Responses to referee comments and changes in the manuscript:

"Sedimentary ancient DNA and pollen reveal the composition of plant organic matter in Late Quaternary permafrost sediments of the Buor Khaya Peninsula (north-eastern Siberia)"

by: Zimmermann et al. 2016

We would like to thank the three referees for their thoughtful comments and suggestions which will improve the clarity and the quality of the paper. Furthermore, we would like to thank anonymous referee 3 that he/she shared his/her knowledge about chloroplast genomes in pollen in the discussion forum. Below, we addressed all comments and questions separately.

Referee comments 1

Page 6, line 23-24 The DNA samples were stored at -20°C and the pollen samples were stored at 4°C. Recommended: The subsample for following sedaDNA and pollen analysis were stored at -20°C and 4°C, respectively.

Thank you for your recommendation. We feel that this reduces the clarity of the sentence. But we shortened and rephrased the sentence to:
page 6, lines 26-27:
"The DNA samples were stored at -20°C and the pollen samples at 4°C."

Page 6, line 24-25 In total, 54 samples were drilled (approximately 3 to 4 samples per metre) for each kind of analysis. Recommended: meter

Thank you for this suggestion. Before we submitted the manuscript it went through a language correction in which meter was changed to metre for language consistency (British English). Therefore, we will leave it at metre.

Referee comments 2

1. What was the source for the plant DNA? Is it possible to determine it? Is it chloroplast, nuclear or mitochondrial DNA? Pollen cells don't contain chloroplasts and if it was possible to determine chloroplast DNA it would be a very reliable method for separation of local and not-local plant taxa.

The targeted DNA sequence is the *trnL*-UAA P6-loop (Taberlet et al. 2007), which is located on the chloroplast genome of plants. As already answered in the discussion forum by Referee 3, chloroplast genomes are indeed present in pollen and therefore (unfortunately) cannot be used to separate local

from non-local plant taxa. We decided to address this point very briefly in the discussion of the manuscript.

page 19, lines 22-23:
"Furthermore, the applied sedaDNA marker is located on the chloroplast genome, which is transmitted through pollen in Pinaceae (reviewed in: Mogensen, 1996)."

2. I don't like how the sedaDNA and pollen diagrams ('stratigrams') are structured. It is very hard to find any taxa on them. I suggest structuring them in ecological way – herbs, shrubs, trees etc. or explain what the principle of such structure was.

We agree and rearranged all stratigraphic diagrams (Figures 2, 4, 6 and 8) in the suggested manner. However, within the groups we retained the original sorting.
The diagrams were plotted using the R command *strat.plot* from the package rioja (Juggins, 2012) and sorted using wa.order. Our intention was to better visualize the changes in species composition from the top of the core to the bottom. We hope that this compromise between an ecological meaningful structure and the application of wa.order is acceptable.

3. I suspect that in the lower stratigraphical zones pollen and other material can be redeposited. Radiocarbon dating, amount of pre-Quaternary pollen and spores indirectly confirm this suggestion. How do you explain the inversion in radiocarbon dates? Maybe, it is contamination from the lower more ancient layers? Why do you not use pollen concentrations that can be an additional marker of redeposition?

Thank you for this suggestion. According to Schirrmeister et al. (2016) the deepest 2.5m of the core imply a former (ancient) active layer. An active layer is prone to disturbances such as erosion, cryoturbation (through seasonal thawing and re-freezing) and potentially also grazing, all of which allow for re-deposition of older material. In the deepest 2.5m the amount of pre-Quaternary pollen and spores is highest (up to ~5%) among all samples and give weight to your suggestion. Additionally, pollen concentration is lowest in this part of the core. For the rest of the core the amount of pre-Quaternary pollen and spores is lower, while pollen concentration is higher. However, the radiocarbon dating is at its limit and therefore associated with high standard deviations, which is the most likely explanation for the age "inversion".

We included:
in the material and methods section page 10 lines 17-18:
"For calculation of concentration of pollen a Lycopodium spore tablet (Batch no. 1031; n=20848 +/- 1460) was added to each sample (Stockmarr 1971)."

in the results section page 16, line 17-18 and in counts are reported in Table 3:
"Furthermore, the major proportion of redeposited, pre-Quaternary pollen can be found in this zone, while total pollen concentration is lowest (Table 3)."

and page 18, line 16-17:
"Similar to the redeposited pollen the major proportion of pre-Quaternary spores can be found between 17.05 m and 18.15 m."

In the discussion we included on page 22 lines 28- page 23, line 2:
"Below the ice wedge the core shows fluctuating ages associated with high standard deviations, which can be explained either by the radiocarbon dating method being at its limit or by reworking of the sediments."

and page 22, line 14-18:
"In the deepest 2.5 m the amount of redeposited pre-Quaternary pollen and spores is highest (up to ~5%) among all samples. According to the sedimentary and hydrogeochemical results presented in Schirrmeister et al. (2016), this part of the core implies an ancient active layer. An active layer is prone to disturbances such as erosion, cryoturbation (through seasonal thawing and refreezing) and potentially also grazing, all of which allow for redeposition of older material."

What part of the core contains higher percentages of exotic DNA sequences (contamination)?

Exotic DNA sequences were found in almost all samples, extraction blanks and PCR negative controls and are reported in S5_cultivated_nonArctic_plants.xlsx. The highest contribution of exotic DNA sequence counts were found at 2.85m depth. We realized that we forgot to change the original sample names with the corresponding depths in the supplementary tables. We apologize for this and added this information in all supplementary tables.

Furthermore, we changed the name of the supplementary table to: S6_exoticDNAsequences.xlsx. The table now also includes for each sample the percentages of exotic DNA sequences of the total read counts (based on best identity = 1.0).

page 9, line 27 – page 10, line 3:
"These exotic DNA sequences were detected in almost all samples but contained mostly less than 1 % of the total number of read counts with best identity of 1.0 (S6). The highest contribution of exotic DNA sequence counts within a sample was at 2.85 m depth from Musaceae."

4. What taxa belong to NPPs in this study? Page 9, line 21: 'pollen, spores and non-pollen palynomorphs (NPPs)'. Table 3 ('Number of non-pollen palynomorphs for each sample') contains algae, fungi, mosses, ferns (!), lycopods (!) and pre-Quaternary spores (!). Page 17, lines 13-14: 'A total of 1,092 NPPs were counted and assigned to 25 taxa, comprising four mosses, two spikemosses, six clubmosses, three ferns, six fungi and four green algae'. Usually, in pollen study to NPPs belong fungi, algae, remnants and eggs of animals etc.; objects which can be determined in the pollen slide after chemical treatment. Spores of higher vascular plants don't relate to NPPs (as a rule, but you can explain your position).

Thank you for pointing this out. We decided to omit the term NPP and instead use "spores and algae". This was corrected throughout the text. Accordingly, we changed the title (page 18, line 10) from: "4.2.3 Non-pollen palynomorphs" to "4.2.3 Spores and algae".

5. I didn't understand for what you calculated terrestrial-aquatic and Poaceae-Cyperaceae ratios. You did not use it in discussion. I didn't also realize where is the interpretation of the PCA analysis in this study.

Thank you for pointing this out, we agree. In some palynological studies the ratio of Poaceae to Cyperaceae is used to assess temporal changes in humidity. We applied this ratio to trace the hydrological development along the core since our results suggested the local presence of a shallow water body in several depths. However, since some Poaceae (e.g. *Arctophila fulva*/*Dupontia fisheri*) are associated with wet conditions while the Cyperaceae *Kobresia* is an indicator for (cryo-)arid climate we chose to compare this ratio with the terrestrial-aquatic ratio. Additionally, the ratios were an instrument to visualize and compare the contribution of terrestrial and swamp/aquatic taxa between samples as well as between the two proxies.

We rephrased page 18 lines 21-24 to:
"Ratios were built for the sedaDNA and pollen datasets to assess whether terrestrial or swamp and aquatic taxa contributed more sequences or pollen grains to a sample and were compared to the corresponding Poaceae-Cyperaceae ratios. This allowed us to trace local hydrological changes and to identify drier (positive values) and wetter phases (negative values) (Fig. 10)."

We now directly refer to the PCA biplots and the ratio plot in the discussion.
page 22, lines 27-28:
"The zonation matches the major structure presented in the PCA biplots of sedaDNA and pollen."

page 23, lines 22-26:
"In several samples the sedaDNA record is dominated by Cyperaceae and other swamp and aquatic taxa, with mostly negative values on the ratio plot, especially for the sedaDNA dataset. This indicates wet conditions on a local scale and probably enhancing organic matter accumulation, as cold and anoxic conditions reduce decomposition rates (Davidson et al., 2000)."

page 26, lines 1-2:
"The sedaDNA results imply profound changes after the LGM, which is displayed in the major structure of the terrestrial PCA biplot."

page 26, line 7:
"In relatively drier periods, indicated by positive values in the ratio plot, the organic matter comprises [...]."

page 26, lines 9-10:
"During moister periods, indicated by negative values in the ratio plot, increased [...]."

6. I suggest constructing the age-depth model for the upper part of the core. CONISS reveals two pollen zones in upper Pleistocene-Holocene part (Fig. 6: TerrPZ3). Maybe, it is the border between YD and Boreal.

Thank you for this suggestion. We performed the age-depth model with Bacon. We included a brief sentence in the methods section, included the ages in the stratigraphic diagrams (Fig. 2, 4, 6, 8) and

adjusted the captions, acknowledged Liv Heinecke for building the age-depth model and included the Figure of the age-depth model in the Supplement (Figure S3.7).

page 12, line 18:

"The age-depth model was built using the Bacon package (Blaauw and Christen, 2011) in R."

The reference was included: *Blaauw, M., Christen J.A.: Flexible paleoclimate age-depth models using an autoregressive gamma process, Bayesian Anal. 6, 457–474, doi: 10.1214/ba/1339616472, 2011*

page 29, line 17:

"We gratefully acknowledge [...], Liv Heinecke for the age-depth model and Bennet Juhls for compiling the maps."

