New Phytologist (2016) doi: 10.1111/nph.14199

Lake sedimentary DNA accurately records 20th century introductions of exotic conifers in Scotland

Per Sjögren[1*], Mary E. Edwards[1,2], Ludovic Gielly[3,4], Catherine T. Langdon[2], Ian W. Croudace[5], Marie Kristine Føreid Merkel[1], Thierry Fonville[2], Inger Greve Alsos[1]

[1)] Tromsø University Museum, UiT – The Arctic University of Norway, Lars Thøringsvei 10, N-9037 Tromsø, Norway

[2)] Department of Geography and Environment, University of Southampton, SO17 1BJ, Southampton, United Kingdom

[3)] Laboratoire d'Ecologie Alpine, Université Grenoble Alpes, F-38000, Grenoble, France

[4)] Laboratoire d'Ecologie Alpine, CNRS, F-38000 Grenoble, France

[5)] Ocean and Earth Science, University of Southampton, National Oceanography Centre, SO14 3ZH, Southampton, United Kingdom

*Author for correspondence. Phone: +47 99562163, E-mail: per.sjoegren@uit.no

Summary

● Sedimentary DNA (sedDNA) has recently emerged as a new proxy for reconstructing past vegetation, but its taphonomy, source area, and representation biases need better assessment. We investigated how sedDNA in recent sediments of two small Scottish lakes reflects a major vegetation change, using well documented 20th century plantations of exotic conifers as an experimental system.

● We used next-generation sequencing to barcode sedDNA retrieved from sub-recent lake sediments. For comparison, pollen was analysed from the same samples.

● The sedDNA record contains 73 taxa (mainly genus or species), all but one of which are present in the study area. Pollen and sedDNA shared 35% of taxa, which partly reflects a difference in source area. More aquatic taxa were recorded in sedDNA, whereas taxa assumed to be of regional rather than local origin were only recorded as pollen.

● The chronology of the sediments and planting records are well aligned, and sedDNA of exotic conifers appears in high quantities with the establishment of plantations around the lakes. SedDNA recorded other changes in local vegetation that accompanied afforestation. There were no signs of DNA leaching in the sediments or DNA originating from pollen.

Keywords: environmental DNA (eDNA), lake sediments, metabarcoding, sedimentary DNA (sedDNA), vegetation change

1. Introduction

Sedimentary DNA (sedDNA) from lakes has potential as a tool for reconstructing past vegetation (Anderson-Carpenter *et al.*, 2011; Pedersen *et al.*, 2015; Thomsen and Willerslev, 2015). Even though studies of sedDNA show promising results (e.g. Willerslev *et al.*, 2003, 2007, 2014; Pansu *et al*., 2015; Alsos *et al*., 2016), to date there have been few investigations concerning important aspects of sedDNA taphonomy, such as i) the source area of plant DNA, ii) whether quantitative relationships exist between vegetation components and sedDNA, iii) how presence and absence are best defined when small quantities of sedDNA are present, and iv) whether sedDNA exhibits vertical mobility in a sedimentary column. Thus, further studies are necessary to reveal the full potential (and possible pitfalls) of using sedDNA as a proxy for vegetation composition (see Birks and Birks, 2015; Pedersen *et al.*, 2015; Thomsen and Willerslev, 2015; Barnes and Turner, 2016).

Previous comparisons show that sedDNA and plant macrofossil records are more floristically similar than those of sedDNA and pollen. As pollen includes a substantial regional component, a primarily local origin of sedDNA is indicated (Jørgensen *et al*., 2012; Parducci *et al*. 2013, 2015; Pedersen *et al*. 2013). For lakes, we hypothesize that terrestrial plant DNA, which can be within plant fragments or bound as molecules to clay or organic particles, could be derived from anywhere in a lake catchment, transported via streams, ground water or overland flow, or directly eroded from the shore (see Barnes and Turner, 2016). If pollen were a source of sedDNA, pollen transported over some distance could complicate the interpretation of local vegetation, as it can in palynological studies (see Sjögren et al., 2008). While DNA can be successfully extracted and amplified from pollen (Parducci *et al*., 2005; Keller *et al.*, 2015; Kraaijeveld *et al*., 2015; Bell *et al*., 2016), it is less clear if pollen actually contributes to the DNA recorded in lake sediments (see Birks *et al*., 2012; Parducci *et al*., 2012a,b). It is therefore important to address if, and to what degree, different sources contribute to the sedDNA record. Finally, given the small molecular size of extra-cellular DNA, there might also be a possibility of movement with or through sediment pore water. Sporadic downward movement of DNA has been recorded in terrestrial sediments (Haile *et al*., 2007; Andersen *et al.*, 2012), but it remains unclear if a similar phenomenon also occurs in lake sediments.

Only a few studies have addressed quantitative questions. Yoccoz *et al.* (2012) demonstrated a relationship between proportions of above-ground plant biomass and DNA abundance in soil, which, though noisy, suggests that the numbers of DNA copies (i.e., read numbers) may

contain quantitative information. A modest but positive correlation has been found between environmental fish DNA read numbers in lake water and fish biomass (Evans *et al*., 2016). When low concentrations of DNA and/or degraded DNA are expected, such as in ancient DNA samples, the tally of a taxon's occurrence in multiple PCR repeats is suggested as a better determinant of its presence or absence than total sequence read numbers (Ficetola *et al.*, 2015).

To evaluate how sedDNA records vegetation changes we took advantage of a "natural" experiment. During the 20th century, large areas of previously open heathland and rough grazing land in southern Scotland were planted with non-native conifers such as *Picea* sp. (spruce) and *Larix* sp. (larch). The native *Pinus sylvestris* (Scots pine) was also planted, augmenting a presence otherwise largely confined to dwindling semi-natural woodlands. While it is conceivable that DNA from ancient pine forests that resides in soils and sediments could be re-deposited in the lakes, DNA of the exotic taxa must be related to modern afforestation. The plantations are well documented by the Forestry Commission (Scotland), and it is possible to determine when and where they were established. A vertical sequence of sediment samples from a lake within a plantation should thus, hypothetically, provide a sedDNA record describing an abrupt and significant change from "conifer-free" open heathland-pasture communities to DNA assemblages dominated by coniferous taxa. The experimental setting mimics common applications of palaeoecology, namely, to determine when tree species colonised an area or how vegetation changed with land-use change. We analysed sedDNA in two short sediment cores from two Scottish lakes situated within well-documented conifer plantations. The sedDNA records were complemented with pollen data, which allowed a direct comparison with a standard, well-studied palaeoecological proxy. We had the following aims:

1) Investigate how the major changes in catchment vegetation composition are reflected in the quantity and quality of the sedDNA record. *Larix, Picea* and *Pinus* are distinguishable in our DNA reference library, and thus the expectation is that their DNA will appear in the sediment record at the depth that corresponds to first planting within the lake catchments—or soon after.
2) Compare the sedDNA record with the pollen record. This allows us to examine differences in sedDNA taxonomic resolution, source area and source dominance in relation to pollen. The contribution of non-local pollen can be documented and checked against the sedDNA record.

3) Assess whether there is vertical movement of sedDNA within lake sediments. The appearance of exotic conifer DNA in the sediment should be abrupt and coincident with the planting horizon, as determined by the isotopically estimated sediment age. Downward leaching would thus lead to an earlier than expected appearance of the DNA.

2. Material and Methods

2.1. Study sites and field sampling

The study area lies in Galloway and Dumfries, vice-county of Wigtownshire, southwest Scotland, a region that has undergone extensive afforestation in the 20th century (Fig. 1). We sampled two lakes within afforested catchments for short sediment cores: Loch of the Lowes (c. 3 ha, 55.004 N, -4.395 W, catchment c. 80 ha) and Spectacle Loch (c. 1.5 ha, 54.986 N, -4.579 W, catchment c. 230 ha). We used a Uwitec ™ gravity surface sampler (60 mm diameter) from an inflatable dinghy. The lakes were sampled in the deepest part (see Fig. 1), at 5.5 m in Loch of the Lowes and 7.8 m in Spectacle Loch. Approximately 35 cm of sediment was recovered in both lakes. Immediately upon sampling, excess water was removed and the core top stabilized with Zorbitrol ™ gel. Cores were returned to the Palaeoenvironmental Laboratory at the University of Southampton within two days and stored at 4°C, then subsequently frozen and shipped to the Tromsø University Museum.

Fig. 1

2.2. Historical record

Based on historical Ordnance Survey (OS) maps (EDINA, 2015a,b) and forestry plantation maps (Forestry Commission Scotland, planting maps: Loch of the Lowes, 2012; Spectacle Loch, 2000) the main changes in vegetation around the lakes can be reconstructed back to the mid-19th century. On the OS map of 1850, all land around Loch of the Lowes is marked as open (Fig. S1a), and maps from 1900 and 1910 show no change in vegetation. Plantating of

conifers, mostly *Pinus sylvestris* but also some *Picea sitchensis*, started in 1938. The bulk of this planting was in an adjacent catchment, and only a small area was planted in the Loch of the Lowes catchment. *P. sylvestris* from this planting is still present today. By 1950, conifers had been planted widely southeast of the lake. Large areas on the north-west side were planted in 1962-1965, and in 1970 practically all land around the lake was afforested. In the 2000s, large areas around the lake were cut down and replanted, primarily with *Picea sitchensis*. Trees stand close to the lake, but *Salix* spp., *Myrica gale* and *Calluna vulgaris* are prominent along the lake shore itself.

In 1840, Spectacle Loch was surrounded by rocky, heathy pastures, with marshland nearby (Fig. S1b). The nearest woodland according to the map was situated c. 1,5 km to the east. The situation was more or less the same in the 1950s. The area surrounding the lake (500 m radius) was planted with conifers between 1952 and 1962, with major plantings adjacent to the lake in 1960 and 1962. Afforestation continued with more planting in 1974-1976, and the entire area (>1 km radius) was covered with conifers by the late 1970s. In 2000, *Larix* forest and mixed *Picea-Pinus* forest dominated the surrounding vegetation (500 m radius), with *Picea sitchensis, Pinus contorta* and *Larix kaempferi* all growing adjacent to the lake. Since then large areas north, east and south of the lake have been felled. In recent years (2008-2013) minor patches of native hardwoods (*Alnus*, *Salix*, and *Betula*) were introduced, but not nearer to the lake than 300 m, and downstream of it. Around the shore there is abundant *Myrica* and *Calluna*, some *Salix* spp. and occasional trees of *Betula pubescens*.

2.3. Sub-sampling

Sub-sampling, extraction and amplification set-ups were performed in a laboratory dedicated for working with ancient DNA at Tromsø Museum. No PCR products have been present in this building. The cores were brought frozen to the laboratory and washed on the outside with chlorine solution. The frozen cores were sawed into 2 cm thick slices; either *in situ* in the plastic tube or when the plastic was removed. All sampling equipment was washed in chlorine solution between the sampling of each slice. Sub-sampling started from the bottom so if any extraneous material was moved, it was introduced upward, not downward. The sediment slices were put in zip-lock bags and kept frozen. Subsequently the slices were allowed to thaw partially, then the outer material of each slice was removed with sterile scalpels (changed after each cut). Material for pollen analysis was taken from the edge of the "cleaned" portion of the

sample and the central sediment piece put into a tube for DNA extraction. Loss-on-ignition analysis and radiometric dating were performed on the remaining material.

2.4. Core dating and age-depth assignments

Radiometric dating methods using ^{137}Cs (half-life of 30 years) and ^{210}Pb (half-life of 22.4 years) are well established (Croudace *et al*., 2012; Miller *et al*., 2014). ^{210}Pb and ^{137}Cs were determined at the National Oceanography Centre (Southampton) using Canberra well-type HPGe gamma-ray spectrometers (Canberra UK Ltd, Didcot). Gamma ray spectra were acquired for 100,000 s for each sample (c. 2 cm resolution) and processed using Fitzpeaks gamma deconvolving software (JF Computing, Stanford in the Vale, UK). The anthropogenic radionuclide ^{137}Cs shows three distinct datable features: the first appearance of ^{137}Cs (~1954), the 1963 "bomb maximum increase" and the 1986 Chernobyl event (e.g., Miller *et al*., 2014). ^{210}Pb activity reaches zero at an age of approximately 66 years (approximately 3 half-lives of ^{210}Pb), i.e. ~1950 AD (see Appleby and Oldfield, 1992; Croudace *et al*., 2012).

