (carrying an adenine) from throughout Europe (Jasińska *et al.*, 2010; Wachowiak *et al.*, 2016), indicating that this SNP is fixed between the two taxa. This was further confirmed in two mixed stands of Scots pine and mountain pine located on Uetliberg mountain (data not shown). We designed PCR primers flanking the SNP to amplify a 109-bp fragment (Methods S1) to distinguish between the two taxa in the 'Binz' pine samples. We applied the marker to the aDNA samples that we considered high quality after screening with the metabarcoding approach for authentic aDNA. The DNA extracts were amplified in four replicates with the above primers, and each PCR product was directly sequenced using the Sanger method. Sequences were aligned in CLC MAIN WORKBENCH v.7.8.1 (Qiagen) software and inspected for the presence of C/A at the species-diagnostic base site.

## Results

### Development of protocols to decontaminate wood surfaces

The results of two PCR replicates indicated that bleach treatment combined with surface removal eliminated contamination from wood powder in two out of four samples and partly eliminated it in one sample, whereas none of the four pollen-contaminated samples were cleaned effectively with this method (Expt 1; Table 1a). Irradiation with the 20-W and 300-W lasers coupled with surface removal also did not eliminate pollen contamination from any of the three samples tested (Table 1b). By contrast, irradiation with the 1000-W laser removed pollen contamination from two out of three samples in Expt 2, and it cleaned five out of 10 samples completely and one sample partly in Expt 3 (Table 1b,c). The 1000-W laser in combination with bleaching and surface removal cleaned seven out of 10 samples completely and one sample partly in Expt 3 (Table 1c). Control samples taken from the contaminated wood surfaces and sequencing of PCR products confirmed that the wood was indeed effectively contaminated with spruce material in all cases (Table 1a–c). None of the total of 16 PCRs of extraction blank controls and 20 PCRs of nontemplate controls yielded amplification products.

### Screening 'Binz' pines for the presence of authentic aDNA

Using the improved cleaning procedure (1000-W laser irradiation coupled with bleaching and surface removal), we analysed a total of 12 wood samples from 11 'Binz' pines. DNA was isolated from both the 'uncleaned wood surface' and 'cleaned wood' and amplified for the plastid *trn*L P6 loop metabarcoding marker to assess surface contamination and the effectiveness of its removal by the cleaning procedure. Each DNA was amplified in five replicates and the amplicons were sequenced. For the 12 'uncleaned wood surface' samples, all PCRs yielded amplification products (Table 2). In the case of the 'cleaned wood', five out of the 12 samples yielded amplification products in more than two PCR replicates; an additional four samples showed amplification products in one or two PCR replicates. None of the five PCR attempts

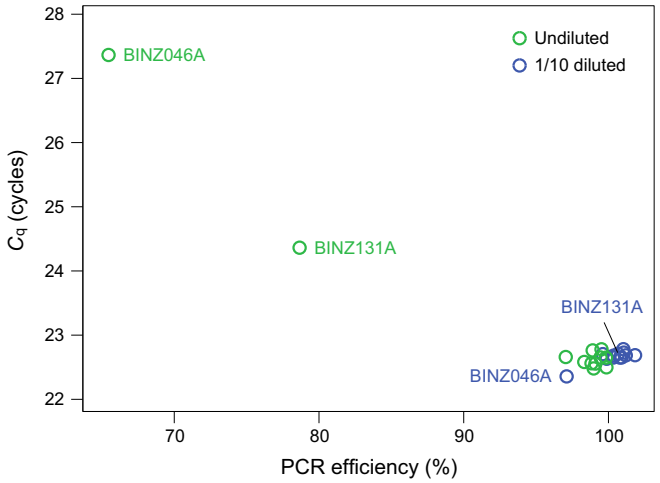



**Fig. 2** PCR inhibitory effect of undiluted and 1/10 diluted DNA extracts prepared from subfossil 'Binz' pine (*Pinus*) wood on the amplification of an internal control gene. $C_q$, quantification cycle; the PCR efficiency of each sample reaction is expressed as a percentage of the standard reaction without DNA extract. Samples showing PCR inhibition are labelled.

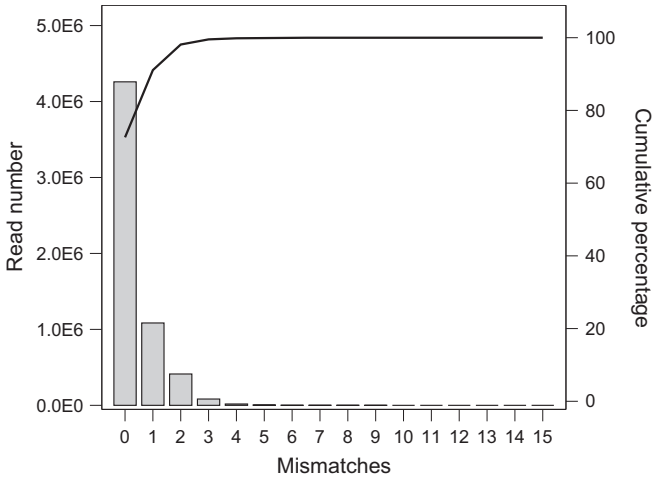



**Fig. 3** Distribution of the number of mismatches between plastid *trn*L P6 loop sequence reads of subfossil 'Binz' pine (*Pinus*) wood and the most similar sequences of the reference library. The left vertical axis shows the number of sequence reads assigned to the most similar reference sequence with 1–15 mismatches. The right vertical axis shows the cumulative percentage of taxonomically assigned sequence reads with maximal mismatches up to each mismatch value.

yielded any product in the remaining three samples (Table 2). Quantification of potential PCR inhibition revealed only two DNA extracts (BINZ046A and BINZ131A) that had a marked inhibitory effect on the amplification of the internal PCR control (IPC; Fig. 2). These extracts, in their undiluted form, caused a reduction in PCR efficiency by 34.5% and 21.3%, respectively, relative to the standard reaction. The 1/10 dilutions of the two extracts resulted in a shift of the quantification cycle $C_q$ by 5.0 and 1.7 cycles, respectively. For the remaining samples, the PCR efficiency of the diluted DNA extracts was only slightly higher compared with the undiluted DNA extracts.