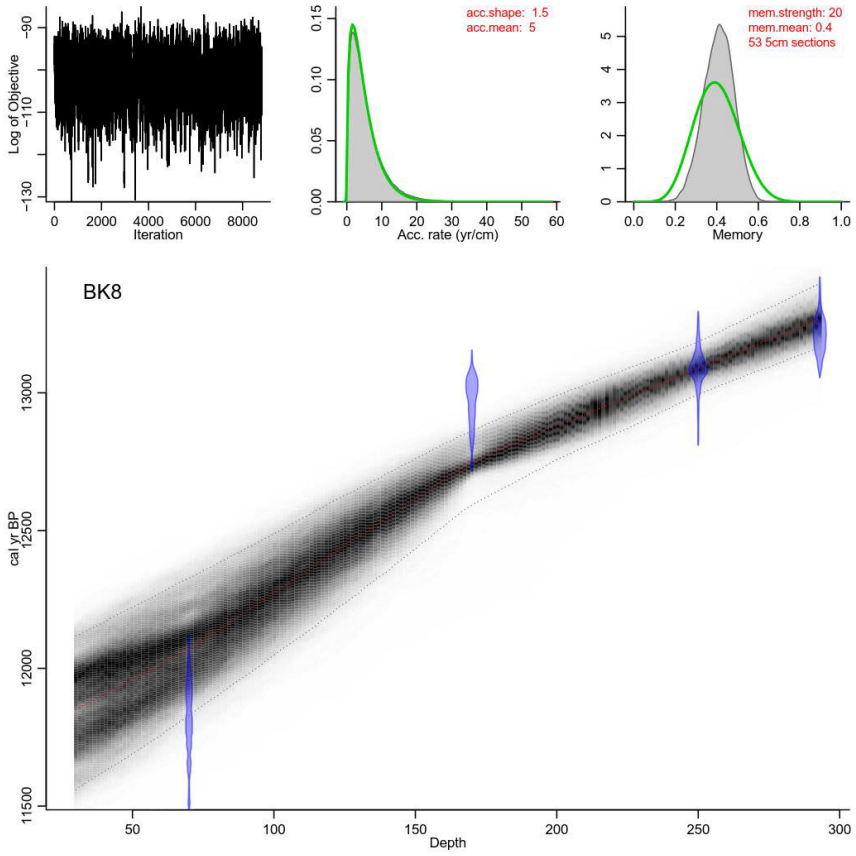

Figure S3.1: Age-depth model for permafrost sediments above the ice wedge. The calibrated ages are reported as calendar years before present (cal yr BP), the depths in cm.

7. Where is S3? Page 12, lines 7-8: 'The complete taxa-list is available in S3'.

This was unintended and is a mistake. The complete list was available in S4. However, since the order of the supplements was changed it is now S5 (S5_DNA_dataset.xlsx). This was corrected.

8. How do you explain the hiatus between last radiocarbon date (9700±50 14C yr BP) on the depth of 0.3 m and modern sample on the depth of 0.1 m?

The permafrost core was drilled at the top of a Yedoma hill (Schirrmeister et al., 2016). Wind and rain probably eroded the younger deposits. We will include this and rephrase the beginning of the discussion chapter 5.3 in the manuscript from:
"The upper part of the core from the ice-wedge up to approximately 0.25 m depth includes the late glacial transition to the early Holocene (approximately 11.4 to 9.7 kyr BP, (13.4–11.1 cal kyr BP)). As emphasized in Andreev et al. (2011) records of the late glacial transition are rare because of active thermoerosion. Hence, our results provide valuable information about the vegetation history in this region and organic matter composition."

page 25, lines 18-24:
"The permafrost core was drilled at the top of a Yedoma hill (Schirrmeister et al. 2016). Wind and rain probably eroded most of the Holocene deposits resulting in a hiatus between the sample of the modern core top and the second sample at 0.25 m depth (11.1 cal kyr BP). The upper part of the core consists of sediments dated to the transition from the late glacial to the early Holocene (13.4–11.1 cal kyr BP). As emphasized in Andreev et al. (2011) records of the late glacial transition are rare because of active thermoerosion. Hence, our results provide valuable information about the vegetation history in this region and organic matter composition."

<u>Specific comments:</u>
Page 1, line 28: 'a shrub tundra spectrum' – spectra (plural). One sample - one spectrum, several samples – several spectra. Or you should use 'pollen complex'.
We rephrased the sentence from:
"Pollen also records a shrub tundra spectrum, mostly seen as changes in relative proportions of the most dominant taxa, while a decrease in taxonomic richness was less pronounced compared to sedaDNA."

To (page 1, lines 28-30):
"Pollen also records a shrub tundra community, mostly seen as changes in relative proportions of the most dominant taxa, while a decrease in taxonomic richness was less pronounced compared to sedaDNA."

Saliceae, Pooideae etc. are the tribes. Please, specify what nomenclature system you used in this study.
Indeed, we did not state the basis for the taxonomic assignment sufficiently clear. The taxonomic assignment for sedaDNA is based on the NCBI taxonomy. The taxonomic assignment is performed using the OBItools command *ecotag* (Boyer et al., 2016) which uses a Last Common Ancestor approach based on the NCBI taxonomy database (Sayers et al., 2009). To clarify which system we used, we will add the following sentence on

page 9 lines 20-21:
"The nomenclature for the taxonomic assignment follows the NCBI taxonomy (Sayers et al., 2009)."

I found the mistakes in Latin. 'Osmuda' – Osmunda (everywhere) Polypodiaceae is the family mainly of tropical ferns. Use Polypodiophyta or specify what nomenclature system you used in this study. 'Botrycoccus' – Botryococcus (in the text) 'Cichorideae' – Cichoriodeae (Fig.6)

Thank you for pointing these out. We corrected this to Botryococcus and Cichoriodeae. We followed Andreev et al. (2011), who uses Polypodiaceae, but we changed it to Polypodiophyta. We changed the Figures 6 and 7 accordingly.

Page 13, Line25. 'assigned to 21 swamp or aquatic taxa' . In S4 only 20.

This was unintended. The list is now corrected: Poinae was transferred from terrestrial to the swamp/aquatic list.

Page 15, Line 11: 'to 53 taxa, including indeterminable and pre-Quaternary pollen.' Indeterminable and pre-Quaternary pollen are not taxa and cannot be included in the taxa list.

This is true. We wanted to say the 53 taxa, the pre-quaternary pollen and indeterminable pollen sum up to a total of 8881 terrestrial pollen grains. We addressed this by changing the sentence to:

page 16, lines 4-5:

"A total of 8,580 terrestrial pollen grains were counted and ascribed to 53 taxa, while 248 were indeterminable and 53 assigned as pre-Quaternary pollen."

We also adjusted the caption of Figure 7:

"PCA biplot of terrestrial taxa from pollen analysis, showing a subset of 23 taxa, indeterminate and pre-Quaternary pollen, [...]."

Page 19, Line 1: 'the under-representation of Salix in comparison to other plant functional types'. Plant functional type is the unit of biome reconstruction. You should rephrase this sentence.

We agree and with regard to shortening the paper (referee 3) we discarded this sentence and the one before.

Line 20: 'Compared to the number of vascular plant taxa (58) and bryophytes (20) recorded by pollen analysis'. Where is the list with moss taxa, determined by pollen analysis?

The list is included in Table 3. With regard to comment number 4 we rephrased the sentence:

Page 21, lines 6-7:

"Compared to the number of vascular plant taxa (58) and bryophytes (4) recorded by pollen analysis, the sedaDNA approach recorded a higher number of both vascular plants (134) and a bryophytes (20)."

Page 23. Line 27: 'In drier periods'. You have reliable chronology for the upper part. When were drier or wetter periods?

We agree that the ages could be included. However, we now believe this could lead to an over-interpretation of the results. Furthermore, with regard to the length of the manuscript and the readability, we decided to refer in the corresponding sentences to not to the ages but to the ratio plot and rephrased the paragraph to:

page 26, lines 7-14:
"In relatively drier periods, indicated by positive values in the ratio plot, the organic matter comprises mostly Poaceae, which were represented by a different composition in comparison to the pre-LGM, with Agrostidinae and Poa MOTU2. During moister periods, indicated by negative values in the ratio plot, increased proportions of wetland plants such as Eriophorum, Equisetum and Ranunculus were recorded. The high proportions of Equisetum in this time interval are supported by our spore record as well as in the palaeogenetic study of Willerslev et al. (2014) and the palynological review of Andreev et al. (2011)."

Page 24, Line 10: 'Published pollen records for 11.7–10.6 kyr BP are dominated by Cyperaceae and Poaceae. Shrub pollen increased at approximately 9 kyr BP (Andreev et al., 2011), with up to 60 % of Betula in the Khorogor Valley near Tiksi (Andreev et al., 2011; Grosse et al., 2007). These results match well with the pollen data presented here.' It is not true. In your study Betula pollen increased at least 2 kyr earlier (Figs 2,6).

The pollen diagram Khg-11 in Grosse et al. (2007) shows an increase in *Betula* sect. Nanae of ~35% and *Betula* sect. Albae of ~25% at 11.540 ± 60 ^{14}C yr BP. In our study the increase is recorded at ~11.400 ± 50 ^{14}C yr BP. Hence, it does match well, but it was not clearly stated why. We therefore rephrased the sentence:

page 26, lines 20-24:
"Shrub pollen increased in the Laptev Sea region approximately at 9 kyr BP (Andreev et al., 2011), while in the Khorogor Valley near Tiksi an increase especially of Betula pollen up to 60% was recorded already at 11.54 ± 60 kyr BP (Grosse et al., 2007; Khg-11). In this study the increase is recorded at 11.4 ± 50 kyr BP and therefore matches well with the pollen data of the Khorogor Valley."

Referee comments 3

I read with interest this manuscript that offers interesting results in good agreement with similar recent studies conducted by other researchers on similar settings from lakes and peat. The manuscript presents the results of ancient pollen and sedimentary DNA study extracted from permafrost from the Buor Khaya Peninsula. It shows several important methodological results of interest for readers working with plant ancient DNA from bulk sediments. It shows first that combination of classical palynological analyses is often not sufficient to resolve flora composition and that DNA is an important complementary tool to investigations of this kind. It confirms (and this should perhaps be strengthened even more by the authors) that DNA signal is local in origin and this fact helps when combined with pollen proxies that we know may have limitations linked to the fact that pollen is transported over long distances. Among papers that should be cited for this result are: Pedersen et al 2016, Sjögren et al 2016 and Parducci et al 2013.

We thank Referee 3 for his/her positive feedback. We will strengthen that our results suggest the local origin for sedaDNA in this study and its benefits in combination with palynological analyses. We included the recommended papers.

Page 19, line 20 – page 20 line 4:

"In most of the samples we did not detect conifer-derived sedaDNA, although they are present in the pollen record. Hence, the Pinaceae pollen presumably originated from extra-regional stands (Birks, 2001; van der Knaap, 1987). Furthermore, the applied sedaDNA marker is located on the chloroplast genome, which is transmitted through pollen in Pinaceae (reviewed in: Mogensen, 1996). If pollen contributed significantly to the sedaDNA record, we would expect to find it at least in samples with high Pinaceae proportions, which we did not. This supports two assumptions about sedaDNA: first, that sedaDNA originates mainly locally (Haile et al., 2007, 2009; Jørgensen et al., 2012; Parducci et al., 2013; Pedersen et al., 2016; Sjögren et al., 2016; Yoccoz et al., 2012) and, second, that it is predominantly derived from roots and other plant parts rather than from pollen (Jørgensen et al., 2012; Levy-Booth et al., 2007; Parducci et al., 2013; Pedersen et al., 2016; Sjögren et al., 2016; Willerslev et al., 2003)."