2.5. Loss-on-Ignition (LOI)

For LOI determination, Loch of the Lowes sediments from the same depths as the sedDNA samples were dried at 105°C overnight, placed in a furnace at 550°C for 4 hours, and weighed after each treatment. For Spectacle Loch, we used freeze-dried sediments representing 2-cm portions of the sediment core, which were weighed, ignited at 550°C for 2 hours and re-weighed (see Heiri *et al.* 2001). The LOI values were calculated as ((dry weight - weight after burning) / dry weight) *100.

2.6. Pollen analysis

Pollen samples were prepared using the acetolysis method (Berglund & Ralska-Jasiewczowa, 1986) and mounted in silicon oil for analysis. As our aim was to register the main changes in dominant taxa rather than to create a floristically detailed pollen diagram, an average of 180 grains were analysed per sample. Trends in DNA values and pollen percentages were compared for the more abundant terrestrial taxa (the planted conifers; *Pinus*, *Picea, Larix,* and all taxa making up 5% or more of the sum of terrestrial pollen or DNA reads; *Quercus*, *Alnus*,

Betula, Salicaceae, *Myrica*, *Calluna* and Poaceae). These taxa constitute on average 92.0 % of the pollen assemblage in both Loch of the Lowes and Spectacle Loch.

2.7. DNA Analysis

Thirteen sediment samples from Loch of the Lowes, twelve sediment samples from Spectacle Loch, eight extraction negative controls and four PCR negative controls were processed. The sediment sample size was 5–8 g. DNA was extracted using the PowerMax soil DNA isolation kit (MO BIO Laboratories, Carlsbad, CA, USA). The manufacturer´s instructions were followed, except that all centrifuge steps were done at 4800 g, and, at step four, the samples were alternately placed in a water bath at 65°C and vortexed for a total of 30-60 min and 10 min, respectively. All samples were finally recovered in 3 ml of elution buffer.

Using a previously described protocol (Alsos *et al.*, 2016), DNA was amplified and massively sequenced in parallel on a Illumina HiSeq 2500 platform, with the one change being that each sample (lake sediment as well as control) underwent six PCR repeats. Thus the total number of repeats was 150 for the sediment samples (25 x 6) and 72 for the controls (12 x 6). All samples were pooled before sequencing. The short and variable P6 loop region of the chloroplast *trn*L (UAA) intron (Taberlet *et al*., 2007) was used as diagnostic marker, amplified with universal primers *"g"* (5'-GGGCAATCCTGAGCCAA-3') and *"h"* (5'-CCATTGAGTCTCTGCACCTATC-3'). In order to segregate sequence reads bioinformatically and assign them to their relevant samples after high-throughput sequencing, unique eight bp-long tags (with at least five differences between tags) were added to the 5' end of each primer (modified from Binladen *et al.,* 2007 and Valentini *et al.,* 2009).

Following the same analysis protocol (Alsos *et al.,* 2016), next-generation sequence data were aligned (*illuminapairedend*), filtered (*ngsfilter*) and trimmed (*obiuniq*, *obigrep* and *obiclean*) using the OBITools software package (Boyer *et al*., 2016; http://metabarcoding.org/obitools/doc/index.html). Resulting barcodes were then assigned to taxa using the *ecotag* program (Yoccoz *et al.*, 2012) with both regional (Sønstebø *et al*., 2010; Willerslev *et al.*, 2014) and global (EMBL release r117) reference libraries, as was done by Alsos *et al.* (2016). After data filtering 11,171,750 reads of 17,145 unique sequences assigned to the 25 sediment samples were retained. A taxon was considered present in a repeat if it was a 100% match and was represented by 10 or more reads. For the full data-set, the limits for inclusion of taxa were as follows: two or more repeats in one sample, one or more repeats in

two adjacent samples, or a total of four or more repeats anywhere in the core. These taxa were considered to have a strong enough DNA signal to justify further analysis.

Exotic taxa were checked for potential PCR errors and tentatively identified with BLASTN 2.2.32+ (Zhang *et al.*, 2000; Morgulis *et al.*, 2008) to determine multiple/alternative taxon assignments (Table S1). Taxa assumed to be false positives based on their occurrences in negative controls were removed from further interpretation (see Table S2). All taxa identified in the DNA record after filtering were checked against the BSBI vice-county records for Wigtownshire (VC74; www.botanical keys.co.uk/flora/vccc/index.html).

3. Results

3.1. Sedimentation, chronology and loss-on-ignition

The cores consisted of dark brown detritus gyttja with no visible changes in the sedimentation regime, as far as it was possible to determine given the frozen and partly covered state when sub-sampled. The top parts of the cores from both lakes were observed in the field to be loose, flocculent sediment, and all samples became soupy after thawing.

The ^{210}Pb total and ^{137}Cs profiles for each lake are fairly similar (Fig. 2). This implies that the erosion and transport of radioactively labelled soil particles can be correlated between the lakes, i.e. that the magnitude of disturbing activities and/or environmental change are similar. In Loch of the Lowes, the three ^{137}Cs marker layers are seen in the data, while the Spectacle Loch profile the 1963 and 1986 events are present but more subtle; the 1963 and 1986 peaks are inferred from changes in the slope profiles (see Fig. 2).

Age-depth models for the two lakes were constructed and show good linear correlations based on the four radiometric marker layers (Fig. 3). The sampling date (2012) for the sediment surface is not included because the sediment-water interface is likely to have been disturbed and/or lost during sampling. Sampling and measurement uncertainties are evaluated to be +/- 1.0 cm for sample depth and +/- 5% for ^{137}Cs activities, based on gamma spectrometry counting statistics. The scale of these uncertainties does not have a significant impact on the age-depth model, and for the crucial period, 1940 to 1970, the ages are evaluated to be correct within a ± 5 year range.

Upland afforestation typically involves disturbance of the soil surface; we anticipated that this might be recorded as an increase in minerogenic input as seen in the LOI values (Fig. 2). In Loch of the Lowes a decrease in LOI begins after 25 cm depth and ends at 19 cm, and in Spectacle Loch a decline begins after 20 cm depth and ends at 13 cm. According to the age-depth model the declines end at 1963 and 1965 respectively.

Fig. 2 & 3

3.2. Temporal changes in sedDNA

The results from the sedDNA analyses are presented as the number of DNA repeats *vs* depth, which provides information on presence-absence and a semi-quantitative estimate of abundance (Fig. 4). Salicaceae, *Myrica gale*, and *Calluna vulgaris*, plus aquatic taxa (*Phragmites australis*, *Myriophyllum alterniflorum*, *Nuphar lutea*, Nymphaeaceae, *Potamogeton* and *Littorella uniflora*) are common throughout the sedDNA records of both lakes. Based on changes in the DNA values of the conifers and broadleaf trees, abundant taxa that show pronounced variations through the record, the Loch of the Lowes record was divided into two zones (LL1, LL2, visual inspection) and the Spectacle Loch record into three zones (SL0, SL1, SL2). Zones 1 and 2 are similar in both lakes and are interpreted together.

Zone 0 (c. <1910 – 1935) at Spectacle Loch pre-dates the record from Loch of the Lowes. The DNA values for broadleaf trees (*Quercus*, *Alnus*, *Betula* and Salicaceae) indicate local presence of these taxa. *Calluna* is also well represented. The vegetation around the lake was likely an open heathland with scattered trees and/or small woods.

Zone 1 (c. 1935 – 1960/65) is characterized by high DNA values of *Myrica*, *Calluna* and Poaceae and low values of all tree taxa, which suggests that open vegetation surrounded the lakes, probably *Calluna* heathland largely lacking local trees. This is in accordance with the historical maps depicting the local vegetation prior to plantation as heathy pasture.

Zone 2 (c. 1960/65 – 2010) starts with an abrupt increase in sedDNA of *Pinus*, *Picea* and *Larix*, and declines in *Myrica*, Poaceae and *Calluna*, which indicate a rapid transition of the local vegetation from open heathland to conifer forest. This is in accordance with the planting of conifers as we know it from the historical records and the present vegetation. The continued presence of shrubs and grasses in the sedDNA likely reflects lake-side vegetation,

small unplanted areas and/or relicts of heathland species in the undergrowth. In Loch of the Lowes sedDNA from *Quercus* and *Betula* appear, suggesting the local establishment of these taxa, likely facilitated by the cessation of grazing.

According to the radiometric dates, the boundary between zone 1 and 2, the main plantation event as detected in the sedDNA records (Fig. 4 and 5), occurred 1960 (~ ± 5 yr) in Loch of the Lowes and 1965 (~ ± 5 yr) in Spectacle Loch. Historical records show that early planting occurred at Loch of the Lowes in 1938, although the most prominent plantings occurred during 1962–1965. The decline in LOI dates to the early 1960's and thus corresponds with this main planting event (Fig. 5). At Spectacle Loch, planting occurred between 1952 and1962, with the main planting events in 1954 and 1960. The decline in LOI begins in the late 1950's, and conifer sedDNA values increase in the early 1960's. The dating of both lake records thus suggests that the time of the main planting aligns well with the appearance of exotic conifer sedDNA in the sediments, within the error of the methods.

Floristically, the DNA record is quite rich (73 unique native taxa), and most taxa are identified to genus or species. All taxa except *Hydrocotyle verticillata* are present in the vice-county. Pyreae includes *Malus domesticus* (apple), and even if the wild species occurs in the region it is treated here as a potential food contaminant. While it may be that some taxa present in vegetation surrounding the lake and as low reads/repeat numbers in the DNA data have been excluded, the fact that the DNA flora is ecologically appropriate indicates that the filtering procedures and thresholds for reads/repeats applied to the sequence data have effectively removed contaminants; the resultant taxa are likely to be true positives.

The record of herbaceous taxa at both lakes fits well with the change in land use during the 20th century. In zone LL1, a suite of taxa typical of moist moorland/rough grazing and/or the lakeside are present: *Pinguicula*, *Potentilla erecta, Ranunculus, Succisa pratensis, Plantago lanceolata.* These largely decline after the establishment of conifer plantations. A range of fern taxa, including *Blechnum, Dryopteris, Thelypteris* and *Phegopteris*, are far more common in the younger (plantation) zone, along with *Viola, Epilobium*, *Chamerion* and *Galium*. Similar changes occur in SL1, where species such as *Plantago lanceolata*, *Prunella vulgaris* and *Succisa* are more common pre-planting and ferns appear post-planting. While many taxa occur in both zones, there is a clear switch in the dominants, reflecting a change in the field layer of the vegetation around the lake as grazing was reduced and the plantations developed.

The sedDNA read data (Fig. 5.) show similar trajectories to the repeats for the conifer taxa and most of the broadleaf trees, but Salicaceae, *Myrica*, *Calluna* and Poaceae diverge. These

are all relatively common taxa growing along the lake shores. Their sedDNA repeat records indicate that they have been locally present throughout the period represented by the sediment column. The record of sedDNA reads, on the other hand, shows major fluctuations, which may relate to quantitative variations in the abundance of the taxa. For example, *Calluna* shows an abrupt decline in sedDNA reads from zone 1 to 2, i.e., when the open pasture-heathland is replaced with conifer plantations.