The Illumina sequencing produced 9248 703 reads in total. Quality filtering and removal of singletons resulted in 5909 144 reads, defining 11 718 unique sequence types 7–138 bp in length. Of these reads, 98.9% could be taxonomically assigned, and the majority could be assigned to the genus level (Table S2). Of the assigned sequences, 99.5% aligned with the closest reference sequences with ≤ 3 mismatches (Fig. 3); the greatest number of mismatches was 15, found in a single sequence type represented by only two reads (Table S2).

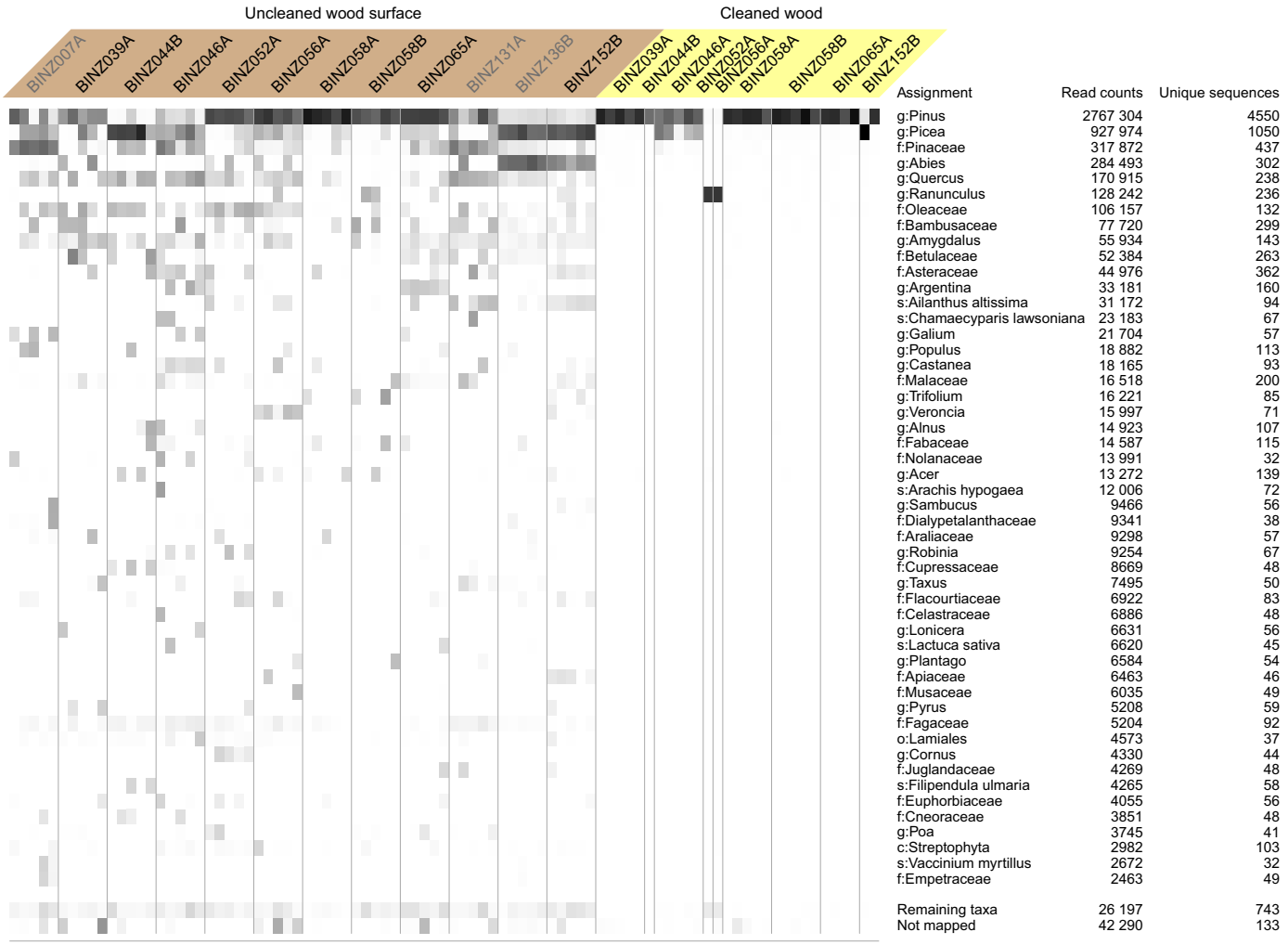
In the DNA of the 'uncleaned wood surface' samples, *Pinus* was the most prevalent taxon. However, in 46.7% of the PCRs the proportion of *Pinus* sequence reads was less than half of all sequence reads, including PCRs derived from samples that did not yield any PCR products after the wood surface was cleaned (Fig. 4). The most common contaminants were tree taxa, i.e. *Picea*, *Abies*, *Quercus*, *Larix/Pseudotsuga* (these last two taxa cannot be distinguished based on the *trn*L marker and their consensus taxon is determined as Pinaceae) and Oleaceae (probably *Fraxinus*), except for the herb taxon *Ranunculus*. In the 'cleaned wood', *Pinus* was prevalent, and only in 13.8% of the PCRs were less than half of the reads assigned to it. In four of the 'cleaned wood' samples (BINZ039A, BINZ058A, BINZ058B and BINZ065A), > 99.9% of the sequence reads of all PCR products were assigned to *Pinus* (Fig. 4; Table S2).