I am less familiar with the type of statistical analyse the authors used and I cannot comment on that in details. I can perhaps comment generally that they seem to me to overanalyse and over interprete their data. For example I didn't understand why to calculate terrestrial-aquatic and Poaceae-Cyperaceae ratios as well as the PCA analysis if these calculations are not later even used and discussed.

We understand this concern and referred to the PCA and ratio plot in the discussion. With regard to the over-interpretation, we are aware that it is difficult to use sedaDNA in a quantitative way due to technical biases. However, we were conservative in our bioinformatic filtering, in our threshold to include taxa in the statistical analysis and in comparing taxonomic richness between samples. Furthermore, we do not rely entirely on sedaDNA but compare it to pollen.

In some palynological studies the ratio of Poaceae to Cyperaceae is used to assess temporal changes in humidity. We applied this ratio to trace the hydrological development along the core since our results suggested the local presence of a shallow water body in several depths. However, since some Poaceae (e.g. *Arctophila fulva*/*Dupontia fisheri*) are associated with wet conditions while the Cyperaceae *Kobresia* is an indicator for (cryo-) arid climate we chose to compare this ratio with the terrestrial-aquatic ratio. Additionally, the ratios were an instrument to visualize and compare the contribution of terrestrial and swamp/aquatic taxa between samples as well as between the two proxies. We hope that we were able to make our intent clearer and that our reasoning is satisfying.

Changes are reported in the reply of Referee 2 comment 5.

I reply here also to one of the other referee to state that pollen contains all three plant DNA genomes (mitochondrial, nuclear and chloroplast) regardless of the type of inheritance of the organelle DNA of the mother plant (paternal or maternal). There is however an increase or a decrease of the mitochondrial or chloroplast organelles in pollen after mitosis 1, depending on inheritance. So Pollen contains cpDNA as well as mtDNA in addition to the nuclear of course. This has been demonstrated in a number of different experimental studies in the lab conducted on fresh pollen (e.g. Nagata et al 1999) as well as in sito in ancient sediments: cpDNA has been amplified from pollen of plants with maternal inheritance and vice versa for mtDNA in conifers for ex. It is now know, but it still require some studies, that plant and in sediments comes mainly from plant tissues like macros and that likely it mainly present in its extracellular forms in sediments

(linked and protected by solid particles). However this is a field that requires lots of studies in the future: taphonomy of plant ancient DNA in sediments.

Thank you very much for this statement and the additional information which you provided in the discussion forum. We agree and are excited to read your paper when it is published.

It is not clear in the results how many samples are analysed for DNA. The authors write of extraction batches of 11 samples. But how many in total? In that paragraph this should be stated more clearly. Were all 32 samples analysed for pollen also analysed for DNA?

We agree, and included a sentence on page 7, lines 2-3:
"All 54 samples were analysed for sedaDNA while only 32 were processed for pollen analysis."

The authors write that they run up to 4 PCR per samples. This is not very good if it means that some samples perhaps have been amplified only one for example. State in a table how many PCR per sample or state the minimum number amplified. Why not amplify the same number for all samples?
We are sorry for this confusion. We pooled exactly two PCR-products for each sample. However, in a few cases up to 4 PCRs were necessary to get two positive PCR-products with clean PCR negative controls. We changed the corresponding on

page 8, lines 9-19 to:
"For each sample we pooled two positive amplifications for sequencing, under the condition that the associated NTCs and extraction blank were negative. The two pooled positive amplifications were purified using the MinElute PCR Purification Kit (Qiagen, Germany), following the manufacturer's recommendations."

Did the authors sequence the positive blanks?
The extraction blanks as well as the PCR negative controls were all negative, indicated by the gel-electrophoreses after the PCRs. Nonetheless, we sequenced all controls. On page 8 lines 5-6 of the first version we wrote: "All extraction blanks and NTCs were included in the sequencing run, using a standardized volume of 10 µl, even though they were negative in the PCRs." Furthermore, the (unfiltered) sequencing results of extraction blanks and PCR controls are reported in Supplement S6. We clarified this by adding the sentence:

page 8, lines 19-20:
"The sequencing results of extraction blanks and PCR controls are reported in the supplementary file S4."
Since we refer to this supplement in an earlier section now, the supplement order was changed accordingly to:

1. S1-S3_msBK8.pdf
    - S1 Sequencing and sequence filtering
    - S2 Technical evaluation and authenticity of the data
    - References
    - S3 Figures
2. S4_negative_controls.xlsx

3. S5_DNA_dataset.xlsx
4. S6_exotic_DNA_sequences.xlsx
5. S7_Pollen_counts.xlsx

Paragraph 6-14 on page 8 should be moved earlier in the paragraph.
We moved the paragraph to the beginning of part 3.3.

I did not understand the sentence on page 18 line 18: The sedDNA record does not contain conifer-derived DNA (with reference from Birks and van der Knaap). Why are these two reference are used here? To confirm what, that DNA of conifer is not detectable by sedDNA. this is not the case. Not clear.
We agree. This is a remnant from an older version of this sentence and we rephrased the sentence

page 19, lines 20-21s:

"In most of the samples we did not detect conifer-derived sedaDNA, although they are present in the pollen record. The conifer pollen presumably originated from extra-regional stands through long-distance dispersal (Birks, 2001; van der Knaap, 1987)."

Page 19, line 17-18. the authors should consider to use MapDamage 2.0 to test for patterns and rates of DNA damage and assure for authenticity of their reads.
We considered using mapDamage 2.0 but came to realize that the program is not adequate for metabarcoding. mapDamage 2.0 requires a (genomic) reference sequence to which the reads of interest are aligned. From this alignment it uses a Bayesian approach to estimate nucleotide misincorporation rates and fragmentation patterns (Jónsson et al., 2013).
The first problem we encountered is that the amplicons we use are highly variable and do not allow for an overall alignment, even if we use closely related taxonomic subsets. The second problem is that the numbers of sequences without an exact match to their corresponding reference sequence are too low to produce an error distribution. The third problem is, that in our study we analysed only sequences with an exact match to a reference sequence. Therefore, we cannot really "prove" that these sequences are authentic. mapDamage would (theoretically) only prove that the sequences, which we discarded are probably authentic. However, if you have any suggestions on how we could apply mapDamage appropriately, we are very interested.

The manuscript and especially discussion is long and can be shortened in my opinion.
We shortened the original discussion by deleting fill words. However, since we had to include some recommended aspects and references in the discussion it is now approximately the same word count as before.

<u>Further improvements:</u>

1. For consistency we changed herbs to forbs and herbaceous taxa to forb taxa, everywhere in the text. Accordingly, we changed the sentence from: "SedaDNA portrays high diversity among grasses, a rich herbaceous flora including swamp and aquatic taxa and high proportions of *Salix*."

page 23 , line 21-22 to:
"SedaDNA portrays high diversity among grasses and forbs including swamp and aquatic taxa well as high proportions of Salix."

2. The reference Stapel et al. 2016 is now published. We updated the reference.

3. We included (authentic) on page 13 line 5:
"No (authentic) tree taxa were detected, although Saliceae, Betula and Alnus occurred there as shrub form. 64.5 % of all terrestrial sequence counts are from shrubs (Saliceae, Betula, Alnus), while 18.1 % are from graminoids and 17.4 % forbs."

4. We found a mistake in the results section regarding the zonation of the terrestrial DNA dataset: correct: TerrDNA Zone 1 (18.48-14.85 m), TerrDNA Zone 2 (14.55–3.3 m)

   This was only a mistake in the text and did not occur in the corresponding plots and it did not influence any results or interpretations.

5. We improved Figure 10 and included the radiocarbon ages.

6. We reduced some text by putting median values in brackets. For examples we changed the sentence: "TerrDNA Zone 1 is dominated by Saliceae ranging from 22–99.8 % with a median of 91 %) [...]."
   page 13, lines 13-14:
   to: "*TerrDNA Zone 1 is dominated by Saliceae ranging from 22–99.8 % (median = 91 %) [...]."*

References

Andreev, A. A., Schirrmeister, L., Tarasov, P. E., Ganopolski, A., Brovkin, V., Siegert, C., Wetterich, S. and Hubberten, H.-W.: Vegetation and climate history in the Laptev Sea region (Arctic Siberia) during Late Quaternary inferred from pollen records, Quat. Sci. Rev., 30(17–18), 2182–2199, doi:10.1016/j.quascirev.2010.12.026, 2011.

Birks, H. H.: Plant macrofossils, in Tracking Environmental Change Using Lake Sediments, vol. 3 Terrestrial, Algal, and Siliceous Indicators, edited by J. P. Smol, H. J. B. Birks, W. M. Last, R. S. Bradley, and K. Alverson, Springer Netherlands, Dordrecht. [online] Available from: http://link.springer.com/10.1007/0-306-47668-1 (Accessed 29 July 2016), 2001.

Blaauw, M., Christen J.A.: Flexible paleoclimate age-depth models using an autoregressive gamma process, Bayesian Anal 6, 457–474, doi: 10.1214/ba/1339616472, 2011

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P. and Coissac, E.: OBITools: a Unix-inspired software package for DNA metabarcoding, Mol. Ecol. Resour., 16, 176–182, doi:10.1111/1755-0998.12428, 2016.

Grosse, G., Schirrmeister, L., Siegert, C., Kunitsky, V. V., Slagoda, E. A., Andreev, A. A. and Dereviagyn, A. Y.: Geological and geomorphological evolution of a sedimentary periglacial landscape in Northeast

503 Siberia during the Late Quaternary, Geomorphology, 86(1–2), 25–51,
504 doi:10.1016/j.geomorph.2006.08.005, 2007.

505 Haile, J., Holdaway, R., Oliver, K., Bunce, M., Gilbert, M. T. P., Nielsen, R., Munch, K., Ho, S. Y. W.,
506 Shapiro, B. and Willerslev, E.: Ancient DNA chronology within sediment deposits: Are Paleobiological
507 reconstructions possible and is DNA leaching a factor?, Mol. Biol. Evol., 24(4), 982–989,
508 doi:10.1093/molbev/msm016, 2007.

509 Haile, J., Froese, D. G., MacPhee, R. D. E., Roberts, R. G., Arnold, L. J., Reyes, A. V., Rasmussen, M.,
510 Nielsen, R., Brook, B. W., Robinson, S., Demuro, M., Gilbert, M. T. P., Munch, K., Austin, J. J., Cooper,
511 A., Barnes, I., Möller, P. and Willerslev, E.: Ancient DNA reveals late survival of mammoth and horse
512 in interior Alaska, Proc. Natl. Acad. Sci., pnas.0912510106, doi:10.1073/pnas.0912510106, 2009.