When compared with historical maps, the numbers of reads and the historical abundance (area planted) of the different conifer taxa follow the same order at Loch of the Lowes: *Picea* is most common, then *Pinus*, then *Larix.* At Spectacle Loch, all three conifers have similar read values in the sedDNA, with exception of the uppermost samples, in which *Picea* and *Larix* are higher than *Pinus* (Fig. 5). All three taxa have been planted adjacent to the lake, and, based on the maps, there is no clear difference in their abundance in the surrounding vegetation. Thus, the relative abundance of conifer sedDNA reads approximates the relative abundance (by area) of planted conifers at each site.

Fig. 4 & 5

3.3. SedDNA compared to the pollen signal

For full pollen diagrams and pollen concentrations of dominant taxa, see Fig. S2–5.

The pollen values of *Pinus* and *Picea* increase more or less simultaneously with the sedDNA or a little later; the lag is most clearly seen in the main rise of *Pinus* at Spectacle Loch (Fig. 5a). A delay in the pollen record compared with the planting event might be expected, as it takes time for the seedlings to reach maturity. A "tail" of *Pinus* pollen (c. 2–4 %) can be seen in both lakes, but this is likely a regional signal (but at Loch of the Lowes there was early *Pinus* planting). *Larix* is not detected by pollen analysis at these sites, which is likely an effect of its poor pollen dispersal capabilities and low pollen productivity (Sjögren *et al.,* 2008, 2010).

Pollen values of *Quercus, Alnus, Betula* and *Salix* give a different picture of the surrounding vegetation development than does the sedDNA. In contrast to the sedDNA, pollen values show continuous presence and no sharp changes. This may reflect masking of local change by regional input. Notably, pollen from *Quercus, Alnus, Betula* are abundant in some zones

where DNA is virtually absent (zone 1, and in Spectacle Loch also zone 2), which show that these relatively high levels of pollen do not provided a detectible DNA signal using the applied method.

Myrica and *Calluna* are well represented in both pollen and sedDNA. (The pollen type *Corylus/Myrica* likely includes a small proportion of *Corylus* pollen, however). Pollen and DNA reads follow roughly similar trends, but trends are different between sites (see Fig. 5c). Poaceae pollen values increase slightly from zone 1 to 2, but sedDNA reads decline. Poaceae includes many different species, and these may affect the pollen and sedDNA records differently. In particular, the emergent aquatic *Phragmites* is common and likely contributes high biomass and/or pollen load to the lake sediments, which could alter patterns of either or both proxies.

In the sedDNA data set (150 repeats) and pollen data set (~4500 terrestrial grains), we identified 97 individual native taxa (assuming lower taxonomic units correspond to higher taxonomic units). Of these, 39 (40%) were unique to sedDNA, 24 (25%) unique as pollen, and 34 (35%) recorded in both. Aquatic and spore taxa were better represented in sedDNA, with 14 (58%) unique to sedDNA, two (8%) unique to pollen and eight (33%) in common. It should be noted, however, that the number of identified taxa in both records depends on the particulars of the methods (e.g. number of DNA repeats per sample, representation of local flora in DNA reference library, number of repeats included to indicate presence, number of pollen grains counted, the pollen analyst's expertise and available time to identify rare pollen types, etc.). Thus, the above results are partly study-specific.

4. Discussion

Our goals were to assess how sedDNA records of plant taxa in lake sediments reflect known changes in vegetation cover and to compare the strengths and weaknesses of sedDNA in relation to pollen analysis. We were interested in the taxonomic clarity and temporal accuracy with which the sedDNA records afforestation, which indicate whether there has been downward leaching of sediments (or, conceivably, any laboratory contamination). The existence of interpretable quantitative trends and the degree of floristic detail available can both be assessed in relation to the pollen record. Finally, we can draw some conclusions about the source area for the sedDNA in these small lakes, based on observations of the catchments.

4.1. The planting period: temporal precision

We know from historical sources that there was a major and relatively abrupt change in the vegetation surrounding the two lakes during the mid-20th century. The sedDNA records major and abrupt change at the transition between zones 1 to 2, i.e., an increase in the conifers *Pinus*, *Picea* and *Larix*, and a decline in heathland taxa, especially *Calluna* (Fig. 4 and 5). At both sites, the sediment age-estimations show that planting dates align well with the rise of the exotic conifer taxa in the DNA records. SedDNA thus works as a temporally and floristically accurate proxy for major changes in local vegetation at both lakes. Sporadic, low values of conifer DNA do occur prior to zone 2, but never exceeds 1 repeat. It is not possible to distinguish this low levels of DNA detection from sequencing errors (Robasky *et al.*, 2014), and in the present study we only regard conifer DNA values ≥2 repeats as proof for local presence of conifer trees (true positives).

The dating shows that these increases occurred about 1960-1965, approximately at the time of the main planting events at the lakes. They are also associated with a slight LOI decline at both sites, which we interpret as reflecting soil disturbance due to mechanized activity, such as ripping of surface peat to improve drainage. This suggests that planting was largely complete before the DNA signal appeared in the lakes. At Loch of the Lowes, the relatively minor area planted in 1938 is only recorded as *Pinus* in a single sedDNA repeat, i.e. below what we considered certain detection level. Even if this planting was close to the lake it was not adjacent, and there are no major inlets on that side of the lake. Vegetation along the shores may thus have functioned to mask or filter the signal. It is likely that lake sedDNA primarily detect terrestrial vegetation that grow in the direct vicinity of the lake or inlet streams. The amount of biomass may also have played a role in when the sedDNA signal appears. Small seedlings may grow for several years prior to their DNA reaching the lake.

If DNA leaching occurred, we would expect DNA of exotic conifers prior to the time of plantation. The increases of conifer sedDNA between zones 1 and 2 are abrupt, and there is little in the way of "tails" of lower values in preceding samples, suggesting that leaching is not a concern. In contrast to what has been found in two studies of terrestrial sediments (Andersen *et al.*, 2012, Haile *et al*., 2007), our study does not indicate leaching of DNA in lake sediments. This is in accordance with other studies that also show that large organic molecules are immobilized in fine grained lake sediments (Smol, 2008).

4.2 Floristic detail and plant abundance

The sedDNA provides quite a rich flora (73 taxa from 25 samples/150 repeats), and all but one taxon are recorded from the area in which the lakes are located. The absence of false positives is due to stringent filtering and application of thresholds to the data (see methods), a procedure broadly applied to environmental DNA records (Ficetola *et al*., 2015; Pedersen *et al.*, 2015, Thomsen and Willerslev 2015; Alsos *et al*., 2016). The sedDNA records several local plant communities. As expected, aquatic macrophytes are well represented, as they can be expected to contribute a large biomass to the lake sediments. In addition, aquatic plant material is less exposed to ultraviolet radiation and temperature fluctuation than terrestrial plant material prior to deposition in the lake sediments, which may improve the DNA quality (see Strickler *et al.*, 2015). The main lake-shore taxa such as *Salix* and *Myrica* are also abundant.

In addition to the changes seen in the tree taxa, the patterns for dominant herbaceous taxa in zones 1 and 2 tend to differ (Fig. 4), which shows that the sedDNA reflects changes in the field layer as well as in the dominant trees. The moist heath/rough pasture components of *Calluna* and a range of forbs show declines as the landscape becomes afforested, while fern abundance and species richness increases. This pattern is particularly clear at Spectacle Loch (see Fig. 4). While we were not able to survey the vegetation of the two lake catchments exhaustively, the patterns of occurrence of the main taxa reflect well their current local presence or (near) absence, supporting our assumption that the sedDNA reflects the presence of local taxa, i.e., those within the hydrologic catchment, especially aquatic and lake-shore vegetation.

The variation in number of repeats can be related to the amount of DNA in the sediments, which, assuming all other factors being constant, in turn would be related to the abundance and proximity of the taxa in the surrounding vegetation (see Barnes and Turner 2016). For most of the common taxa in our study, the number repeats reached maximum values, and were therefore non- informative about variations in abundance. For example, at Loch of the Lowes, *Calluna* and Poaceae show little change in repeat values through the record. However, both taxa show an abrupt decline in sedDNA reads from zone 1 to 2, i.e., when the open pasture-heathland is replaced with conifer plantations (Fig. 5c). DNA read data could thus supplement DNA repeat data to detect shifts in abundance among common species.

4.3. SedDNA compared with pollen

Our results show, as expected, that sedDNA and pollen data sense the surrounding vegetation differently. Pollen of anemophilous taxa, such as conifer and broadleaf trees as well as most graminoids, has a large source area, and potentially it could to a great extent represent regional vegetation (see Sjögren *et al.*, 2008, 2010, 2015). DNA primarily represents local vegetation (near the lakes and within their hydrological catchments, Alsos *et al.*, 2016). Pollen values are affected by species-specific pollen productivity and dispersal properties (e.g., *Pinus vs Larix*), which are likely to affect DNA far less severely. In this test, using *Pinus*, *Picea* and *Larix*, we conclude that sedDNA gives a clearer signal of afforestation, particularly in the case of *Larix*. It also most likely gives a more realistic representation of the extent of broadleaved trees in the lake catchments in zones 1 and 2, particularly zone 2, when it is known that the catchments were afforested with conifers, yet there are still relatively high broadleaf pollen values. Our study strongly supports the idea that the DNA in lake sediments primarily originates from the hydrologic catchment, as suggested previously. Especially vegetation in or in direct proximity to the lake or inlet streams seem well represented.

In the present data set there are several examples of high pollen abundance and no DNA presence for angiosperm taxa (e.g., *Quercus, Betula* and *Alnus* in zone 1; Fig 5b). This suggests that sedDNA does not originate from pollen, at least for angiosperms. The reasons may be the relatively low biomass of pollen in sediment and/or the low copy numbers of cpDNA in pollen grains, as cpDNA is maternally inherited in most angiosperms (Nagata et al., 1999; Zhang *et al.*, 2003; Ellis *et al*., 2008).

Our study does not allow us to say conclusively that DNA of conifers cannot be derived from pollen. In conifers, cpDNA is paternally inherited. However, previous studies indicate that even when pollen of gymnosperms is recorded, it may be absent in DNA from the same sample (Parducci *et al.*, 2012b). In the present data-set *Pinus* pollen do occur during the pre-planting period with no clear correspondence in the sedDNA signal, although the pollen was only present in low abundances so the evidence is not as clear as with the angiosperms. In the case of Loch of the Lowes we know that *Pinus* was planted at the lake during zone 1 time, so neither can we dismiss the possibility that the weak DNA value that is present (1 repeat) is related to biomass derived from vegetative remains. An obvious follow-up experiment would be to assess DNA in a lake that has never had conifers in its catchment but is still close enough to a conifer stand to have high conifer pollen influx.

5. Conclusions

The sedDNA records from Loch of the Lowes and Spectacle Loch accurately depict the major 20th-century vegetation change—afforestation with conifers—known from historical data. The observed patterns are consistent with the sedDNA primarily reflecting local, i.e, within-catchment vegetation, as compared with the mix of local and regional vegetation portrayed by the pollen data. The level of floristic detail in the sedDNA is good and shows changes in minor taxa as well as in dominants. Aquatic taxa and taxa that dominate the biomass are especially well recorded, but forbs and cryptogams are also represented. The results of this study show that when carefully executed, sedDNA studies of lake sediments can provide reliable records, temporally and floristically, of local vegetation change. We recommend that future studies adopt the multiple-repeat approach, which increases the probability of detecting rare species and provides good opportunities of semi-quantification. For abundant taxa, the number of reads is more appropriate for semi-quantification. Further calibration of source areas and spatial biases in lake-sedDNA in a range of ecological settings is also desirable.

Acknowledgements

We thank Kim Davies and Helen MacKay for field assistance and Frédéric Boyer for help in raw-sequence data handling. We are grateful Rob Souter and Alistair Henderson of the Forestry Commission, Scotland, for help with site access and plantation data. Critical and insightful comments from three reviewers improved the final version. The work was supported by the Research Council of Norway (grant nos. 213692/F20 and 230617/E10 to Alsos).