Out of the 32 PCRs of nontemplate controls and 30 PCRs of extraction blank controls, only three amplified visible PCR bands: two nontemplate PCR controls performed in the modern DNA laboratory for the 'uncleaned wood surface' samples, and an extraction blank control in the aDNA laboratory. The first and second of these controls showed a preponderance of *Picea* and *Pinus*, respectively, whereas *Acer* was predominant in the third control (Table S2).

To further authenticate the DNA of the 'cleaned wood' samples that almost exclusively revealed *Pinus* sequences, their C>T plus G>A misincorporation rates were compared with that of the 'modern needle control'. Additionally, we analysed the *Pinus* sequences of the 'uncleaned wood surface' samples and, where present, *Picea* sequences of 'uncleaned wood surface' and 'cleaned wood' samples to identify and characterize contamination (Fig. 5; Table S3). In the *Pinus* 'cleaned wood' samples, the rates were on average 40.7 times higher than in the *Pinus* 'modern needle control' and 2.1 times higher than in the *Pinus* 'uncleaned surface wood' samples, while the rates of the contaminant *Picea* sequences did not differ significantly from that of the *Pinus* 'modern needle control' (Table S3A,B). The noise misincorporation rates of all samples had the same order of magnitude as the C>T plus G>A misincorporation rate in the *Pinus* 'modern needle control', and they did not increase in the 'cleaned wood' samples (Table S3A,C). In summary, all the DNA extracts having >99.9% of their sequence reads assigned to *Pinus* after wood cleaning (BINZ039A, BINZ058A, BINZ058B and BINZ065A) showed patterns of misincorporation typical of post-mortem damage.

## Species identification of 'Binz' pines

To distinguish between Scots pine and mountain pine in the studied 'Binz' pine samples, we only considered the DNA




| Assignment | Read counts | Unique sequences |
|---|---|---|
| g:Pinus | 2767 304 | 4550 |
| g:Picea | 927 974 | 1050 |
| f:Pinaceae | 317 872 | 437 |
| g:Abies | 284 493 | 302 |
| g:Quercus | 170 915 | 238 |
| g:Ranunculus | 128 242 | 236 |
| f:Oleaceae | 106 157 | 132 |
| f:Bambusaceae | 77 720 | 299 |
| g:Amygdalus | 55 934 | 143 |
| f:Betulaceae | 52 384 | 263 |
| f:Asteraceae | 44 976 | 362 |
| g:Argentina | 33 181 | 160 |
| s:Ailanthus altissima | 31 172 | 94 |
| s:Chamaecyparis lawsoniana | 23 183 | 67 |
| g:Galium | 21 704 | 57 |
| g:Populus | 18 882 | 113 |
| g:Castanea | 18 165 | 93 |
| f:Malaceae | 16 518 | 200 |
| g:Trifolium | 16 221 | 85 |
| g:Veroncia | 15 997 | 71 |
| g:Alnus | 14 923 | 107 |
| f:Fabaceae | 14 587 | 115 |
| f:Nolanaceae | 13 991 | 32 |
| g:Acer | 13 272 | 139 |
| s:Arachis hypogaea | 12 006 | 72 |
| g:Sambucus | 9466 | 56 |
| f:Dialypetalanthaceae | 9341 | 38 |
| f:Araliaceae | 9298 | 57 |
| g:Robinia | 9254 | 67 |
| f:Cupressaceae | 8669 | 48 |
| g:Taxus | 7495 | 50 |
| f:Flacourtiaceae | 6922 | 83 |
| f:Celastraceae | 6886 | 48 |
| g:Lonicera | 6631 | 56 |
| s:Lactuca sativa | 6620 | 45 |
| g:Plantago | 6584 | 54 |
| f:Apiaceae | 6463 | 46 |
| f:Musaceae | 6035 | 49 |
| g:Pyrus | 5208 | 59 |
| f:Fagaceae | 5204 | 92 |
| o:Lamiales | 4573 | 37 |
| g:Cornus | 4330 | 44 |
| f:Juglandaceae | 4269 | 48 |
| s:Filipendula ulmaria | 4265 | 58 |
| f:Euphorbiaceae | 4055 | 56 |
| f:Cneoraceae | 3851 | 48 |
| g:Poa | 3745 | 41 |
| c:Streptophyta | 2982 | 103 |
| s:Vaccinium myrtillus | 2672 | 32 |
| f:Empetraceae | 2463 | 49 |
| Remaining taxa | 26 197 | 743 |
| Not mapped | 42 290 | 133 |

**Fig. 4** Taxa identified in the DNA obtained from 'uncleaned wood surface' (marked in brown) and 'cleaned wood' (marked in yellow) of subfossil 'Binz' pine (*Pinus*) samples using the plastid *trn*L P6 loop metabarcoding marker. The 50 most abundant taxa are displayed based on total sequence read counts. Shading displays square root-transformed read counts for each successful PCR amplification. The total number of unique sequence types assigned to a given taxon is indicated; letters in front of taxon names refer to species (s), genus (g) or family (f) level. Samples labelled in grey did not yield any PCR products after wood surface cleaning.

extracts that amplified predominantly pine DNA with high rates of C>T and G>A misincorporations in multiple PCRs (BINZ039A, BINZ058A, BINZ058B and BINZ065A). Using our species-distinctive marker, three of them, originating from BINZ039 and BINZ058, amplified successfully. At the diagnostic sequence position 27, the nucleotide was cytosine (as opposed to adenine) in all PCR replicates, indicating that these two individuals are Scots pines (Fig. S1).