513 Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., Orlando, L. mapDamage2.0: fast approximate
514 Bayesian estimates of ancient DNA damage parameters. Bioinformatics, 29(13):1682-1684.
515 doi:10.1093/bioinformatics/btt193, 2013

516 Jørgensen, T., Haile, J., Möller, P., Andreev, A., Boessenkool, S., Rasmussen, M., Kienast, F., Coissac,
517 E., Taberlet, P., Brochmann, C., Bigelow, N. H., Andersen, K., Orlando, L., Gilbert, M. T. P. and
518 Willerslev, E.: A comparative study of ancient sedimentary DNA, pollen and macrofossils from
519 permafrost sediments of northern Siberia reveals long-term vegetational stability, Mol. Ecol., 21(8),
520 1989–2003, doi:10.1111/j.1365-294X.2011.05287.x, 2012.

521 Juggins, S.: rioja: Analysis of Quaternary science data, R package version 0.7-3., 2012.

522 van der Knaap, W. O.: Transported pollen and spores on Spitsbergen and Jan Mayen, Pollen Spores,
523 24, 449–453, 1987.

524 Levy-Booth, D. J., Campbell, R. G., Gulden, R. H., Hart, M. M., Powell, J. R., Klironomos, J. N., Pauls, K.
525 P., Swanton, C. J., Trevors, J. T. and Dunfield, K. E.: Cycling of extracellular DNA in the soil
526 environment, Soil Biol. Biochem., 39(12), 2977–2991, doi:10.1016/j.soilbio.2007.06.020, 2007.

527 Mogensen, H. L.: The Hows and Whys of Cytoplasmic Inheritance in Seed Plants, Am. J. Bot., 83(3),
528 383, doi:10.2307/2446172, 1996.

529 Parducci, L., Matetovici, I., Fontana, S. L., Bennett, K. D., Suyama, Y., Haile, J., Kjaer, K. H., Larsen, N.
530 K., Drouzas, A. D. and Willerslev, E.: Molecular- and pollen-based vegetation analysis in lake
531 sediments from central Scandinavia, Mol. Ecol., 22(13), 3511–3524, doi:10.1111/mec.12298, 2013.

532 Pedersen, M. W., Ruter, A., Schweger, C., Friebe, H., Staff, R. A., Kjeldsen, K. K., Mendoza, M. L. Z.,
533 Beaudoin, A. B., Zutter, C., Larsen, N. K., Potter, B. A., Nielsen, R., Rainville, R. A., Orlando, L., Meltzer,
534 D. J., Kjær, K. H. and Willerslev, E.: Postglacial viability and colonization in North America's ice-free
535 corridor, Nature, 537(7618), 45–49, 2016.

536 Sayers, EW., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., DiCuccio,
537 M., Edgar, R., Federhen, S., Feolo, M., Geer, L.Y., Helmberg, W., Kapustin, Y., Landsman, D., Lipman,
538 D.J., Madden, T.L., Maglott, D.R., Miller, V., Mizrachi, I., Ostell, J., Pruitt, K.D., Schuler, G.D., Sequeira,
539 E., Sherry, S.T., Shumway, M., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusova, T.A., Wagner, L.,
540 Yaschenko, E., Ye, J. Database resources of the National Center for Biotechnology Information.
541 Nucleic Acids Res., 37(Database issue):D5-15., 2009

Schirrmeister, L., Schwamborn, G., Overduin, P. P., Strauss, J., Fuchs, M. C., Grigoriev, M., Yakshina, I., Rethemeyer, J., Dietze, E. and Wetterich, S.: Yedoma Ice Complex of the Buor Khaya Peninsula (southern Laptev Sea), Biogeosciences Discuss., 1–36, doi:10.5194/bg-2016-283, 2016.

Sjögren, P., Edwards, M. E., Gielly, L., Langdon, C. T., Croudace, I. W., Merkel, M. K. F., Fonville, T. and Alsos, I. G.: Lake sedimentary DNA accurately records 20th Century introductions of exotic conifers in Scotland, New Phytol., n/a-n/a, doi:10.1111/nph.14199, 2016.

Stockmarr, J.: Tablets with spores used in absolute pollen analysis, Pollen et Spores, 13:615-621, 1971

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C. and Willerslev, E.: Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding, Nucleic Acids Res., 35(3), doi:10.1093/nar/gkl938, 2007.

Willerslev, E., Hansen, A. J., Binladen, J., Brand, T. B., Gilbert, M. T. P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D. A. and Cooper, A.: Diverse Plant and Animal Genetic Records from Holocene and Pleistocene Sediments, Science, 300(5620), 791–795, doi:10.1126/science.1084114, 2003.

Yoccoz, N. G., Bråthen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., Von Stedingk, H., Brysting, A. K., Coissac, E., Pompanon, F., Sønstebø, J. H., Miquel, C., Valentini, A., De Bello, F., Chave, J., Thuiller, W., Wincker, P., Cruaud, C., Gavory, F., Rasmussen, M., Gilbert, M. T. P., Orlando, L., Brochmann, C., Willerslev, E. and Taberlet, P.: DNA from soil mirrors plant taxonomic and growth form diversity, Mol. Ecol., 21(15), 3647–3655, doi:10.1111/j.1365-294X.2012.05545.x, 2012.

Sedimentary ancient DNA and pollen reveal the composition of plant organic matter in Late Quaternary permafrost sediments of the Buor Khaya Peninsula (north-eastern Siberia)

Heike Hildegard Zimmermann[1,2], Elena Raschke[1,3], Laura Saskia Epp[1], Kathleen Rosmarie Stoof-Leichsenring[1], Georg Schwamborn[1], Lutz Schirrmeister[1], Pier Paul Overduin[1], Ulrike Herzschuh[1,2]

[1]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Periglacial Research Unit, Telegrafenberg A43, 14473 Potsdam, Germany
[2] Institute of Earth and Environmental Sciences, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam-Golm, Germany
[3]Arctic and Antarctic Research Institute, Bering St. 38, 199397 St. Petersburg, Russia

Correspondence to: Heike H. Zimmermann (heike.zimmermann@awi.de), Ulrike Herzschuh (ulrike.herzschuh@awi.de)

Abstract. Organic matter deposited in ancient, ice-rich permafrost sediments is vulnerable to climate change and may contribute to the future release of greenhouse gases; it is thus important to get a better characterization of the plant organic matter within such sediments. From a Late Quaternary permafrost sediment core from the Buor Khaya Peninsula, we analysed plant-derived sedimentary ancient DNA (sedaDNA) to identify the taxonomic composition of plant organic matter, and undertook palynological analysis to assess the environmental conditions during deposition. Using sedaDNA we identified 154 taxa and from pollen and non-pollen palynomorphs we identified 86 taxa. In the deposits dated between 54 and 51 kyr BP, sedaDNA records a diverse low-centred polygon plant community including recurring aquatic pond vegetation while from the pollen record we infer terrestrial open-land vegetation with relatively dry environmental conditions at a regional scale. A fluctuating dominance of either terrestrial or swamp/aquatic taxa in both proxies allowed the local hydrological development of the polygon to be traced. In deposits dated between 11.4 and 9.7 kyr BP (13.4–11.1 cal kyr BP), sedaDNA shows a taxonomic turnover to moist shrub tundra and a lower taxonomic richness compared to the older samples. Pollen also records a shrub tundra community, mostly seen as changes in relative proportions of the most dominant taxa, while a decrease in taxonomic richness was less pronounced compared to sedaDNA. ~~Pollen also records a shrub tundra spectrum, mostly seen as changes in relative~~

proportions of the most dominant taxa, while a decrease in taxonomic richness was less pronounced compared to sedaDNA. Our results show the advantages of using sedaDNA in combination with palynological analyses when macrofossils are rarely preserved. The high resolution of the sedaDNA record provides a detailed picture of the taxonomic composition of plant-derived organic matter throughout the core and palynological analyses prove valuable by allowing for inferences of regional environmental conditions.

Keywords: paleoenvironmental DNA, trnL, metabarcoding, vegetation history, polygon, Arctic

1. Introduction

Decomposition of soil organic carbon from ancient permafrost deposits may augment ongoing global warming, but how intensely will depend on the character of the organic matter. Most soil organic carbon in permafrost was deposited during the Late Pleistocene and Holocene (cf. Schirrmeister et al., 2011a). Throughout the last glacial (Marine Isotope Stages 4 to 2, ca. 71–10.5 kyr BP), much of north-eastern Siberia was non-glaciated (Hubberten et al., 2004). The region was exposed to extremely cold conditions, which resulted in the formation of deep permafrost with an estimated thickness of 500 m (Duchkov et al., 2014) and in lowlands of polygonal tundra environments. Ice-wedge polygons developed after many cycles of wintertime frost cracking and subsequent springtime melt-water infiltration into these cracks. Ice-wedges create ridges surrounding polygonal depressions (e.g. Lachenbruch, 1962; Leffingwell, 1915; Minke et al., 2007). The cold and water-saturated conditions in the depressions caused the accumulation of organic matter in soil horizons, as decomposition rates were low (Davidson and Janssens, 2006; Hugelius et al., 2014). Over time, ice- and organic matter-rich permafrost containing massive syngenetic ice wedges was formed (Schirrmeister et al., 2011b). These Late Pleistocene deposits are called Ice Complex or Yedoma and are estimated to store 83 ± 12 Pg organic carbon (Hugelius et al., 2014), which accounts for more than 10 % of the total organic carbon pool in permafrost globally.

Permafrost is susceptible to future climate change-induced increases in ground temperatures, which can lead to active layer deepening (Romanovsky et al., 2010) and thermoerosion (Grosse et al., 2011).

When permafrost sections thaw, microbial-driven decomposition rates increase and release climate relevant greenhouse gases that further enhance climate warming in a positive feedback loop (Knoblauch et al., 2013; Wagner et al., 2007). To improve assessments of the potential greenhouse gas release from organic matter, knowledge about its composition is of great relevance as it allows for the inference of organic matter decomposability (Cornwell et al., 2008; Hobbie, 1992). Furthermore, knowledge about the conditions under which the organic matter accumulated in the past can help us understand how it accumulates today (Lyell, 1830). Recently, the sediments, cryolithology and stratigraphy of outcrops on the western coast of the Buor Khaya Peninsula in north-eastern Siberia were described, including organic carbon quantity, quality and degradability (Günther et al., 2013b; Schirrmeister et al., 2016; Stapel et al., 2016; Strauss et al., 2012, 2013, 2015) ~~(Günther et al., 2013b; Schirrmeister et al., 2016; Stapel et al., in review; Strauss et al., 2012, 2013, 2015)~~. However, the palaeobotany of the site has not been studied. It represents a more southerly example of highly degraded Ice Complex in the central Laptev region between the Lena Delta to the west and the Indigirka lowlands to the east. The closest palaeobotanical reconstruction was carried out at the north-eastern coast of the peninsula using palaeolake sediments, but only of Mid Holocene origin (Willerslev et al., 2014). Hence a record which describes the floristic composition of organic matter from the last glacial is needed to complete our understanding of the source and quality of organic carbon in these deposits.