Author Contribution

MEE and IGA planned and designed the research, MEE carried out the coring, LG, MKFM, PS, and IGA performed the DNA analysis, CTL performed the pollen analysis, IWC did the radiometric dating, TF compiled and interpreted the historical maps and planting patterns, and PS analysed and presented the data. PS wrote the manuscript with input from MEE, LG and IGA.

Conflict of Interest

Ludovic Gielly is one of the co-inventors of patents related to g-h primers and the subsequent use of the P6 loop of the chloroplast *trn*L (UAA) intron for plant identification using degraded template DNA. These patents only restrict commercial applications and have no impact on the use of this locus by academic researchers

References

Alsos IG, Sjögren P, Edwards ME, Landvik JY, Gielly L, Forwick M, Coissac E, Jakobsen LV, Føreid MK, Pedersen MW. 2016. Sedimentary ancient DNA from Lake Skartjørna, Svalbard: assessing the resilience of arctic flora to Holocene climate change. *The Holocene* **26**: 627–642.

Alsos IG, Coissac E, Edwards ME, Føreid MK, Gielly L, Sjögren P, Taberlet P, Yoccos, N. 2015. Plant DNA in sediments: to which degree do they represent the flora? *Genome* **58**: 163–303.

Anderson-Carpenter LL, McLachlan JS, Jackson ST, Kuch M, Lumibao CY, Poinar HN. 2011. Ancient DNA from lake sediments: Bridging the gap between paleoecology and genetics. *BMC Evolutionary Biology* **11**:30.

Appleby PG, Oldfield F. 1992. Applications of ^{210}Pb to sedimentation studies. In Ivanovich, M, Harmon RS, eds. *Uranium Series Disequilibrium*; New York, USA: Oxford University Press: 731–778.

Andersen K, Bird KL, Rasmussen M, Haile J, Breuning-Madsen H, Kjaer KH, Orlando L, Gilbert MT, Willerslev E. 2012. Meta-barcoding of 'dirt' DNA from soil reflects vertebrate biodiversity. *Molecular Ecology* **21**: 1966–1979.

Barnes MA, Turner CR. 2016. The ecology of environmental DNA and implications for conservation genetics. *Conservation genetics* **17**: 1-17.

Bell KL, Burgess KS, Okamoto KC, Aranda R, Brosi BJ. 2016. Review and future prospects for DNA barcoding methods in forensic palynology. *Forensic Science International: Genetics* **21**: 110–116**.**

Berglund BE, Ralska-Jasiewczowa M. 1986. Pollen analysis and pollen diagrams. In: Berglund BE, ed. *Handbook of Holocene Palaeoecology and Palaeohydrology*. Chichester, UK: Wiley & Sons: 455–484.

Binladen J, Gilbert MTP, Bollback JP, Panitz F, Bendixen C, Nielsen R, Willerslev E. 2007. The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. *PloS One* **2**: e19.

Birks HJB, Birks HH. 2015. How have studies of ancient DNA from sediments contributed to the reconstruction of Quaternary floras? *New Phytologist*, doi: 10.1111/nph.13657

Birks HH, Giesecke T, Hewitt GM, Tzedakis PC, Bakke J, Birks HJB. 2012. Comment on "Glacial Survival of Boreal Trees in Northern Scandinavia". *Science* **338**: 742.

Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E. 2016. OBITOOLS: a unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources* **16**: 176-182.

Croudace IW, Warwick PE, Morris JE. 2012. Evidence for the Preservation of Technogenic Tritiated Organic Compounds in an Estuarine Sedimentary Environment. *Environmental Science and Technology* **46**: 5704–5712.

EDINA. 2015a. "Historical maps of Loch of the Lowes", [PDF map], Scale 1:10,560, Kirkcudbrightshire Sheet XXIX, Ordnance Survey County Series 1st Edition 1849-1899, Published 1853, Kirkcudbrightshire Sheet XXXII.NE, Ordnance Survey County Series 1st Revision 1888-1914, Published 1910, National Grid 1:10560, Published 1957, National Grid 1:10000, Published 1978, [TIFF geospatial data], National Library of Scotland and Landmark Information Group. Using: EDINA Historic Digimap Service, http://digimap.edina.ac.uk/, Created: October 2015.

EDINA. 2015b. "Historical maps of Spectacle Loch", [PDF map], Scale 1:10,560, Wigtownshire Sheet VIII, Ordnance Survey County Series 1st Edition 1849-1899, Published 1847, Wigtownshire Sheet VIII.SE, Ordnance Survey County Series 1st Revision 1888-1914, Published 1909, National Grid 1:10560, Published 1957, National Grid 1:10000, Published 1978, [TIFF geospatial data], National Library of Scotland and Landmark Information Group. Using: EDINA Historic Digimap Service, http://digimap.edina.ac.uk/, Created: October 2015.

Ellis JR, Bentley KE, McCauley DE. 2008. Detection of rare paternal chloroplast inheritance in controlled crosses of the endangered sunflower Helianthus verticillatus. *Heredity* **100**: 574–580.

Evans NT, Olds BP, Renshaw MA, Turner CR, Li Y, Jerde CL, Mahon AR, Pfrender ME, Lamberti GA, Lodge DM. 2016. Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Molecular Ecology Resources* **16**: 29–41.

Ficetola GF, Pansu J, Bonin A, Coissac E, Giguet-Covex C, De Barba M, Gielly L, Lopes CM, Boyer, F., Pompanon, F, *et al.* 2015. Replication levels, false presences, and the estimation of presence / absence from eDNA metabarcoding data. *Molecular Ecology Resources* **15**: 543–556.

Haile J, Holdaway R, Oliver K, Bunce M, Gilbert MTP, Nielsen R, Munch K, Ho SYW, Shapiro B, Willerslev E. 2007. Ancient DNA chronology within sediment deposits:

are paleobiological reconstructions possible and is DNA leaching a factor? *Molecular Biology and Evolution* **24**: 982-989.

Heiri O, Lotter AF, Lemcke G. 2001. Loss on ignition as a method for estimating organic and carbonate content in sediments: reproducibility and comparability of results. *Journal of Paleolimnology* **25**: 101–110.

Jørgensen T, Haile J, Möller P, Andreev A, Boessenkool S, Rasmussen M, Kienast F, Coissac E, Taberlet P, Brochmann C, *et al.* 2012. A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. *Molecular Ecology* **21**: 1989–2003.

Keller A, Danner N, Grimmer G, Ankenbrand M, von der Ohe K, von der Ohe W, Rost S, Härtel S, Steffan-Dewenter I. 2015. Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. *Plant Biology* **17**: 558–566.

Kraaijeveld K, de Weger LA, Ventayol García M, Buermans H, Frank J, Hiemstra PS, den Dunnen JT. 2015. Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Molecular Ecology Resources* **15**: 8-16.

Miller H, Croudace IW, Bull JM, Cotterill CJ, Dix JK, Taylor RN. 2014. A 500 year sediment lake record of anthropogenic and natural inputs to Windermere (English Lake District) using double-spike lead isotopes, radiochronology and sediment microanalysis. *Environmental Science and Technology* **48**: 7254−7263.

Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schäffer AA. 2008. Database Indexing for Production MegaBLAST Searches. *Bioinformatics* **24**: 1757–1764.

Nagata N, Saito C, Sakai A, Kuroiwa H, Kuroiwa T. 1999. The selective increase or decrease of organellar DNA in generative cells just after pollen mitosis one controls cytoplasmic inheritance. *Planta* **209**: 53–65.

Pansu J, Giguet-Covex C, Ficetola GF, Gielly L, Boyer F, Zinger L, Arnaud F, Poulenard J, Taberlet P, Choler P. 2015. Reconstructing long-term human impacts on plant communities: an ecological approach based on lake sediment DNA. *Molecular Ecology* **24**: 1485–1498.

Parducci L, Edwards ME, Bennett KD, Alm T, Elverland E, Tollefsrud MM, Jørgensen T, Houmark-Nielsen M, Larsen NK, Kjær KH *et al*. 2012a. Response to Comment on "Glacial Survival of Boreal Trees in Northern Scandinavia". *Science* **338**: 742.

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

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

Parducci L, Matetovici I, Fontana SL, Sonia L, Bennett KD, Suyama Y, Haile J, Kjær KH, Larsen, NK, Drouzas AD *et al*. 2013. Molecular- and pollen-based vegetation analysis in lake sediments from central Scandinavia. *Molecular Ecology* **22**: 3511–3524.

Parducci L, Väliranta M, Salonen JS, Ronkainen T, Matetovici I, Fontana SL, Eskola T, Sarala P, Suyama Y. 2015. Proxy comparison in ancient peat sediments: pollen, macrofossil and plant DNA. *Philosophical Transactions of the Royal Society B* **370**: 20130382.

Pedersen MW, Ginolhac A, Orlando L, Olsen J, Andersen K, Holm J, Funder S, Willerslev E, Kjær KH. 2013. A comparative study of ancient environmental DNA to pollen and macrofossils from lake sediments reveals taxonomic overlap and additional plant taxa. *Quaternary Science Reviews* **75**: 161–168.

Pedersen MW, Overballe-Petersen S, Ermini L, Der Sarkissian C, Haile J, Hellstrom M, Spens J, Thomsen PF, Bohmann K, Cappellini E *et al.* 2015. Ancient and modern environmental DNA. *Philosophical Transactions of the Royal Society B* **370**: 20130383.

Robasky K., Lewis NE, Church GM. 2014. The role of replicates for error mitigation in next-generation sequencing. *Nature Reviews Genetics* **15**: 56–62.

Sjögren P, Connor SE, van der Knaap WO. 2010. The development of composite dispersal functions for estimating absolute pollen productivity in the Swiss Alps. *Vegetation History and Archaeobotany* **19**: 341–349.

Sjögren P, van der Knaap WO, Huusko A, van Leeuwen JFN. 2008. Pollen productivity, dispersal and representation factors of major tree taxa in the European Alps based on pollen trap results. *Review of Palaeobotany and Palynology* **152**: 200–210.

Sjögren P, van der Knaap WO, van Leeuwen JFN. 2015. Pollen dispersal properties of Poaceae and Cyperaceae: first estimates of their absolute pollen productivities. *Review of Palaeobotany and Palynology* **216**: 123–131.

Smol JP. 2008. *Pollution of lakes and rivers: a paleoenvironmental perspective*. Oxford, UK: Blackwell Publishing.

Sønstebø JH, Gielly L, Brysting AK, Elven R, Edwards M, Haile J, Willerslev E, Coissac E, Rioux D, Sannier J, *et al.* 2010. Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. *Molecular Ecology Resources* **10**: 1009–1018.

Strickler KM, Fremier AK, Goldberg CS. 2015. Quantifying effects of UV-B, temperature and pH on eDNA degradation in aquatic microcosms. *Biological Conservation* **183**: 85–92.

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

Thomsen PF, Willerslev E. 2015. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation* **183**: 4–18.

Valentini A, Miquel C, Nawaz MA, Bellemain E, Coissac E, Pompanon F, Gielly L, Cruaud C, Nascetti G, Winker P, *et al*. 2009. New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the trnL approach. *Molecular Ecology Resources* **9**: 51–60.

Willerslev E, Cappellini E, Boomsma W, Nielsen R, Hebsgaard MB, Brand TB, Hofreiter M, Bunce M, Poinar HN, Dahl-Jensen D, *et al*. 2007. Ancient biomolecules from deep ice cores reveal a forested southern Greenland. *Science* **317**: 111–114.

Willerslev E, Davison J, Moora M, Zobel M, Coissac E, Edwards ME, Lorenzen ED, Vestergård M, Gussarova G, Haile J, *et al.* 2014. Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature* **506**: 47–51.

Willerslev E, Hansen AJ, Binladen J, Brand TB, Gilbert MT, Shapiro B, Bunce M, Wiuf C, Gilichinsky DA, Cooper A. 2003. Diverse Plant and Animal Genetic Records from Holocene and Pleistocene Sediments. *Science* **300**: 791–795.