## Discussion

Ancient DNA studies in plants have focused on observations of past genetic diversity and the assessment of selection during the domestication of cultivated plants (Gutaker & Burbano, 2017). However, genomic resources of tree species – most importantly draft genome sequences – are becoming accessible at an increasing rate. Moreover, novel molecular and analytical techniques, such as the cost-effective exome capture-based approach recently developed by Schmid *et al.* (2017) to obtain genome-scale data from subfossil leaf remains, are now available for wood samples. However, technical advancements specific to the physical structure and chemical characteristics of subfossil wood still need to be developed in order to obtain and confirm authentic aDNA. These include DNA extraction methods that increase aDNA yield and techniques to ascertain that the DNA recovered is free from modern and ancient DNA contamination. For plant remains, a particular risk of modern DNA contamination arises from pollen, especially from wind-pollinated tree species, which typically produce large amounts of pollen (Gugerli *et al.*, 2005). This risk is further increased by the sample handling processes, as a consequence of cross-contamination with modern or ancient wood powder from the same or other species (Deguilloux *et al.*, 2002). DNA contamination originating from the same species causes the greatest concern, as authentic and contaminant DNA often cannot be distinguished based on fixed genetic differences (Llamas *et al.*, 2017).

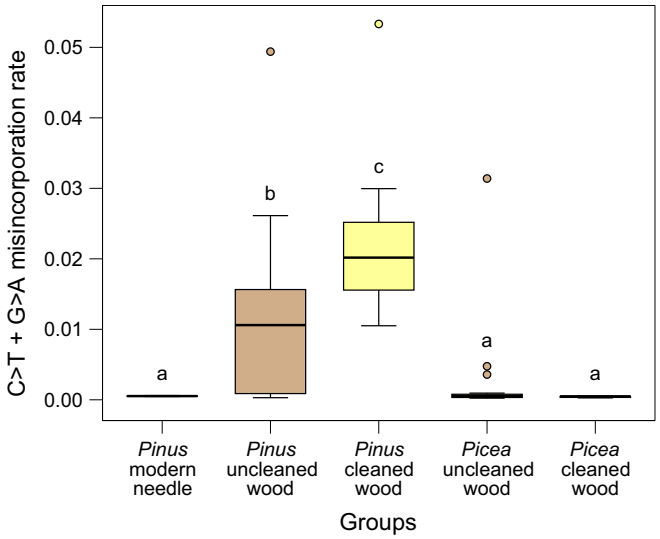



**Fig. 5** Rates of C>T plus G>A misincorporations in pine (*Pinus*) and spruce (*Picea*) sequences obtained from 'uncleaned wood surface' (marked in brown) and 'cleaned wood' (marked in yellow) of subfossil 'Binz' pine samples and in modern pine needles. Summary data are presented in box plot format: boxes indicate the interquartile range; the line drawn across the box represents the median; whiskers show the largest or smallest observations that fall within 1.5 times the box size; any observations outside these values are shown as individual circles. Misincorporation rates marked with different lowercase letters (a, b, c) differed significantly from each other. In the '*Pinus* cleaned wood' group, we considered only sequences derived from BINZ039A, BINZ058A, BINZ058B and BINZ065A that were free of contaminants (see Fig. 4).

The recently introduced molecular and computational techniques to filter authentic aDNA based on cytosine deamination patterns have primarily been developed for cases where the contaminating DNA is of recent origin, and therefore may be useless in cases where the contaminating DNA displays post-mortem damage patterns similar to that of the endogenous aDNA (Llamas *et al.*, 2017). In addition, such methods are wasteful because they disregard the aDNA fragments without deaminated cytosines (Gansauge & Meyer, 2014). Consequently, aDNA authentication remains reliant on thorough sample cleaning before DNA extraction (Korlević *et al.*, 2015). In the present study, we evaluated different decontamination techniques, and applied them to subfossil 'Binz' pine wood.

## Decontamination of subfossil wood surfaces

The primary method used to clean wood samples has been bleaching in combination with surface removal of both modern (Deguilloux *et al.*, 2002; Tnah *et al.*, 2012) and ancient wood (Wales *et al.*, 2014; Gómez-Zeledón *et al.*, 2017). However, its efficiency has never been tested experimentally. In this study, we assessed the effectiveness of cleaning procedures through experimental contamination followed by its elimination, and the subsequent measurement of the presence of residual contaminant DNA. Our results clearly show that even repeated treatment with high concentrations of bleach coupled with surface removal is ineffective at eliminating DNA contamination from pollen, and often also from wood powder, probably the two most common contaminants of wood samples. The higher resistance of pollen than wood powder can be explained by the former's larger initial DNA content and the pollen wall built from sporopollenin, one of the most durable organic substances. To complement bleaching and surface removal, we introduced laser treatment, a widely used industrial cleaning procedure for the removal of surface debris, contaminants and coatings (Büchter, 2012). Out of the three laser systems we tested, only the 1000-W laser was successful in eliminating experimentally added pollen. Treatment of the wood with this laser resulted in partial carbonization of the outermost wood surface. Temperatures in the inner part of the wood only increased slightly and did not exceed 40–45°C. In the majority of samples, we found the combination of laser treatment and thorough bleaching coupled with surface removal to be the most effective method to eliminate the DNA contamination. Unexpectedly, we found samples that were more contaminated after laser treatment and bleaching compared with laser treatment alone. This discrepancy may be attributable to microscopic cracks and to spreading of the remaining pollen on the wood surface during bleaching.