Plant macrofossil and/or pollen are the palaeobotanical records usually used to study the permafrost soil organic matter composition and environmental conditions during its deposition. Plant macrofossils, which can often be identified to species level, mostly originate directly from former vegetation at the study site, allowing past local environmental conditions to be inferred (Birks, 2001). However, identifiable macrofossils are usually preserved in low quantities and preservation varies strongly among taxa, which hinders a quantitative exploitation of this proxy data (Kienast et al., 2001). Pollen, in contrast, is preserved in sufficient amounts to quantify its composition, but in arctic treeless ecosystems it originates from across a variable extent from extra-local to local sources (Birks, 2001; van der Knaap, 1987) and taxonomic resolution is mostly limited to genus- and family level (Beug, 2004; Moore et al., 1991). A comparatively new proxy is the analysis of sedimentary ancient DNA (sedaDNA), which originates from disseminated material within sediments (Haile et al., 2009; Rawlence et al., 2014). Its

Feldfunktion geändert

application to permafrost sediments was introduced by Willerslev et al. (2003) who reported the successful retrieval of plant DNA as old as 300,000 to 400,000 years. Since then, several studies prove that DNA is exceptionally well preserved in permafrost sedimentary archives with respect to vascular plants (Willerslev et al., 2003), bryophytes (Epp et al., 2012), fungi (Bellemain et al., 2013; Epp et al., 2012; Lydolph et al., 2005), bacteria (e.g. Wagner et al., 2007; Willerslev et al., 2004b), invertebrates (Epp et al., 2012), birds (Epp et al., 2012), and mammals (e.g. Arnold et al., 2011; Haile et al., 2009; Willerslev et al., 2003, 2014). The constantly cool temperatures of permafrost lead to reduced microbial and enzymatic degradation (Levy-Booth et al., 2007) and limited hydrolytic damage, since up to 97 % of the water is frozen (Willerslev et al., 2003, 2004a). SedaDNA is supposed to be of local origin (Boessenkool et al., 2014; Haile et al., 2007; Jørgensen et al., 2012; Parducci et al., 2013, Pedersen et al., 2016, Sjögren et al., 2016) and can be preserved extracellularly, even when macrofossil evidence is absent (Arnold et al., 2011; Willerslev et al., 2003). In comparison to arctic pollen and macrofossils, the taxonomic resolution of sedaDNA exceeds that of pollen in almost all groups of higher plants and is close to the resolution of macrofossils (Jørgensen et al., 2012; Pedersen et al., 2013; Sønstebø et al., 2010).

Feldfunktion geändert

The purpose of this study is to combine sedaDNA and pollen analyses to reconstruct past local and regional flora from a Late Quaternary permafrost sediment core, which was recovered from the western coast of the Buor Khaya Peninsula, a sparsely investigated region of the Late Pleistocene Ice Complex. The sediments contain moderate organic carbon content generally 2–5 % (Schirrmeister et al., 2016; Stapel et al., 2016; Strauss et al., 2015) ~~(Schirrmeister et al., 2016; Stapel et al., in review; Strauss et al., 2015)~~. Given that organic matter is the substrate for microbial turnover and the future release of greenhouse gases, it is crucial to get a better characterization of the plant organic matter within those sediments. At the same time it is necessary to understand the environmental conditions that prevailed when the organic matter was deposited to allow for inferences of modern and future processes under comparable environmental conditions. We therefore aim to answer the following questions: (1) What is the taxonomic composition of plant organic matter stored locally in the ancient permafrost sediments? (2) What were the environmental conditions during the time of organic matter accumulation?

2. Geographical settings

The Buor Khaya Peninsula belongs to the north-east Siberian Yana-Indigirka lowlands and is surrounded by the Buor Khaya Bay and the Yana Bay in the southern Laptev Sea (Fig. 1). It is underlain by continuous permafrost with ground temperatures of less than -10°C (Schirrmeister et al., 2016). The study area has a subarctic continental climate with a short growing season and long severe winters. Mean annual precipitation is 321.5 mm, and the mean annual air temperature is -12.8°C (mean temperature of the warmest month: 8.7°C, mean temperature of the coldest month: -32.5°C) at the closest meteorological station in Tiksi (WMO 218240) based on data between 1981 and 2010.

The treeless landscape is covered by polygonal tundra and a mosaic of thermokarst and alas depressions and Yedoma 'hills' up to 37 m above sea level (Günther et al., 2013a). Thermokarst depressions formed by permafrost thaw and ground subsidence as a consequence of increasing temperatures after the Last Glacial Maximum (Czudek and Demek, 1970), whereas Yedoma hills and uplands are relicts of Late Pleistocene accumulation plains (Schirrmeister et al. 2013). According to the Circumpolar Arctic Vegetation Map (CAVM Team, 2003) the modern vegetation cover can be classified mostly as erect dwarf-shrub tundra, in places tussock-sedge, dwarf-shrub, moss tundra and the northernmost part of the peninsula as sedge, moss, dwarf-shrub wetlands. On raised microsites, such as ridges of low-centred ice-wedge polygons, sedge, moss, dwarf-shrub wetlands also include components of non-tussock sedge, dwarf-shrub, moss tundra.

3. Material and Methods

3.1 Core material

In April 2012, an 18.9-m long core with frozen sediment (BK-8) was drilled from the top of an eroding Yedoma hill at 34 m a.s.l. (Fig. 1) located at about 100 m from the cliff edge on the western coast of the Buor Khaya Peninsula (71.420° N, 132.111° E). Detailed descriptions of the fieldwork and cryolithological properties (i.e. ground ice and sediment features) are published in Günther et al. (2013) and Schirrmeister et al. (2016). Radiocarbon dating was performed on bulk plant macro remains by accelerator mass spectrometry at the Cologne AMS laboratory, Germany (Schirrmeister et al., 2016)

and radiocarbon ages are given in Table 1. Based on field observations, core descriptions and analytical data sets, Schirrmeister et al. (2016) subdivided the core into six cryolithological units (Table 1). Absolute ice-contents varied between 66 and 84 wt %, which is typical for Yedoma (Schirrmeister et al., 2011b, 2016). Plant remains were sparsely distributed throughout the core and were composed mainly of fine rootlets, grass fragments, small woody pieces and very few seeds and fruits.

3.2 Subsampling of the permafrost core

In January 2014 the core segments were cut into two halves. One was stored as an archive and the second was subsampled. The opening of the core and the subsampling took place at the German Research Centre for Geosciences Helmholtz Centre Potsdam in the climate chamber at a temperature of -10°C, where no genetic experiments were performed either before or after. The drilling mud was removed using a band saw. Approximately 3 mm of the surface was removed with a clean straight draw knife of 225 mm length (Wilh. Schmitt & Comp. GmbH & Co. KG, Germany). Before use, the draw knife was cleaned with 5 % Deconex-solution (Th. Geyer, Germany), rinsed in purified water, followed by DNA-ExitusPlus treatment (VWR, Germany) and rinsing in purified water. Finally, the draw knife was soaked in 96 % technical Ethanol (Carl Roth GmbH & Co. KG, Germany) and flamed. Then, about 1 mm of the newly exposed sediment was removed with a small clean knife (Th. Geyer, Germany). Each knife was used for one draw only. Before use, the small knife was cleaned with 5 % Deconex-solution, rinsed in purified water, followed by DNA-Away® (Carl Roth GmbH & ca. KG, Germany) treatment, and rinsed in purified water. Finally, each side of the knife was UV-irradiated, at close distance for 10 minutes in a CL1000 ultraviolet crosslinker (UVP, USA), as recommended by Champlot et al. (2010).

The samples were drilled through the clean surface using a TCT hole saw with one tooth and an outer diameter of 25 mm (Esska.de GmbH, Germany). Before use, they were cleaned in the same manner as the small knives. The ends of the retrieved cylindrical sediment pieces were cut using sterile disposable scalpel tips. After cutting the first end, the scalpel tip was cleaned using Deconex, rinsed with purified water and Ethanol, and finally flamed to prevent contamination from one end to the other. The DNA samples were stored at -20°C and the pollen samples at 4°C. ~~The DNA samples were stored at -20°C and~~

~~the pollen samples were stored at 4°C.~~ In total, 54 samples were drilled (approximately 3 to 4 samples per metre) for each kind of analysis. All 54 samples were analysed for sedaDNA while only 32 were processed for pollen analysis. The ice-wedge segment was not sampled for DNA analyses, because it was shattered into pieces, which were too small for an intact piece from the inside to be drilled out.

3.3 Molecular genetic laboratory work

DNA isolation and PCR setup was performed in the palaeogenetic laboratory of the Alfred-Wegener-Institute Helmholtz Centre for Polar and Marine Research in Potsdam, Germany. This lab is dedicated to ancient DNA isolation and PCR setup and is located in a building devoid of any molecular genetics lab work. The lab is cleaned frequently by the researchers and subjected to nightly UV-irradiation. All laboratory work was performed in a UVC/T-M-AR DNA/RNA cleaner-box (BIOSAN, Latvia). DNA isolations and PCR setups were performed on different days using dedicated sets of pipettes and equipment. Further precautions to reduce contamination included UV-irradiation of 10x buffer, BSA, $MgSO_4$ and DEPC-treated water for 10 minutes in a UV crosslinker in thin-walled PCR reaction tubes approximately 1 cm below the UV light bulbs (similar to recommendations of Champlot et al. (2010)).

Total DNA was isolated from approximately 5 g of frozen permafrost sediment using the PowerMax® Soil DNA Isolation Kit (Mo Bio Laboratories, Inc. USA). For the initial lysis and homogenization step, 0.8 mg peqGOLD proteinase K (VWR, Germany), 0.5 ml 1M Dithiotreitol (VWR, Germany) and 1.2 ml C1 solution and samples were added to 15 ml PowerBead solution, vortexed for 10 minutes and incubated overnight at 56°C on a nutating mixer (VWR, Germany) under gentle agitation. All following steps were carried out according to the kit manufacturer's recommendations, using 1.6 ml elution buffer and extending the incubation time to 10 minutes for the final elution. One extraction blank was included for each isolation batch of 11 samples and processed in the same way as the samples.