Yoccoz NG, Bråthen KA, Gielly L, Haile J, Edwards ME, Goslar T, von Stedingk H, Brysting AK, Coissac E, Pompanon F, *et al.* 2012. DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology* **21**: 3647–3655.

Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology* **7**: 203–214.

Zhang Q, Liu Y, Sodmergen. 2003. Examination of the Cytoplasmic DNA in Male Reproductive Cells to Determine the Potential for Cytoplasmic Inheritance in 295 Angiosperm Species. *Plant and Cell Physiology* **44**: 941–951.

Supporting material

Fig. S1. Historical maps.

Fig. S2. Pollen percentage diagram for Loch of the Lowes.

Fig. S3. Pollen percentage diagram for Spectacle Loch.

Fig. S4. Selected pollen and DNA data from Loch of the Lowes.

Fig. S5. Selected pollen and DNA data from Spectacle Loch.

Table S1. Re-assignment of exotic taxa.

Table S2. Negative controls.

Figure captions

Fig. 1. Overview and site maps. The bathymetric contours are approximate and delineate i) the 5-m depth at Loch of the Lowes and ii) 3, 5, and 7 m at Spectacle Loch. Cores were retrieved from the deepest part of each lake.

Fig. 2. Radiometric sediment age estimations and Loss-on-ignition (LOI). The middle minima in the LOI curves (shaded) are assumed to reflect increased erosion related to the main planting events. Dates are based on the first appearance of ^{210}Pb activity (~1950) and the three ^{137}Cs marker layers (the first appearance of ^{137}Cs ~1954, the 1963 "bomb maximum" increase and the 1986 Chernobyl event). Dating events within brackets are present but more subtle.

Fig. 3. Age-depth models based on linear interpolation of ^{210}Pb and ^{137}Cs marker layers. The water–mud interphase is likely disturbed and not included in the models.

Fig. 4. DNA repeat records (≥10 reads). Taxa with a minimum of two repeats in a single sample, a single repeat in two adjacent samples and/or ≥ four repeats in total are presented for each record. Taxa to the far right are aliens or common food plants (*Malus domesticus* included in Pyreae).

Fig. 5a-c. DNA and pollen results of selected taxa: 5a) coniferous tree; 5b) broadleaf trees; and 5c) shrubs and grasses. Only planted conifers and terrestrial taxa with ≥5% of the pollen or DNA repeat records are presented. Pollen is presented as a percentage of the terrestrial pollen sum. Conifer cover %: Percentage cover of conifer plantations within 500 m of the catchment in relation to modern planted areas as estimated from available historical maps. Note differences in scales on both x and y axes. The youngest part of the age scale (>1986) is extrapolated from the age-depth model. *: *Salix* and *Myrica/Corylus*-type pollen in 5b and 5c, respectively.

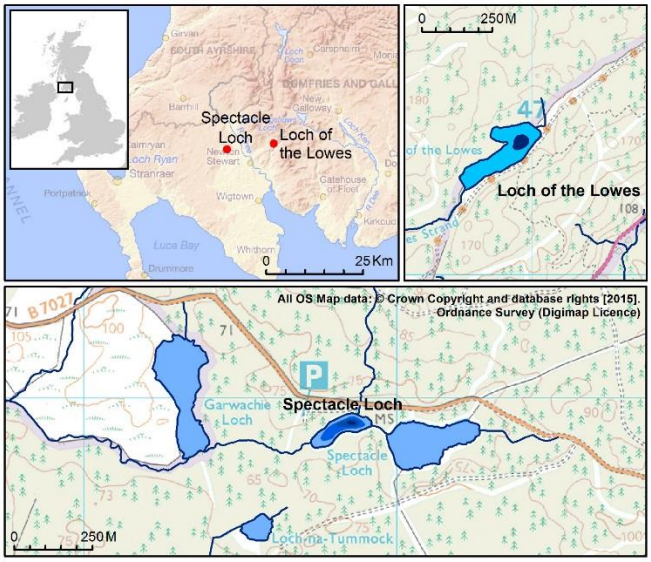
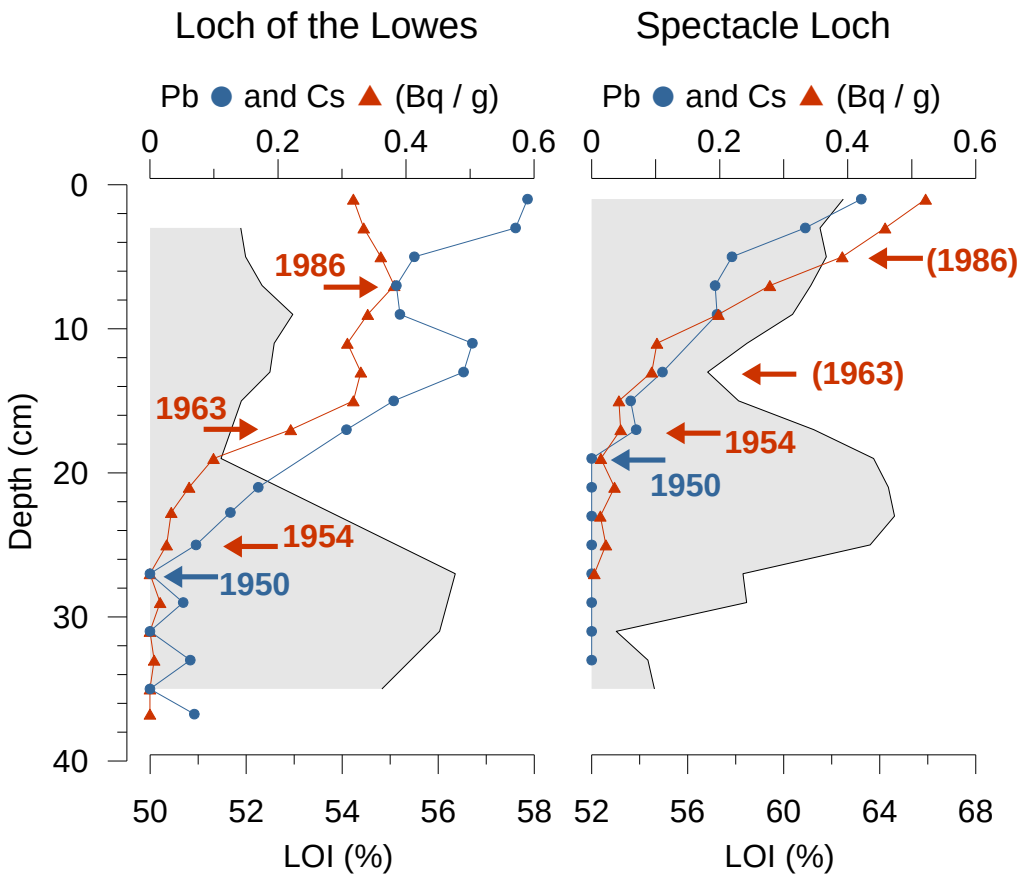
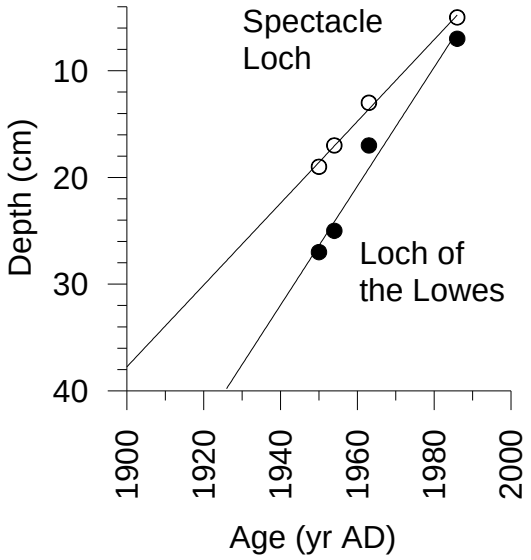


Fig. 1. Overview and site maps.


Loch of the Lowes
Spectacle Loch
Pb ● and Cs ▲ (Bq / g)
0
0.2
0.4
0.6
Depth (cm)
0
10
20
30
40
1986
1963
1954
1950
(1986)
(1963)
1954
1950
50
52
54
56
58
52
56
60
64
68
LOI (%)
LOI (%)


Spectacle
Loch
Loch of
the Lowes
Depth (cm)
10
20
30
40
1900
1920
1940
1960
1980
2000
Age (yr AD)

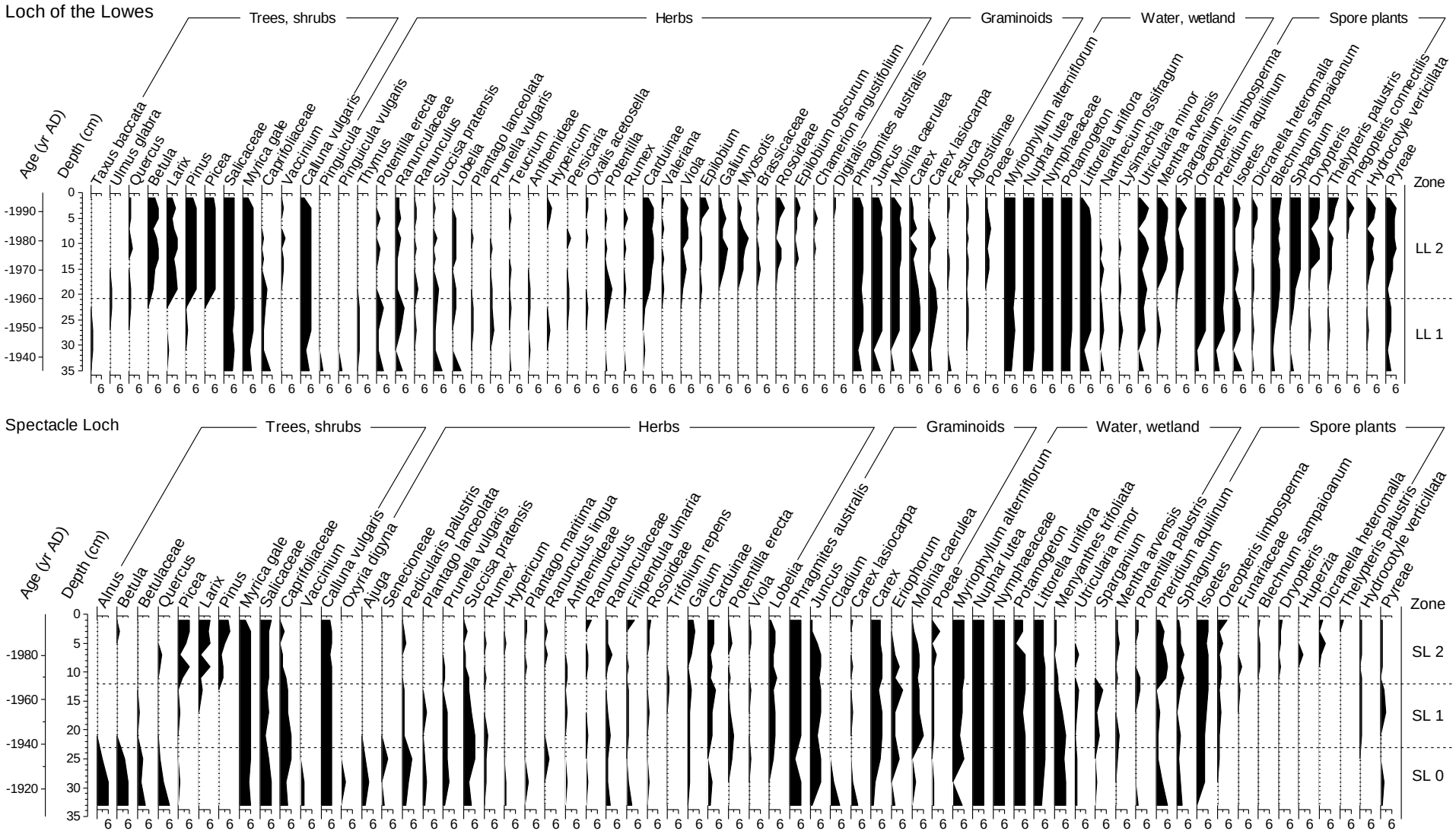


Fig. 4 DNA repeat records (≥10 reads). Taxa with minimum two repeats in a single sample, a single repeat in two adjacent samples and/or ≥ four repeats in total are presented for each record. Taxa to the far right are alien or common food plants (*Malus domesticus* included in Pyreae).