To validate our cleaning protocol in reality, we compared how the cleanness (i.e. the taxonomic assembly and C>T plus G>A misincorporation rates in the *trn*L P6 loop) of DNA extracts from the 'Binz' pines was changed by the cleaning procedure. Analysis of plant DNA recovered from the 'uncleaned wood surface' revealed considerable DNA contamination from wind-pollinated taxa that occur in the surroundings of the sites where samples were excavated and stored. This finding suggests that air-borne pollen was abundant and contaminated the wood samples. Comparison of the 'uncleaned wood surface' and 'cleaned wood' samples shows that the sequence read number of taxa other than *Pinus* dropped four-fold on average after cleaning. The C>T plus G>A misincorporation rates in the *Pinus* 'cleaned wood' samples increased significantly compared with those of the *Pinus* 'uncleaned wood surface' samples. Notably, the lower C>T plus G>A misincorporation rates of the 'uncleaned wood surface' compared with the 'cleaned wood' samples suggest that modern *Pinus* DNA contamination was present on the wood surface. *Picea* C>T plus G>A misincorporation rates were very similar to that of the modern *Pinus* sample, indicating that the contamination probably originated from pollen or wood powder of modern samples.

## Screening samples for aDNA authenticity

We considered a sample to contain authentic aDNA once the following three criteria were met: (1) DNA of the studied taxon must be repeatedly detectable, (2) the sample must be free of substantial amounts of contaminant DNA (i.e. other plant taxa), and (3) the potentially endogenous DNA must show an elevated C>T plus G>A misincorporation rate. We applied these criteria in a stepwise approach: in the first step, a DNA sample had to amplify multiple times with the *trn*L barcoding primers. If this criterion was fulfilled, then in the second step sequencing had to confirm a high prevalence of DNA of the studied taxon in each PCR amplicon. If this criterion was met, the C>T plus G>A misincorporation rate had to be markedly higher than in DNA of modern control samples.

In practice, we considered an aDNA extract of the 'Binz' pines to contain authentic aDNA if (1) the *trn*L metabarcoding fragment amplified in at least two out of five PCR replicates, (2) the prevalence of *Pinus* sequence reads was >90% in all *trn*L PCR replicates of the given DNA extract, and (3) the C>T plus G>A misincorporation rate was at least 10 times higher than that of the modern reference. Based on these criteria, we identified four out of 12 DNA extracts, BINZ039A, BINZ058A, BINZ058B and BINZ065A, as containing authentic aDNA. The results shown in Fig. 3 and summarized in Table S4 indicate that our results are very robust and our judgement of aDNA authenticity would not have been altered even if the criterion thresholds had been modified significantly.

### Ancient DNA quality

DNA retrieval from wood is challenging even in the case of modern wood, where DNA is usually highly degraded and may contain enzyme-inhibiting substances (Deguilloux *et al.*, 2002; Jiao *et al.*, 2012; Tnah *et al.*, 2012). Samples therefore often fail to provide enough DNA to ensure successful genetic analysis (Rachmayanti *et al.*, 2009). As an example, Schroeder *et al.* (2016) could successfully genotype only 58% of their analysed > 200 processed and treated timber samples. In our case, four out of 12 pine samples dated to > 13 000 yr BP yielded authentic *Pinus* aDNA for the 84-bp *trn*L region (including 39 bp of primers). Three out of these four samples also repeatedly amplified the 109-bp *trn*F-*trn*L region (including 44 bp of primers). We attribute the lower success rate of the latter region to its larger size. In the only study testing PCR inhibition in ancient wood DNA, Gómez-Zeledón *et al.* (2017) found significant enzyme inhibition, yet it did not affect the assays to the extent that it prevented successful genotyping. We observed strong effects of enzyme inhibition in only two of our DNA extracts, which, together with the results of Gómez-Zeledón *et al.* (2017), suggests that inhibition is not a major factor hindering genetic analyses of subfossil wood.

### Identification of the 'Binz' pines

The 'Binz' wood specimens could be identified as Scots pine or mountain pine based on wood anatomical analysis, similar to specimens collected from previous subfossil pine excavations from the area near Zurich (Kaiser, 1979; Schaub *et al.*, 2005, 2008a,b; Kaiser *et al.*, 2012). However, the exact identification of these two species based on wood morphology is not yet possible. Using our *trn*F-*trn*L genetic marker, we could firmly identify two 'Binz' pines as Scots pine.

Paleobotanical evidence attests that colonization by *Pinus* on the Swiss Plateau, which includes the Zurich area, started *c.* 14 000 yr BP (Ammann *et al.*, 1994, 2013), suggesting that the Binz site may have been colonized by an early pine population. Of the two pine species currently growing on Uetliberg, mountain pine is better adapted to harsh environmental conditions because it has a broader ecological niche, including greater cold and drought tolerance compared with Scots pine (Landolt *et al.*, 2010; Wachowiak *et al.*, 2015). Even though the ecological conditions during the Late Glacial period may have been more suitable for mountain pine, our results clearly indicate that Scots pine was present at the Binz location *c.* 13 900 yr BP.