The PCR reactions were performed with the trnL g and h primers (Taberlet et al., 2007). Both primers were modified on the 5' end by unique 8 bp tags which varied from each other in at least five base pairs to distinguish samples after sequencing (Binladen et al., 2007) and were additionally elongated by NNN tagging to improve cluster detection on the sequencing platform (De Barba et al., 2014). The PCR reactions contained 1.25 U Platinum® *Taq* High Fidelity DNA Polymerase (Invitrogen, USA), 1 x HiFi

buffer, 2 mM $MgSO_4$, 0.25 mM mixed dNTPs, 0.8 mg Bovine Serum Albumin (VWR, Germany), 0.2 mM of each primer and 3 µl DNA in a final volume of 25 µl. PCRs were carried out in a TProfessional Basic thermocycler (Biometra, Germany) with initial denaturation at 94°C for five minutes, followed by 50 cycles of 94°C for 30 s, 50°C for 30 s, 68°C for 30 s and a final extension at 72°C for 10 min. To trace possible contamination one No Template Control (NTC) was included in each PCR and treated identically to the samples and the extraction blanks. 2 % agarose (Carl Roth GmbH & Co. KG, Germany) gels were used to check if the PCR was successful and whether the products matched the expected size.

For each sample we pooled two positive amplifications for sequencing, under the condition that the associated NTCs and extraction blank were negative. ~~All samples were subjected to up to four PCR runs. Once two positive amplifications were obtained, while the associated NTC and extraction blank remained negative, they were used for sequencing.~~ The two pooled positive amplifications were ~~pooled and~~ purified using the MinElute PCR Purification Kit (Qiagen, Germany), following the manufacturer's recommendations. Elution was carried out twice with DEPC-treated ultra-purified water to a final volume of 40 µl. The DNA concentrations were estimated with the dsDNA BR Assay and the Qubit® 2.0 fluorometer (Invitrogen, USA) using 1 µl of the purified amplifications. To avoid bias based on differences in DNA concentration between samples, they were pooled in equimolar concentrations. All extraction blanks and NTCs were included in the sequencing run, using a standardized volume of 10 µl, even though they were negative in the PCRs. The sequencing results of extraction blanks and PCR controls are reported in the supplementary file S4. Library preparation and sequencing on the Illumina HiSeq platform (2 x 125 bp) were performed by the Fasteris SA sequencing service (Switzerland).

~~DNA isolation and PCR setup was performed in the palaeogenetic laboratory of the Alfred-Wegener-Institute Helmholtz Centre for Polar and Marine Research in Potsdam, Germany. This lab is dedicated to ancient DNA isolation and PCR setup and is located in a building devoid of any molecular genetics lab work. The lab is cleaned frequently by the researchers and subjected to nightly UV irradiation. All laboratory work was performed in a UVC/T-M-AR DNA/RNA cleaner box (BIOSAN, Latvia). DNA isolations and PCR setups were performed on different days using dedicated sets of pipettes and~~

~~equipment. Further precautions to reduce contamination included UV-irradiation of 10x buffer, BSA, $MgSO_4$, and DEPC-treated water for 10 minutes in a UV-crosslinker in thin-walled PCR reaction tubes approximately 1 cm below the UV light bulbs (similar to recommendations of Champlot et al. (2010)).~~

3.4 Analysis of sequence data and taxonomic assignments

The sequence quality was checked using FastQC (Andrews, 2010) (S1, Fig. S3.1). Filtering, sorting and taxonomic assignments of the sequences were performed using OBITools (Boyer et al., 2016). Forward and reverse reads were aligned to produce single sequences using *illuminapairedend*. These sequences were assigned to their samples based on exact matches to their tag-combination using *ngsfilter*, followed by *obigrep* to exclude sequences shorter than 10 bp and *obiuniq* with which duplicated sequences were merged while keeping the information to which sample the sequences originally belonged. Rare sequences occurring with less than 10 read counts across the dataset were excluded as probable artefacts using *obigrep*. Sequence variants probably attributable to PCR or sequencing errors were excluded by *obiclean* (Boyer et al., 2016).

Two reference databases were used for taxonomic assignments as described in Epp et al. (2015): the first is based on the quality-checked and curated Arctic and Boreal vascular plant and bryophyte reference libraries (composed of 1664 vascular plant and 486 bryophyte species) published by Sønstebø et al. (2010), Willerslev et al. (2014) and Soininen et al. (2015); the second is based on the EMBL Nucleotide Database standard sequence release 117 (Kanz et al., 2005; www.ebi.acuk/embl/). The sequences were assigned to taxon names based on sequence similarity to each of the reference databases using *ecotag*. The nomenclature for the taxonomic assignment follows the NCBI taxonomy (Sayers et al., 2009). When the same taxonomic names were given more than once to different sequences we attached the affix MOTU (Molecular Operational Taxonomic Unit).

To further remove noise in the dataset, sequences occurring less than 10 times in a sample were excluded using R v. 3.0.3 (R Core Team, 2014). Only sequences that displayed a best identity value of 1.0 to an entry of a reference database were kept and assigned its taxonomic name (S5~~S4~~, PANGAEA doi: follows). Sequences which were assigned to cultivated plants or those highly unlikely to occur in the Arctic were excluded from our analyses as probable contamination (S6~~5~~). These exotic DNA

sequences were detected in almost all samples but contained mostly less than 1 % of the total number of read counts with best identity of 1.0 (S6). The highest contribution of exotic DNA sequence counts within a sample was at 2.85 m depth from Musaceae. As a conservative measure we excluded sequences assigned to the PACMAD-clade (including *Muhlenbergia richardsonis*), since they are identical with *Zea mays* and authentic sequences cannot be distinguished from probable contamination. Nevertheless, the excluded sequences comprised only a small proportion (3.3 %) of all Poaceae sequences and an under-representation would thus have only a minor impact. Sequences appearing in the extraction blanks and NTCs comprised 3.4 % of all sequence counts and were also excluded (S46). In most cases these sequences were only found in the extraction blanks (55.6 %) or NTCs (33.3 %) and not in the samples, while approximately 10 % of were probably derived from wrong sample assignment through tag-jumps.

3.5 Pollen sample treatment and analysis

Thirty-two samples were selected for pollen analysis. From each sample approximately 3 g (wet weight) were taken for sample preparation. For fluid samples (unfrozen due to storage at 4°C) 1 ml of sediment was taken using a syringe. Standard preparation following Faegri and Iversen (1989) including KOH, HCl and HF treatment was used to extract pollen, spores and algae~~non pollen palynomorphs (NPPs)~~ from the sediment. For calculation of concentration of pollen a *Lycopodium* spore tablet (Batch no. 1031; n=20848 +/- 1460) was added to each sample (Stockmarr 1971). Pollen of terrestrial and aquatic plants as well as common spores and algae~~NPPs~~ were analysed using a light microscope (Zeiss Axioskop 2) under 400–600x magnification (S7). At least 300 pollen grains, spores and algae ~~NPPs~~ were identified in each sample following sample sizes in Andreev et al. (2011). Published pollen atlases (Beug, 2004; Kupriyanova and Alyoshina, 1972, 1978; Moore et al., 1991; Savelieva et al., 2013; Sokolovskaya, 1958) and a pollen reference collection at the Arctic and Antarctic Research Institute (Sankt-Petersburg) and the Alfred Wegener Institute were used for taxonomic identification of pollen and spores. ~~NPPs~~ Spores were determined according to van Geel (2001), van Geel et al. (1983), van Geel and Aptroot (2006). Freshwater algae were determined using Jankovská and Komárek (2000) and Komárek and Jankovská (2001).

3.6 Statistical analyses and visualization

Analyses for sedaDNA and pollen were carried out in the same manner. After first inspection of the data we assigned Cyperaceae to the local component (Moore et al., 1991) of the swamp and aquatic taxa to get a better picture about changes in the terrestrial sedaDNA and pollen signals. We thus separated the data of both proxies into two datasets: (1) terrestrial and (2) swamp and aquatic. Recorded bryophytes for sedaDNA and spores ~~NPPs for pollen~~ were described, but since they were sparse, they could not be analysed statistically. Due to the higher taxonomic resolution of the sedaDNA, *Kobresia* (Cyperaceae) remained in the terrestrial sedaDNA set, while Poinae – either *Arctophila fulva* or *Dupontia fisheri* – were included in the swamp and aquatic set as both can occur within low-centred polygons (Aiken et al., 2007). Rarefaction curves were produced and rarefied taxon richness calculated using *rarecurve* and *rarefy*, to compare taxon richness at a particular number of sequences or pollen grains, determined by the lowest number of retrieved sequences or pollen grains among all samples (Heck et al., 1975; Hurlbert, 1971). This was performed for the terrestrial and swamp and aquatic datasets separately.

For statistical analyses only taxa that were present in at least six samples in the terrestrial and in three samples in the taxonomically-poorer swamp and aquatic datasets were included. Relative proportions of taxa within a sample were calculated on the basis of their sequence count or pollen sum in each dataset. A double square-root transformation was performed on the relative proportions of the sedaDNA dataset to mitigate the effect of overrepresented and rare sequences, while square-root transformation was applied to the pollen dataset, since differences between counts were not as pronounced. A constrained hierarchical clustering approach (CONISS (Grimm, 1987)) was performed with clusters constrained by depth using *chclust*. The relative proportions were thereafter transformed to Euclidean distances by *vegdist*. Zoning was guided by the broken-stick model (Bennett, 1996; MacArthur, 1957) using *bstick*, but with the condition that a minimum of five samples was necessary to assign a zone for the DNA datasets to avoid inflation of several very small zones. For pollen analysis a minimum of four samples was necessary to assign a zone, since fewer samples were taken compared to the DNA dataset. The stratigrams were plotted for each dataset separately with *strat.plot*.

Ratios of terrestrial to swamp and aquatic taxa and Poaceae to Cyperaceae were built (Mensing et al., 2008) to assess which contributed most to a sample, following Eq. (1) and Eq. (2):

$$\text{ratio} = \frac{\text{terrestrial} - \text{aquatic}}{\text{terrestrial} + \text{aquatic}}, \quad (1)$$

$$\text{ratio} = \frac{\text{Poaceae} - \text{Cyperaceae}}{\text{Poaceae} + \text{Cyperaceae}}, \quad (2)$$

For this, the sums of sequence counts or pollen grains of (1) all terrestrial, (2) all swamp and aquatic taxa, (3) Poaceae and (4) Cyperaceae in a sample were built. The ratios range between -1.0 and 1.0 with negative values indicating dominance of swamp and aquatic or Cyperaceae sequences or pollen grains and positive values indicating dominance of terrestrial or Poaceae sequences, while a value of zero indicates an equal contribution of both.

A Principal Component Analysis (PCA) was applied on the double square-root transformed relative proportions for sedaDNA and square-root transformed relative proportions for pollen using *rda* in order to portray the major structure in the multivariate dataset. Loadings for PC1 and PC2 axes were extracted and visualized in biplots. For better visibility, only terrestrial taxa that explained the most are plotted (n = 20 for sedaDNA; n = 25 for pollen).