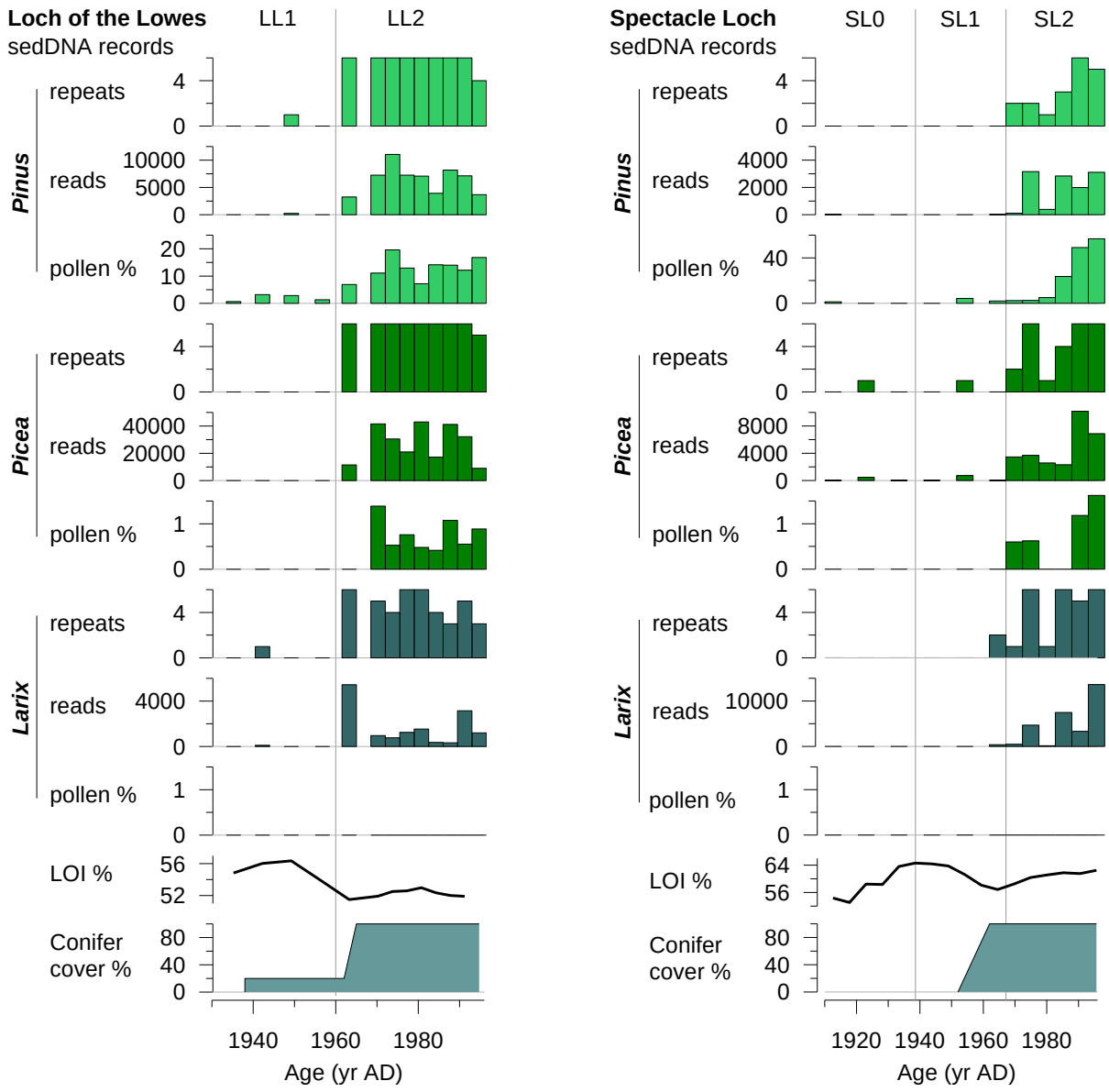


Fig. 5a. Coniferous trees

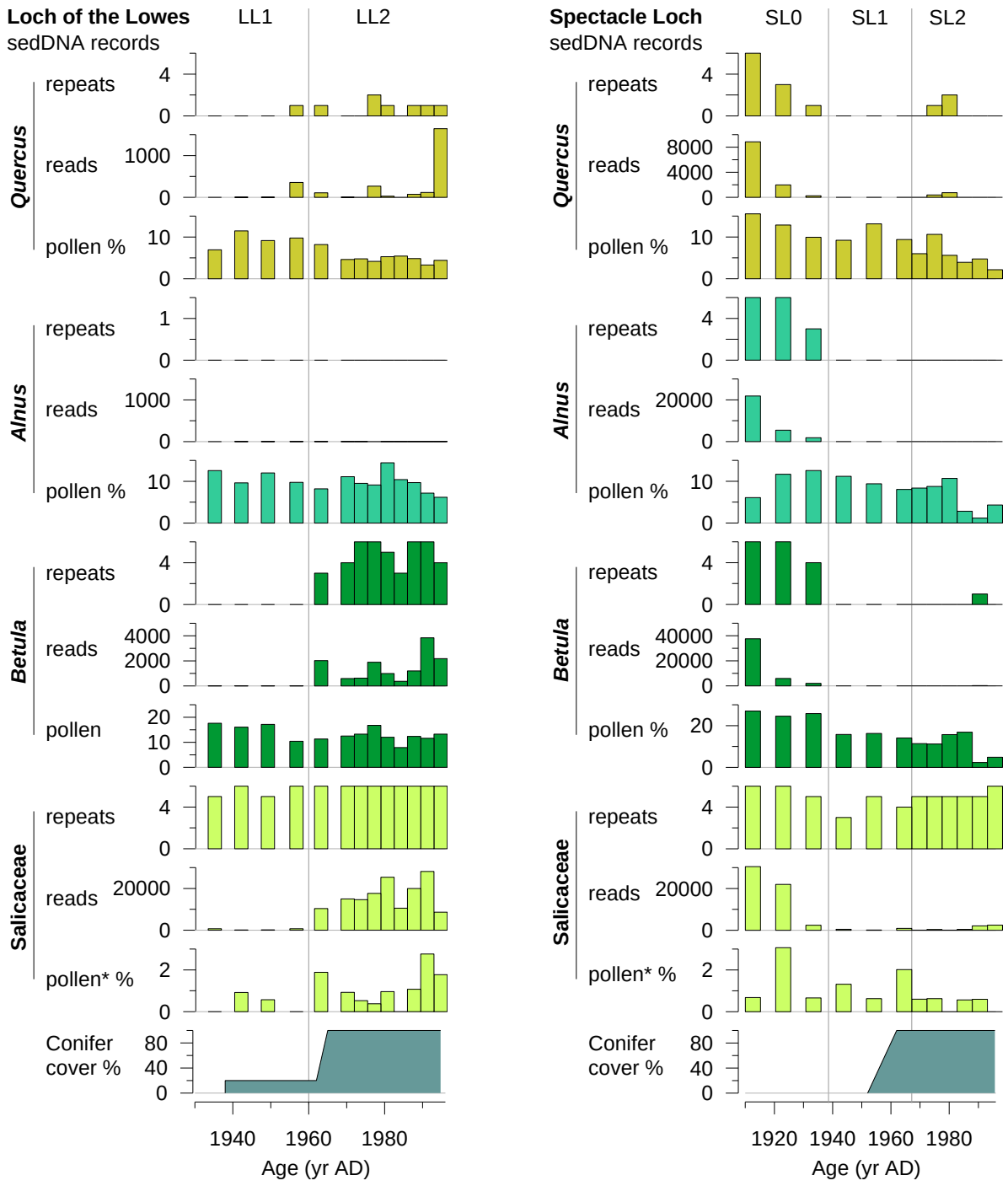


Fig. 5b Broadleaf trees.

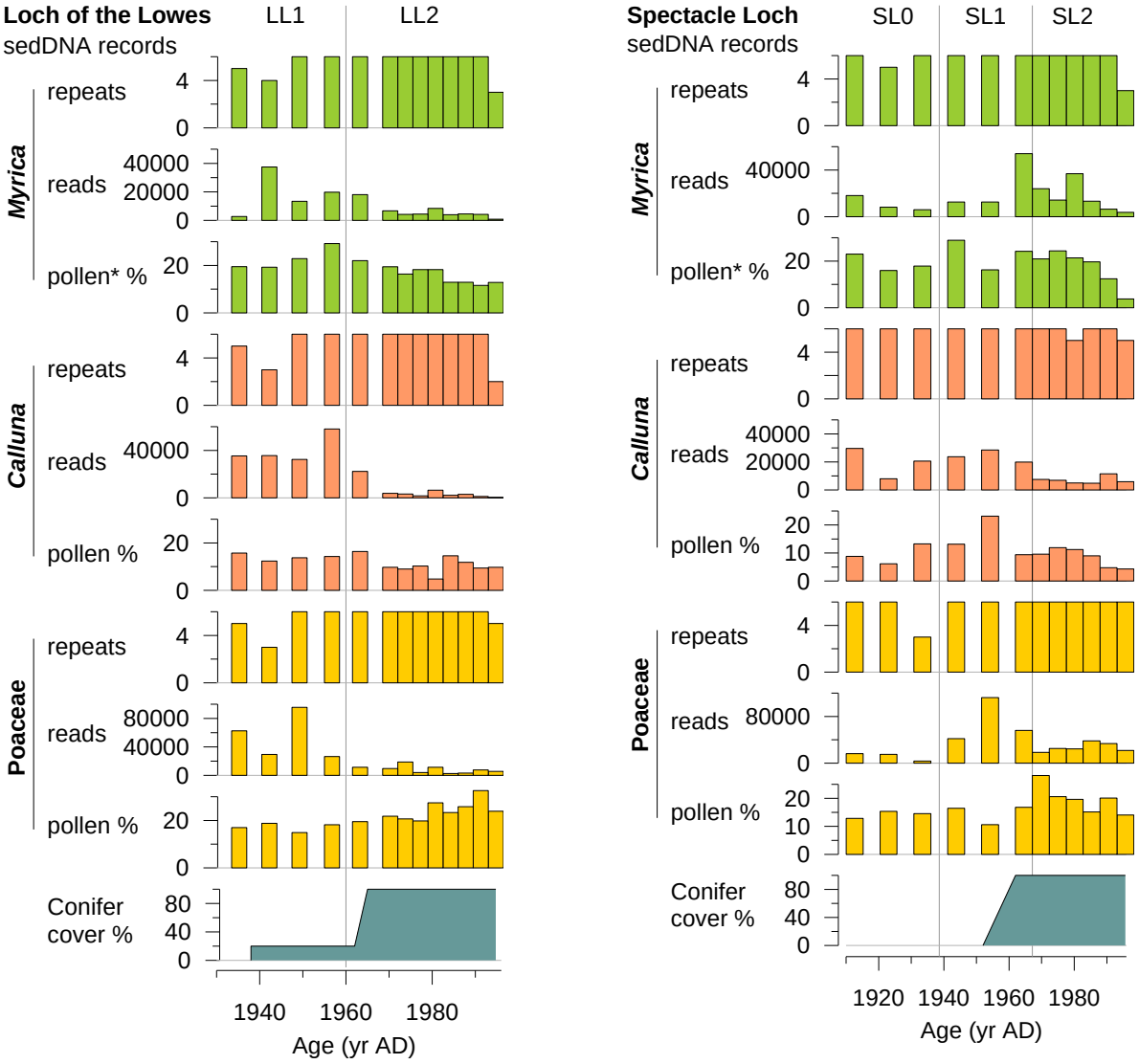
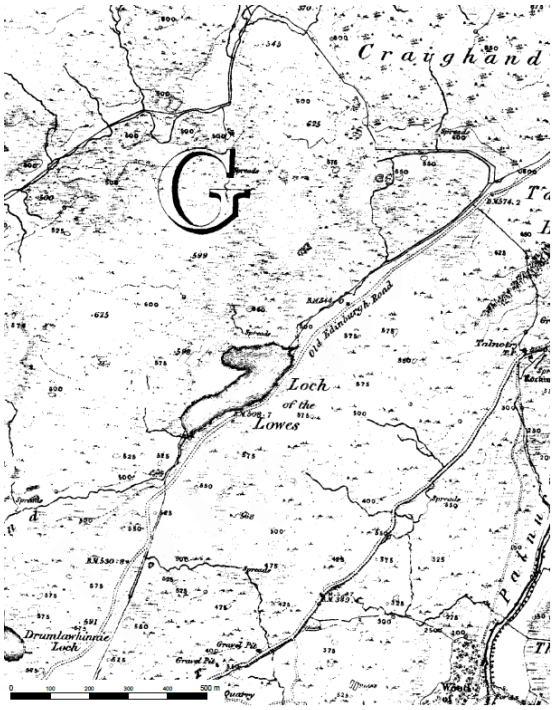
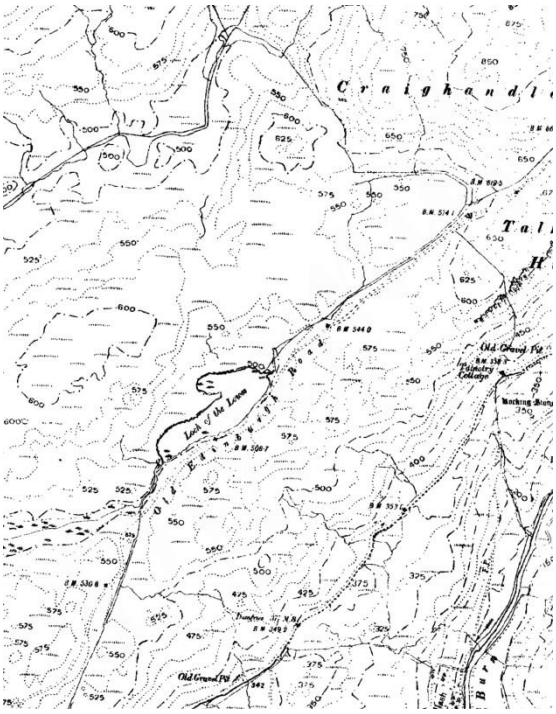