### Conclusions

In this study, we demonstrated that subfossil wood, collected and archived without special precautions, is highly prone to contamination with modern plant material, and standard methods are inefficient in eliminating this contamination. Our approach of laser irradiation coupled with bleaching and surface removal efficiently removes exogenous DNA contamination derived from wood powder and pollen. Application of this procedure to material from >13 000-yr-old pine stumps, together with plastid DNA metabarcoding and assessment of post-mortem DNA damage, allowed us to identify several samples containing authentic aDNA. These techniques can be applied to any subfossil wood, and they are likely to improve the accessibility of relict wood to aDNA studies in order to elucidate population genetic, evolutionary and ecological processes. In the case of the 'Binz' pines, the wood material offers unique opportunities to study genetic variation through time in a tree population that lived in a period of rapid climatic change, including both periods of climate warming and periods of cooling. Studying the relationship between genetic changes and the local environmental conditions derived from tree-ring data has the potential to provide direct insights into the processes of local adaptation. Such studies involve genome-scale sequencing techniques such as shotgun sequencing and capture-based methods. We suggest metabarcoding with authentication criteria as a cost-effective method for selecting DNA samples free of contaminant DNA for the costlier downstream applications, particularly when working with subfossil wood or other ancient plant remains archived in collections.

## Acknowledgements

We thank Simon Dummermuth and Robin Winiger for laboratory assistance, Morten Rasmussen for technical advice, the staff of Génome Québec Innovation Center (Montréal, Canada) for carrying out the sequencing work, and the staff of Achermann Unternehmungen (Kloten, Switzerland) for assistance during the laser experiments. Melissa Dawes and two anonymous referees kindly commented on an earlier version of the manuscript. This study was financed by the WSL internal grant 'Evaluating the potential of the subfossil "Binz" pines for ancient DNA analyses'. U.B. received funding from the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Program I (NPU I grant no. LO1415), L.P. from the Swedish Research Council (grant 2013-D0568401), and F.R. from the Swiss National Science Foundation (SNF Grant 200021L_157187/1).

## Author contributions

C.S., U.B., F.G. and B.L. conceived the study. L.P. conducted preliminary DNA extraction and sequence analysis. D.N., F.R.

and U.B. prepared the subfossil wood samples and provided their age data. B.L., C.S. and S.B. performed the laboratory work, and B.L., M.H., C.S. and S.Z. analysed the data. B.L., C.S. and U.B. wrote the manuscript with support from the other co-authors. All authors read and approved the final version of the manuscript. F. G., U.B. and C.S. were joint senior authors.

## References

**Ammann B, Lotter AF, Eicher U, Gaillard M-J, Wohlfarth B, Haeberli W, Lister G, Maisch M, Niessen F, Schlüchter C. 1994.** The Würmian Late-glacial in Iowland Switzerland. *Journal of Quaternary Science* **9**: 119–125.

**Ammann B, van Leeuwen JFN, van der Knaap WO, Lischke H, Heiri O, Tinner W. 2013.** Vegetation responses to rapid warming and to minor climatic fluctuations during the Late-Glacial Interstadial (GI-1) at Gerzensee (Switzerland). *Palaeogeography, Palaeoclimatology, Palaeoecology* **391**: 40–59.

**Barta JL, Monroe C, Kemp BM. 2013.** Further evaluation of the efficacy of contamination removal from bone surfaces. *Forensic Science International* **231**: 340–348.

**Becker B, Kromer B, Trimborn P. 1991.** A stable-isotope tree-ring timescale of the Late Glacial/Holocene boundary. *Nature* **353**: 647–649.

**Boessenkool S, Hanghøj K, Nistelberger HM, Der Sarkissian C, Gondek AT, Orlando L, Barrett JH, Star B. 2017.** Combining bleach and mild pre-digestion improves ancient DNA recovery from bones. *Molecular Ecology Resources* **17**: 742–751.

**Briggs AW, Stenzel U, Johnson PLF, Green RE, Kelso J, Prüfer K, Meyer M, Krause J, Ronan MT, Lachmann M, Pääbo S. 2007.** Patterns of damage in genomic DNA sequences from a Neandertal. *Proceedings of the National Academy of Science, USA* **104**: 14616–14621.

**Brotherton P, Endicott P, Sanchez JJ, Beaumont M, Barnett R, Austin J, Cooper A. 2007.** Novel high-resolution characterization of ancient DNA reveals C > U-type base modification events as the sole cause of post mortem miscoding lesions. *Nucleic Acids Research* **35**: 5717–5728.

**Brown TA, Barnes IM. 2015.** The current and future applications of ancient DNA in Quaternary science. *Journal of Quaternary Science* **30**: 144–153.

**Büchter E. 2012.** A green way to clean – laser cleaning. *Laser Technik Journal* **9**: 36–38.

**Büntgen U, Myglan VS, Ljungqvist FC, McCormick M, Di Cosmo N, Sigl M, Jungclaus J, Wagner S, Krusic PJ, Esper J *et al.* 2016.** Cooling and societal change during the Late Antique Little Ice Age from 536 to around 660 AD. *Nature Geoscience* **9**: 231–236.