The statistical analyses were performed in R v. 3.0.3 (R Core Team 2014) using the packages "vegan" (Oksanen et al., 2011), "rioja" (Juggins, 2012) and "analogue" (Simpson, 2007; Simpson and Oksanen, 2016). The age-depth model was built using the Bacon package (Blaauw and Christen, 2011) in R. The datasets are deposited at https://www.pangaea.de/ (doi: follows after acceptance).

4. Results

4.1 SedaDNA

4.1.1 SedaDNA of terrestrial plants

In total, 7,238,506 sequence counts were assigned to 113 terrestrial plant taxa, of which 40 were identified to species level, 49 to genus level, 22 to sub-tribe or family level, and one to order level.

Rarefaction analysis to compare richness on a similar level of sampling effort is based on the minimum number of observed sequence counts (n = 1064, 14.85 m depth). All curves of samples reach an asymptotic state indicating a sufficient sampling effort in all samples (Fig. S3.2). The complete taxa-list is available in ~~S3~~S5.

No (authentic) tree taxa were detected, although Saliceae, *Betula* and *Alnus* occurred there as shrub form. 64.5 % of all terrestrial sequence counts are from shrubs (Saliceae, *Betula*, *Alnus*), while 18.1 % are from graminoids and 17.4 % ~~herbs~~forbs. The dataset is particularly dominated by Saliceae sequences, which comprise 64 % of all terrestrial sequence counts, followed by Pooideae (13 %), Agrostidinae MOTU1 (3 %), *Ranunculus* MOTU1 (3 %) and *Betula* (3 %).

According to the CONISS results the terrestrial plant sedaDNA dataset can be divided into three zones (TerrDNA Zone) (Fig. 2). TerrDNA Zone 1 (18.48–14.85 ~~10.15~~ m) displays low to high rarefied richness ranging from 2–26 taxa, with a median of 11 taxa. The lowest richness is found among the deepest samples from 14.85 m downwards. TerrDNA Zone 1 is dominated by Saliceae ranging from 22–99.8 % (~~with a~~ median ~~of~~ = 91 %) accompanied by low proportions of Pooideae up to 24 % (~~with a~~ median =~~of~~ 1 %). The most frequent ~~herbaceous~~forb and graminoid taxa are *Dryas* (up to 20 %), *Festuca* MOTU1 (up to 19 %), Anthemideae MOTU1 (up to 7 %) and *Potentilla* (up to 6 %).

TerrDNA Zone 2 (~~9.9~~14.55–3.3 m) frames the ice-wedge (8.35–3.4 m) and exhibits the highest rarefied taxonomic richness, especially among graminoids, while ~~herbs~~forbs range from 6–24 taxa, median 17. Similar to TerrDNA Zone 1, Saliceae sequences dominate ranging from 0–83 %~~,~~ (median = 60 %) while Anthemideae (up to 50 %), Pooideae (up to 56 %), Poeae MOTU1 (up to 36 %), *Potentilla* (up 31 %) and *Dryas* (up 13 %) have the highest proportions among graminoids and ~~herbs~~forbs. The sample at 3.3 m depth is the only sample of this zone above the ice-wedge and shows an intermediate composition between TerrDNA Zone 2 and 3. It has lower richness in comparison to the samples below the ice-wedge, but contains sequences of Anthemideae, *Oxytropis*, *Festuca* MOTU2 and *Plantago*, which are absent in TerrDNA Zone 3. However, it is the only sample of this zone to record *Betula*, although at a lower proportion than samples of TerrDNA Zone 3.

TerrDNA Zone 3 (2.85–0.1 m) displays a low to intermediate taxonomic richness ranging from 2–13 taxa~~,~~ (median = 8) ~~taxa~~. Several graminoid (e.g. Poeae MOTU1, *Festuca*, *Pleuropogon sabinei*) and

herbaceousforb taxa (e.g. Anthemideae, *Potentilla*, *Hulteniella integrifolia*) are not recorded. Characteristic is the presence of *Betula* sequences with proportions ranging from 0–63 % (with a median =of 2 –%). Pooidae (up to 96 %, median = 16 %) and Saliceae (up to 86 %, median = 8 %) sequences dominate with median proportions of 16 % and 8 %, respectively. Taxonomic composition is further characterized by Agrostidinae MOTU1 and 2 (up to 100 % in the uppermost sample), *Ranunculus* MOTU1 (up to 74 %), *Parrya* (up to 44 %), *Pyrola grandiflora* (up to 26 %), *Delphinium* (up to 25 %), Ranunculaceae (up to 22 %) and Asteraceae (up to 20 %).

The PCA biplot of the first two axes (Fig. 3), jointly explaining 41.5 % of the variance in the dataset, reveals the major structure of the terrestrial sedaDNA. Along PC1 Saliceae and *Potentilla* exhibit the highest loadings of 0.67 and 0.63, respectively, while *Betula* and Agrostidinae MOTU1 have the lowest negative loadings of -0.54 and -0.46, respectively. Along PC2 Saliceae and *Potentilla* have the highest loadings of 0.52 and 0.23, while Pooideae and Poeae MOTU1 show the lowest negative loadings of -1.14 and -0.53, respectively. Samples of TerrDNA Zone 1 are placed mostly in the upper right quadrant and are partly intermixed with those of TerrDNA Zone 2, which are mostly placed in the lower right quadrant. TerrDNA Zone 3 samples are located in the upper and lower left quadrants and intermix only with the sample from 3.3 m.

4.1.2 SedaDNA of swamp and aquatic plants

In total, 4,591,277 sequence counts were assigned to 21 swamp or aquatic taxa, of which five were assigned to species level, 15 to genus level and one to sub-tribe level. The most dominant sequences of the dataset are *Carex aquatilis* comprising 66 % of all sequence counts, followed by *Eriophorum* MOTU1 (11 %), *Carex* MOTU1 (8 %), *Caltha palustris* (5 %) and Poinae (*Arctophilla fulva* or *Dupontia fisheri*) (4 %). Rarefaction is based on the minimum number of 20 sequence counts at 1.2 m depth. The rarefaction curves for all samples reach an asymptotic state indicating sufficient sampling effort in all samples (Fig. S3.3).

The swamp and aquatic plant DNA dataset is divided into two zones (Fig. 4). AquaDNA Zone 1 (18.48–3.3 m) comprises all samples below the ice-wedge plus a single sample above the ice-wedge at

3.3 m. The rarefied taxonomic richness of this zone ranges between zero and five with a median of one. In the deepest part of this zone (14.85–18.48 m) three out of ten samples exhibit aquatic sequences from a single taxon with small numbers of sequence counts (up to 85). AquaDNA Zone 1 is dominated by *Caltha palustris* up to 97 %, *Carex* MOTU1 up to 100 % and *Carex aquatilis* up to 100 %. Hydrophytes such as *Stuckenia* and *Hippuris* are characteristically present with proportions up to 63 % and 4 %, respectively.

AquaDNA Zone 2 (2.85–0.1 m) comprises all samples above the ice-wedge, except for the sample at 3.3 m depth. This zone exhibits a lower rarefied taxonomic richness than AquaDNA Zone 1 (Fig. 4). Samples between 2.4 and 1.2 m display high proportions of *Equisetum* up to 100 % ~~with a~~(~~-~~median ~~=of~~ 54 %) followed by *Eriophorum* MOTU1 up to 99 %~~,~~ (median = 12 %). Near surface parts at 1.1–0.1 m depth are dominated by Poinae up to 50 %, *Carex* MOTU1 up to 45 % and *Tephroseris* up to 17 %. Hydrophytes are not present whereas *Caltha palustris* is present at 0.55 m with a proportion of less than 1 %. Sedges are only represented by *Carex aquatilis* and *Carex* MOTU1.

The PCA biplot of the first two axes (Fig. 5), jointly explain 48.79 % of the variance in the dataset. Along PC1 *Caltha palustris* and *Carex aquatilis* demonstrate the highest loadings of 0.86 and 0.85, respectively, while *Equisetum* exhibits the only negative loading of -1.11. Along PC2 Poinae and *Caltha palustris* exhibit the highest loadings of 0.91 and 0.62, respectively, while *Carex* MOTU1 and *Carex aquatilis* have the lowest negative loadings of -0.79 and -0.46, respectively. AquaDNA Zone 1 samples are located mostly in the right quadrants and partly intermix with those of AquaDNA Zone 2, which are mostly located in the left quadrants.

4.1.3 SedaDNA of bryophytes and algae

In total, 8,482 sequence reads were assigned to 19 bryophyte taxa and one alga (Table 2), which constitutes 0.07 % of all sequences. Seven taxa were assigned to species level, including the freshwater alga *Cosmarium botrytis* (John et al., 2002), four to genus level, five to family level and four to order level. The observed number of taxa varies between one and six. 82.31 % of all cryptogam sequence reads are assigned to *Sphagnum* (37.39 %) and *Sphagnum russowii* (44.92 %) at 3.3 m; the depth at which the highest number of bryophyte taxa is recorded.

4.2 Pollen

4.2.1 Pollen of terrestrial plants

A total of 8,580 terrestrial pollen grains were counted and ascribed to 53 taxa, while 248 were indeterminable and 53 assigned as pre-Quaternary pollen. ~~A total of 8,881 terrestrial pollen grains were counted and ascribed to 53 taxa, including indeterminable and pre-Quaternary pollen.~~ Seven taxa were assigned to species type or section level, 24 taxa to genus level, 21 to family level, and one to order level. Rarefaction is based on the minimum number of 137 pollen counts at 11.7 m depth. The rarefaction curves did not reach saturation implying that the sampling effort was insufficient to display the sample diversities (Fig. S3.4, Fig. S3.5).

The core is divided into three terrestrial palynological zones (Fig. 6). TerrPZ 1 (18.48–17.2 m) displays the highest rarefied richness ranging from 18 to 21 taxa (~~, with a~~ median ~~=of~~ 20)~~ taxa~~. This zone is distinguished by the presence of pollen from shrubs, especially *Betula* sect. *Nanae* with proportions up to 15 %, *Alnus fruticosa*-type up to 8 % and *Salix* up to 11 %. Poaceae with proportions up to 50 %, Brassicaceae up to 10 % and *Artemisia* up to 7 % dominate among the graminoid and ~~herbaceous~~forb taxa. Pollen from trees comprise, among others, *Pinus Haploxylon*-type up to 3 % and *Larix* up to 2 %. Furthermore, the major proportion of redeposited, pre-Quaternary pollen can be found in this zone, while total pollen concentration is lowest (Table 3).