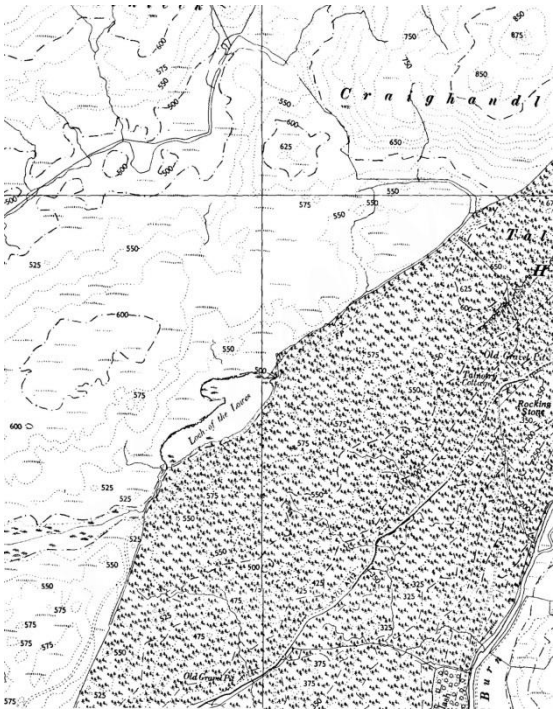
Fig. 5c Shrubs and grasses.



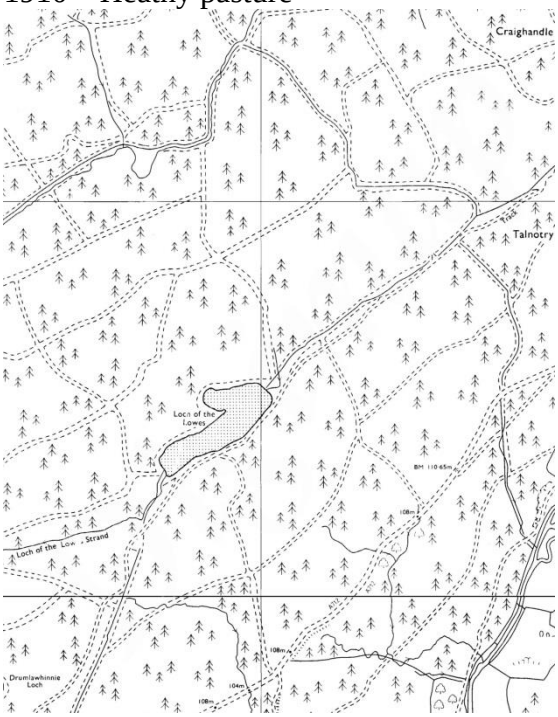
1850 – Heathy pasture



1910 – Heathy pasture



1950 – Heathy pasture, conifer forest

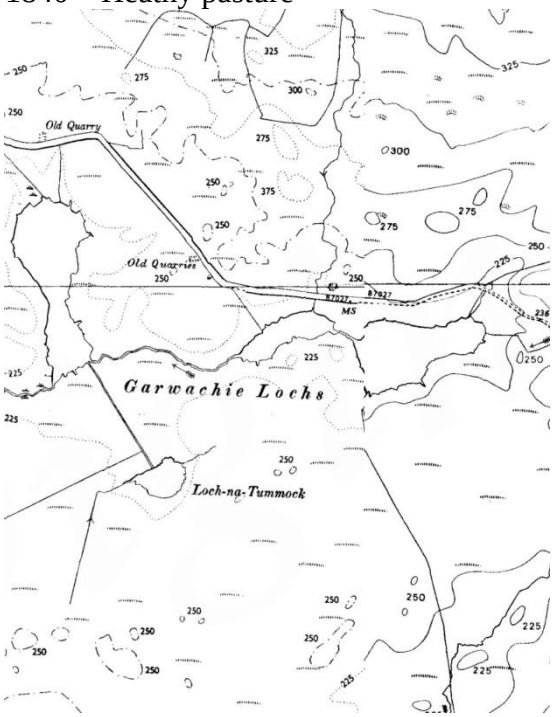


1970 – Conifer forest

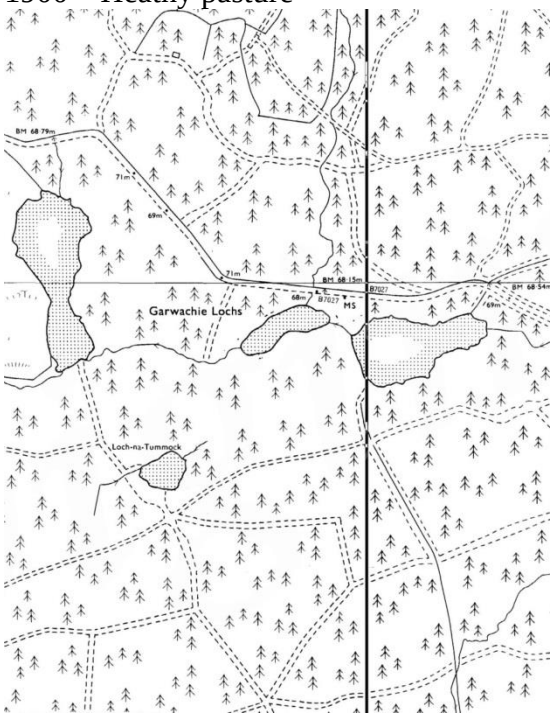
Fig. S1a Historical maps of Loch of the Lowes. The scale bar in the top left image indicates 500 m.

1840 – Heathy pasture

1900 – Heathy pasture



1950 – Heathy pasture



1970 – Conifer forest

Fig. S1b Historical maps of Spectacle Loch. The scale bar in the top left image indicates 500 m.

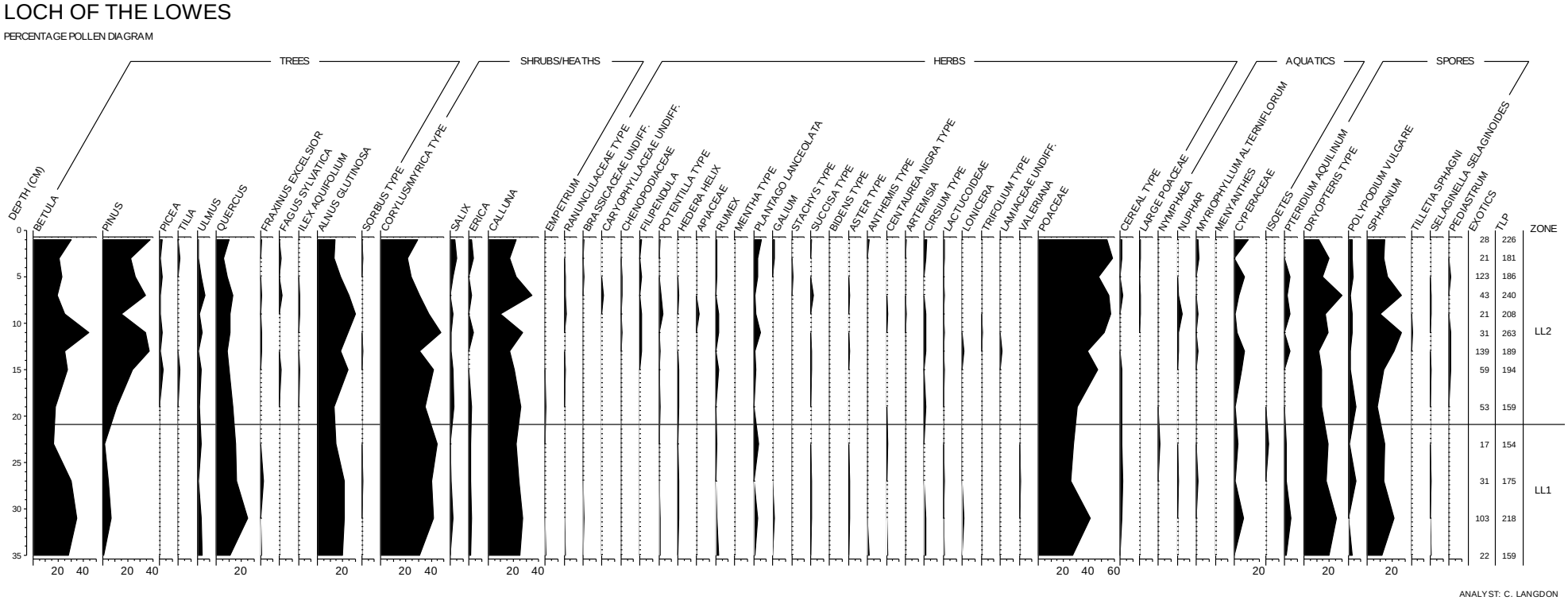


Fig. S2. Pollen percentage diagram for Loch of the Lowes. Percentage values are based on total land pollen (TLP).