**Büntgen U, Tegel W, Nicolussi K, McCormick M, Frank D, Trouet V, Kaplan JO, Herzig F, Heussner K-U, Wanner H, Lüterbacher J, Esper J. 2011.** 2500 years of European climate variability and human susceptibility. *Science* **331**: 578–582.

**Champlot S, Berthelot C, Pruvost M, Bennett EA, Grange T, Geigl E-M. 2010.** An efficient multistrategy DNA decontamination procedure of PCR reagents for hypersensitive PCR applications. *PLoS ONE* **5**: e13042.

**Deguilloux M-F, Bertel L, Celant A, Pemonge MH, Sadori L, Magri D, Petit RJ. 2006.** Genetic analysis of archaeological wood remains: first results and prospects. *Journal of Archaeological Science* **33**: 1216–1227.

**Deguilloux M-F, Pemonge M-H, Petit RJ. 2002.** Novel perspectives in wood certification and forensics: dry wood as a source of DNA. *Proceedings of the Royal Society of London B: Biological Sciences* **269**: 1039–1046.

**Dumolin-Lapègue S, Pemonge M-H, Gielly L, Taberlet P, Petit RJ. 1999.** Amplification of oak DNA from ancient and modern wood. *Molecular Ecology* **8**: 2137–2140.

**Edgar RC. 2010.** Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

**Edgar RC, Flyvbjerg H. 2015.** Error filtering, pair assembly, and error correction for next-generation sequencing reads. *Bioinformatics* **31**: 3476–3482.

**Fortea J, de la Rasilla M, García-Tabernero A, Gigli E, Rosas A, Lalueza-Fox C. 2008.** Excavation protocol of bone remains for Neandertal DNA analysis in El Sidrón Cave (Asturias, Spain). *Journal of Human Evolution* **55**: 353–357.

**Gansauge M-T, Meyer M. 2014.** Selective enrichment of damaged DNA molecules for ancient genome sequencing. *Genome Research* **24**: 1543–1549.

**Gigli E, Rasmussen M, Civit S, Rosas A, de la Rasilla M, Fortea J, Gilbert MTP, Willerslev E, Lalueza-Fox C. 2009.** An improved PCR method for endogenous DNA retrieval in contaminated Neandertal samples based on the use of blocking primers. *Journal of Archaeological Science* **36**: 2676–2679.

**Gómez-Zeledón J, Grasse W, Runge F, Land A, Spring O. 2017.** TaqMan qPCR pushes boundaries for the analysis of millennial wood. *Journal of Archaeological Science* **79**: 53–61.

**Gugerli F, Parducci L, Petit RJ. 2005.** Ancient plant DNA: review and prospects. *New Phytologist* **166**: 409–418.

**Gutaker RM, Burbano HA. 2017.** Reinforcing plant evolutionary genomics using ancient DNA. *Current Opinion in Plant Biology* **36**: 38–45.

**Jasińska AK, Wachowiak W, Muchewicz E, Boratyńska K, Montserrat JM, Boratyński A. 2010.** Cryptic hybrids between *Pinus uncinata* and *P. sylvestris*. *Botanical Journal of the Linnean Society* **163**: 473–485.

**Jiao L, Yin Y, Xiao F, Sun Q, Song K, Jiang X. 2012.** Comparative analysis of two DNA extraction protocols from fresh and dried wood of *Cunninghamia lanceolata* (Taxodiaceae). *Iawa Journal* **33**: 441–456.

**Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013.** mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*: btt193.

**Kaiser F. 1979.** *Ein spätglazialer Wald im Dättnau bei Winterthur/Schweiz*. PhD thesis, University of Zurich, Switzerland.

**Kaiser KF, Friedrich M, Miramont C, Kromer B, Sgier M, Schaub M, Boeren I, Remmele S, Talamo S, Guibal F *et al.* 2012.** Challenging process to make the Lateglacial tree-ring chronologies from Europe absolute–an inventory. *Quaternary Science Reviews* **36**: 78–90.

**King CE, Debruyne R, Kuch M, Schwarz C, Poinar HN. 2009.** A quantitative approach to detect and overcome PCR inhibition in ancient DNA extracts. *BioTechniques* **47**: 941–949.

**Korlević P, Gerber T, Gansauge M-T, Hajdinjak M, Nagel S, Aximu-Petri A, Meyer M. 2015.** Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. *BioTechniques* **59**: 87–93.

**Landolt E, Bäumler B, Erhardt A, Hegg O, Klötzli F, Lämmler W, Nobis M, Rudmann-Maurer K, Schweingruber F, Theurillat J-P. 2010.** *Flora indicativa. Ecological indicator values and biological attributes of the flora of Switzerland and the Alps*. Bern, Switzerland: Haupt.

**Lauber K, Wagner G, Gygax A, eds. 2012.** *Flora Helvetica Bern*. Switzerland: Haupt.

**Liepelt S, Sperisen C, Deguilloux M-F, Petit RJ, Kissling R, Spencer M, de Beaulieu J-L, Taberlet P, Gielly L, Ziegenhagen B. 2006.** Authenticated DNA from ancient wood remains. *Annals of Botany* **98**: 1107–1111.

**Llamas B, Valverde G, Fehren-Schmitz L, Weyrich LS, Cooper A, Haak W. 2017.** From the field to the laboratory: controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. *STAR: Science & Technology of Archaeological Research* **3**: 1–14.

**Martin M. 2011.** Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* **17**: 10–12.

**May A, Abeln S, Buijs MJ, Heringa J, Crielaard W, Brandt BW. 2015.** NGS-eval: NGS Error analysis and novel sequence VAriant detection tooL. *Nucleic Acids Research* **43**: W301–W305.