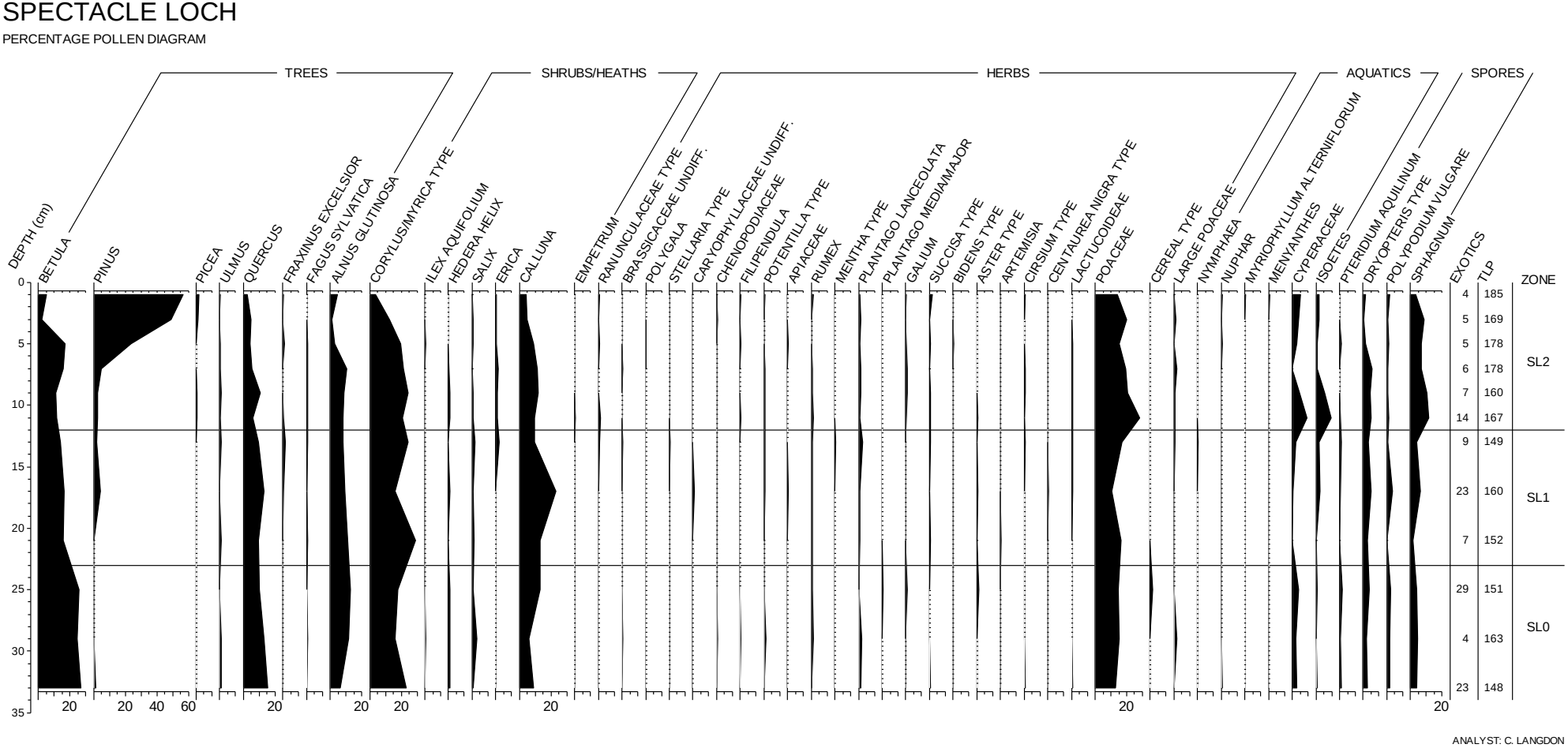


Fig. S3. Pollen percetage diagram for Spectacle Loch. Percentage values are based on total land pollen (TLP).

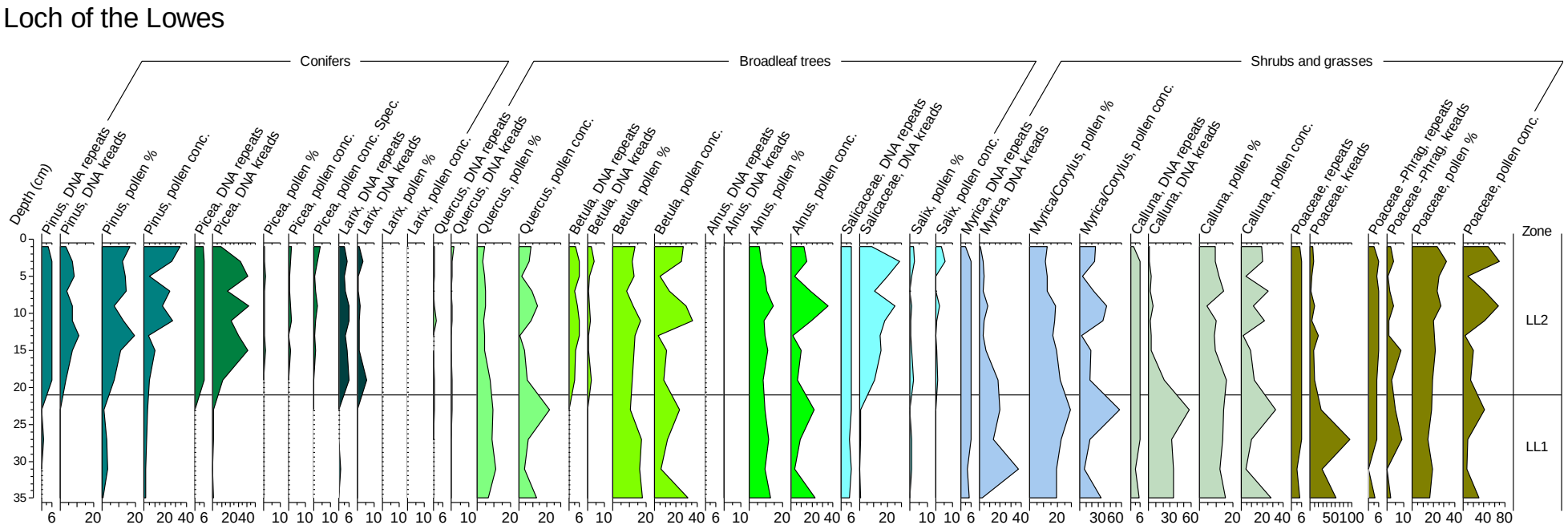


Fig. S4. Comparison of number of DNA sequence reads, number of DNA PCR repeats, and percentage pollen from Loch of the Lowes, Scotland. DNA reads are given in 1000 of reads (sum of all six PCR repeats). Pollen percentage values (pollen %) are based on total land pollen. Pollen concentrations values (pollen conc.) are given as 1000 of pollen per cm^3. ″Picea pollen conc. spec″ is based on a higher count of exotics than done for the other samples. ″Poaceae -Phrag.″ is the remaining Poaceae DNA repeats and reads when *Phragmites australis* is removed.

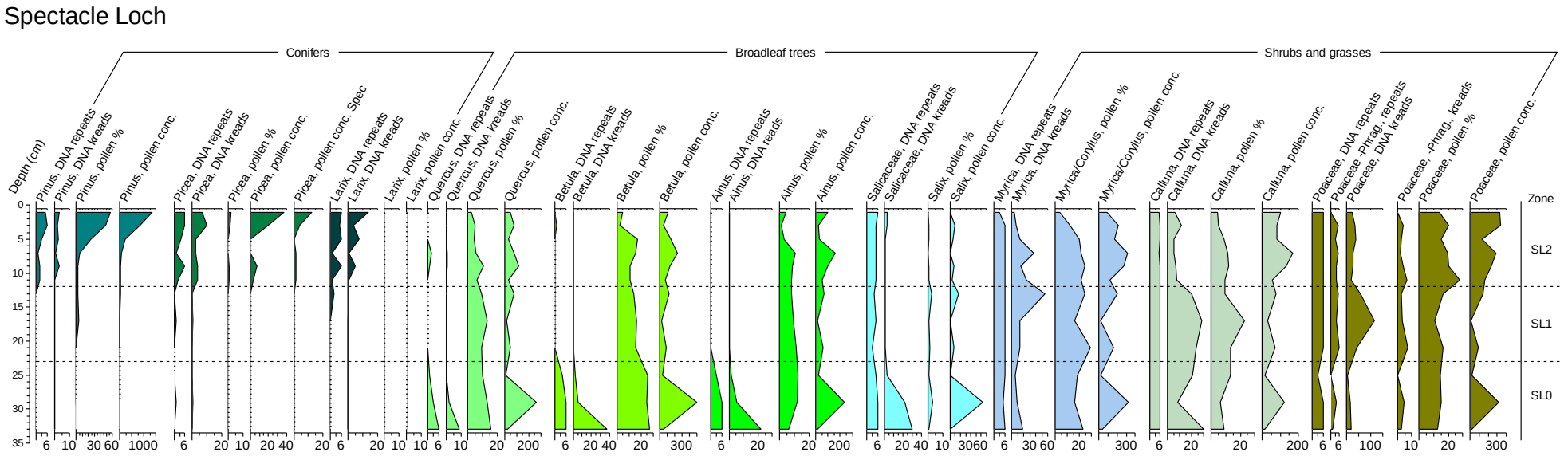


Fig. S5. Comparison of number of DNA sequence reads, number of DNA PCR repeats, and percentage pollen from Spectacle Loch, Scotland. DNA reads are given in 1000 of reads (sum of all six PCR repeats). Pollen percentage values (pollen %) are based on total land pollen. Pollen concentrations values (pollen conc.) are given as 1000 of pollen per cm^3. ″Picea pollen conc. spec″ is based on a higher count of exotics than done for the other samples. ″Poaceae -Phrag.″ is the remaining Poaceae DNA repeats and reads when *Phragmites australis* is removed.

Table S1. Re-assignment of exotic taxa to native alternative provided by BLAST (100% similarity) or to higher taxonomic level when error in PCR is likely (differ only in subsequent number of identical bp). *Pedicularis parviflora* was removed as it had identical distribution among samples as the native *Pedicularis palustris*.

Original assignment	**Re-assignment**	**Comment**
Epilobium alsinifolium	*Epilobium obscurum*	BLASTed to native taxon
Lomelosia cretica	*Succisa pratensis*	BLASTed to taxon already present
Pedicularis parviflora	*(Pedicularis palustris)*	Identical distribution among samples
Pinguicula algida	*Pinguicula vulgaris*	BLASTed to taxon already present
Ranunculus ficariifolius	Ranuculaceae	PCR error
Ranunculus reptans	Ranuculaceae	PCR error
Vaccinium ovalifolium	*Vaccinium* sp.	PCR error

Table S2. SedDNA taxa recorded in both at least one sediment sample (out of 25 samples, 150 repeats) and one negative control sample (out of 12 samples, 72 repeats) and also in at least one of the sediment samples. Only records with ≥10 reads per repeat are included. Note that for common taxa the source of the DNA in the negative controls likely is the sediment samples themselves; in addition, the potential ecological impact of a false positive repeat is generally much smaller than for a rare taxon. Of the present taxa we consider the ones with a ratio of ≥0.35 between the relative occurrence in the negative controls in relation to the sediment samples as problematic, and these are removed from further interpretation. Most other taxa have very low values and/or high occurrence in the repeats, and the ecological interpretation for these is considered unproblematic.

Taxa	Number of repeats in sediment samples	Fraction of repeats in sediments / fraction of repeats in negative controls
Calluna vulgaris	140	0.01
Phragmites australis	140	0.01
Pinus	72	0.03
Molinia caerulea	90	0.05
Ranuculaceae	37	0.06
Nuphar lutea	149	0.07
Poeae	27	0.08
Carduinae	73	0.09
Pyreae	61	0.14
Alnus	15	0.14
Myriophyllum alterniflorum	135	0.14
Ranunculus	13	0.16
Potamogeton	140	0.18
Salicaceae	136	0.26
Nymphaeaceae	150	0.33
Musaceae	6	0.35
Holcus	6	0.35
Triticeae	11	0.38
Cucurbitaceae	5	0.42
Asteraceae	19	0.44
Veronica	6	0.69
Zea mays	15	0.83
Prunus	6	1.04