**Nievergelt D, Hellmann L, Büntgen U. 2014.** Ein absolut datierbarer Schweizer Jahrringkalender bis ins Spätglazial? Einem Traum einen Schritt näher *GeoPanorama* **17**: 15–18.

**Olivieri C, Ermini L, Rizzi E, Corti G, Bonnal R, Luciani S, Marota I, De Bellis G, Rollo F. 2010.** Characterization of nucleotide misincorporation patterns in the iceman's mitochondrial DNA. *PLoS ONE* **5**: e8629.

**Parducci L, Petit RJ. 2004.** Ancient DNA–unlocking plants' fossil secrets. *New Phytologist* **161**: 335–339.

**Pilli E, Modi A, Serpico C, Achilli A, Lancioni H, Lippi B, Bertoldi F, Gelichi S, Lari M, Caramelli D. 2013.** Monitoring DNA contamination in handled vs. directly excavated ancient human skeletal remains. *PLoS ONE* **8**: e52524.

**Rachmayanti Y, Leinemann L, Gailing O, Finkeldey R. 2009.** DNA from processed and unprocessed wood: factors influencing the isolation success. *Forensic Science International: Genetics* **3**: 185–192.

**Rawlence NJ, Lowe DJ, Wood JR, Young JM, Churchman GJ, Huang Y-T, Cooper A. 2014.** Using palaeoenvironmental DNA to reconstruct past environments: progress and prospects. *Journal of Quaternary Science* **29**: 610–626.

**Renaud G, Slon V, Duggan AT, Kelso J. 2015.** Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biology* **16**: 224.

**Rohland N, Harney E, Mallick S, Nordenfelt S, Reich D. 2015.** Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**: 20130624.

**Sampietro ML, Gilbert MTP, Lao O, Caramelli D, Lari M, Bertranpetit J, Lalueza-Fox C. 2006.** Tracking down human contamination in ancient human teeth. *Molecular Biology and Evolution* **23**: 1801–1807.

**Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S. 2012.** Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS ONE* **7**: e34131.

**Schaub M, Büntgen U, Kaiser KF, Kromer B, Talamo S, Andersen KK, Rasmussen SO. 2008a.** Lateglacial environmental variability from Swiss tree rings. *Quaternary Science Reviews* **27**: 29–41.

**Schaub M, Kaiser KF, Frank DC, Büntgen U, Kromer B, Talamo S. 2008b.** Environmental change during the Allerød and Younger Dryas reconstructed from Swiss tree-ring data. *Boreas* **37**: 74–86.

**Schaub M, Kaiser KF, Kromer B, Talamo S. 2005.** Extension of the Swiss Lateglacial tree-ring chronologies. *Dendrochronologia* **23**: 11–18.

**Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ *et al.* 2009.** Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75**: 7537–7541.

**Schmid S, Genevest R, Gobet E, Suchan T, Sperisen C, Tinner W, Alvarez N. 2017.** HyRAD-X, a versatile method combining exome capture and RAD sequencing to extract genomic information from ancient DNA. *Methods in Ecology and Evolution* **8**: 1374–1388.

**Schroeder H, Cronn R, Yanbaev Y, Jennings T, Mader M, Degen B, Kersten B. 2016.** Development of molecular markers for determining continental origin of wood from white oaks (*Quercus* L. sect. *Quercus*). *PLoS ONE* **11**: e0158221.

**Skoglund P, Northoff BH, Shunkov MV, Derevianko AP, Pääbo S, Krause J, Jakobsson M. 2014.** Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proceedings of the National Academy of Sciences, USA* **111**: 2229–2234.

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

**Tani N, Tsumura Y, Sato H. 2003.** Nuclear gene sequences and DNA variation of *Cryptomeria japonica* samples from the postglacial period. *Molecular Ecology* **12**: 859–868.

**Tnah LH, Lee SL, Ng KKS, Bhassu S, Othman RY. 2012.** DNA extraction from dry wood of *Neobalanocarpus heimii* (Dipterocarpaceae) for forensic DNA profiling and timber tracking. *Wood Science and Technology* **46**: 813–825.

**Wachowiak W, Stephan BR, Schulze I, Prus-Głowacki W, Ziegenhagen B. 2006.** A critical evaluation of reproductive barriers between closely related species using DNA markers – a case study in *Pinus*. *Plant Systematics and Evolution* **257**: 1–8.

**Wachowiak W, Trivedi U, Perry A, Cavers S. 2015.** Comparative transcriptomics of a complex of four European pine species. *BMC Genomics* **16**: 234.

**Wachowiak W, Żukowska WB, Wójkiewicz B, Cavers S, Litkowiec M. 2016.** Hybridization in contact zone between temperate European pine species. *Tree Genetics & Genomes* **12**: 48.

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

## Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

**Fig. S1** Taxonomic identification of subfossil 'Binz' pines.

**Table S1** Plastid *trn*L P6 loop reference library

**Table S2** Taxonomic assignment of the plastid *trn*L P6 loop in DNA obtained from the subfossil 'Binz' pines

**Tables S3** Nucleotide misincorporation rates in DNA obtained from modern pine needles and from uncleaned and cleaned subfossil 'Binz' pine wood

**Table S4** Evaluation of the DNA authenticity of the 'Binz' pine DNA extracts

**Methods S1** Primer sequences.

**Methods S2** Laboratory practices and precautions against contamination.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.