For TerrPZ 2 (17.05–8.75 m) relatively high taxonomic richness is registered ranging from 15–26 taxa, median 19 taxa. This zone is characterized by a decrease in pollen from trees and shrubs with decreasing depth and a dominance of ~~herbaceous~~forb and graminoid taxa, especially Poaceae with proportions up to 79 %. *Artemisia*, Asteraceae, Caryophyllaceae, Brassicaceae, Ranunculaceae and Saxifragaceae dominate among other ~~herbs~~forbs. Furthermore, a remarkable proportion of cf. *Dryas* of up to 2 % is seen in this zone.

TerrPZ 3 (2.85–0.1 m) displays a lower rarefied richness of 11–19 taxa (~~,~~ median = 14)~~ taxa~~. This zone is dominated by *Betula* sect. *Nanae* (up to 54 %), *Salix* (up to 38 %) and Poaceae (up to 51 %). In the samples of 0.85 m and 1.3 m large proportions of *Salix* are present of 30 % and 38 %, respectively. The

proportion of *Artemisia* decreases and ranges between 1 % and 2 %. The uppermost sample is characterized by large increases of *Alnus fruticosa*-type (up to 17 %) and Ericales (up to 18 %) reflecting the recent pollen spectrum.

The first two axes in the biplot (Fig. 7), jointly explain 55 % of the variance in the dataset. Along PC1 Poaceae and *Artemisia* exhibit the highest loadings of 0.44 and 0.23, respectively, while *Betula* sect. *Nanae* and *Salix* have negative loadings of -0.94 and -0.17, respectively. Along PC2 the highest positive loadings belong to Poaceae and *Salix* with 0.31 and 0.28, respectively, while Brassicaceae and *Alnus fruticosa*-type exhibit the lowest negative loadings of -0.37 and -0.31, respectively. The upper right quadrant includes mostly shrub taxa, the upper left quadrant mostly coniferous trees and the lower quadrants mostly open-land taxa. Samples of TerrPZ 1 plot mostly in the lower left and right quadrants, samples of TerrPZ 2 plot in the upper and lower left quadrants and samples of TerrPZ3 plot in the upper and lower left quadrants, without intermixing.

4.2.3 Pollen and spores of swamp and aquatic plants

A total of 2,816 aquatic pollen grains and spores were counted and ascribed to eight taxa, of which six were assigned to genus level and two to family level. The numbers of aquatic pollen grains are low and vary between 15 and 220 per sample. Rarefaction curves, based on the minimum number of pollen grains (n = 15; 16 m depth), do not reach saturation, implying insufficient sampling effort and an underestimate of the sample's diversity (Fig. S3.6). The broken-stick model recommends three zones, but under the condition that each zone must consist of at least four samples, two aquatic palynological zones are assigned (Fig. 8).

AquaPZ 1 (18.48–16.0 m) consists of the deepest part of the core. Rarefied richness ranges from one to two with a median of two. Cyperaceae dominate with proportions of 73–96 % (~~and a~~ median ~~=of~~ 92 %). The sum of proportions from other taxa accounts for less than 10 %, except for the sample at 16.0 m, which was co-dominant with *Equisetum* at 27 %, while the sample at 17.0 m is co-dominant with cf. *Potamogeton* at 13 %.

AquaPZ 2 (15.2–0.1 m) comprises nearly the whole core. Rarefied richness ranges from one to two with a median of one. The proportions of Cyperaceae range from 92 % up to 100 % (~~with a~~ median =~~of~~ 98 %). The sum of all further taxa account for less than 10 % in each sample.

The first two axes in the biplot (Fig. 9), jointly explain 75 % of the variance in the aquatic pollen dataset. *Equisetum* displays the only positive loading along PC1 of 0.67, while cf. *Potamogeton* has the lowest negative loading of -0.42. Along PC2 cf. *Potamogeton* has the highest positive loading of 0.58 and Cyperaceae has the lowest negative loading of 0.17. Samples from AquaPZ 1 are located in the upper quadrants of the biplot while most of the samples from AquaPZ 2 are located in the lower quadrants.

4.2.3 Spores and algae~~Non-pollen palynomorphs~~

A total of 1,092 ~~NPPs~~ spores and algae were counted and assigned to 25 taxa, comprising four mosses, two spikemosses, six clubmosses, three ferns, six fungi and four green algae (Table 3). In samples below the ice-wedge the most abundant NPPs are green algae, especially *Pediastrum*, ~~*Botrycoccus*~~*Botryococcus* and *Zygnema*-type and Bryales. The most abundant fungal spores are *Sporormiella* and *Glomus*, with *Glomus* being only recorded in the deepest parts of the core from 14.65 m downwards~~, except for~~as well as in the sample at 8.9 m. Similar to the redeposited pollen the major proportion of pre-Quaternary spores can be found between 17.05 m and 18.15 m. Above the ice-wedge *Gelasinospora*, *Zygnema*-type, ~~*Botrycoccus*~~*Botryococcus*, Bryales and *Sphagnum* are relatively abundant.

4.3 Ratios of terrestrial to swamp and aquatic taxa and Poaceae to Cyperaceae

Ratios were built for the sedaDNA and pollen datasets to assess whether terrestrial or swamp and aquatic taxa contributed more sequences or pollen grains to a sample and were compared to the corresponding Poaceae-Cyperaceae ratios. This allowed us to trace local hydrological changes and to identify drier (positive values) and wetter phases (negative values) (Fig. 10). Generally, sedaDNA and pollen show similar trends for both ratios, with an exception between 11.7 m and 12.12 m depth. The Poaceae-Cyperaceae ratio of the sedaDNA exhibits highly fluctuating ratios across the core and mostly

follows the pattern of the terrestrial-aquatic ratio. Eighteen sedaDNA samples are dominated by swamp and aquatic taxa: four above the ice-wedge and 14 below. The pollen ratios show more moderate fluctuations and only two samples are dominated by swamp and aquatic taxa at 11.7 m and 12.12 m depth. However, in six samples Cyperaceae dominate over Poaceae, one at 2.85 m and five between 11.7 m and 14.65 m, whilst samples between 13.0 m and 14.0 m have equal contributions of Poaceae and Cyperaceae pollen.

5. Discussion

5.1 Quality and proxy value of sedaDNA and pollen data

All samples from the BK-8 sediment core contained plant-derived DNA and pollen. The two proxies are known to complement each other (e.g. Jørgensen et al., 2012) and differences in the obtained data result mostly from the spatial scale at which sedaDNA and pollen originate (local vs. regional signal) and technical biases, which lead to variations in the taxonomic richness, the level of taxonomic resolution and the strength of taphonomic processes in both proxies.

The different spatial scales of sedaDNA (local) and pollen (local to extra-regional) records are an important aspect of the differences identified in the taxa spectra and thus indicate complementarity rather than direct comparability of the proxies. ~~Our results agree with several other studies which demonstrate first, that sedaDNA originates mainly locally (Haile et al., 2007, 2009; Jørgensen et al., 2012; Parducci et al., 2013; Yoccoz et al., 2012) and, second, that it is predominantly derived from roots and other plant parts rather than from pollen (Jørgensen et al., 2012; Levy-Booth et al., 2007; Willerslev et al., 2003).~~ In most of the samples we did not detect conifer-derived sedaDNA, although they are present in the pollen record. Hence, the Pinaceae pollen presumably originated from extra-regional stands (Birks, 2001; van der Knaap, 1987). Furthermore, the applied sedaDNA marker is located on the chloroplast genome, which is transmitted through pollen in Pinaceae (reviewed in: Mogensen, 1996). If pollen contributed significantly to the sedaDNA record, we would expect to find it at least in samples with high Pinaceae proportions, which we did not. This supports two assumptions about sedaDNA: first, that sedaDNA originates mainly locally (Haile et al., 2007, 2009; Jørgensen et al., 2012; Parducci et al.,

2013; Pedersen et al., 2016; Sjögren et al., 2016; Yoccoz et al., 2012) ~~(Haile et al., 2007, 2009; Jørgensen et al., 2012; Parducci et al., 2013; Yoccoz et al., 2012),~~ and, second, that it is predominantly derived from roots and other plant parts rather than from pollen (Jørgensen et al., 2012; Levy-Booth et al., 2007; Parducci et al., 2013; Pedersen et al., 2016; Sjögren et al., 2016; Willerslev et al., 2003). ~~(Jørgensen et al., 2012; Levy-Booth et al., 2007; Willerslev et al., 2003). The sedaDNA record does not contain conifer-derived DNA (Birks, 2001; van der Knaap, 1987), although they display high proportions in the pollen record, presumably from long-distance-dispersed pollen.~~ Overall, we find a steady dominance of Saliceae (which we interpret as *Salix*), Poaceae and Cyperaceae sequences across all samples of the core. This is likely caused by the huge below-ground biomass of these taxa in tundra environments, which can far exceed the above-ground biomass. According to Iversen et al. (2015), the ratio of below-~~ground~~ to above-ground biomass in tundra is highest for sedges and grasses, followed by shrubs, and is lowest for forbs. ~~They also showed that these ratios are larger in tundra compared to other biomes such as forests (Iversen et al., 2015). Accordingly Yoccoz et al. (2012) propose that a higher biomass above ground than below ground led to the under-representation of *Salix* in comparison to other plant functional types in the soil of boreal environments.~~ Since *Salix* is also found throughout the pollen record, we assume that it was locally present throughout the investigated time frame ~~and is not only a result of *Salix* roots, which can grow down to depths of 30 to 40 cm (Iversen et al., 2015). Hence, our results~~ This further supports the general view that sedaDNA mainly presents a local signal.

Technical and taphonomic biases of pollen data are well known. For example, standard pollen sample preparation, as applied in this study, may (partly) destroy *Luzula* and *Larix* pollen grains (Moore et al., 1991). Hydrophytes are largely under-represented in the pollen dataset when compared to the sedaDNA results, which may be caused by low pollen production or insufficient sedimentation, as pollen from e.g. *Potamogeton*~~, for example,~~ tends to float on the water surface for pollination (Cox, 1988; Preston and Croft, 1997). While t~~T~~aphonomic biases in sedaDNA are still not well understood and part of ongoing research, especially for lake sediments (Alsos et al., 2015), t~~T~~he technical biases of sedaDNA ~~during DNA isolation, PCR and sequencing~~ are ~~also~~ known and have been reviewed in ~~(~~Hansen et al.~~,~~ (2006)~~,~~ Schnell et al.~~,~~ (2015)~~;~~ and Thomsen and Willerslev~~,~~ (2015). ~~In this study, w~~We found an inflation of unique sequence types, attributable to PCR-errors, across the whole dataset, except for those ~~taxa~~ which

Feldfunktion geändert
Formatiert: Deutsch (Deutschland)
Formatiert: Deutsch (Deutschland)

we deemed as possible contamination ~~(cultivated and non-Arctic plants)~~. Otherwise, our dataset shows low probabilities for erroneous base-calls indicated by high sequencing qualities. We therefore assume that the taxa included in the analyses are authentic ~~and the PCR errors are caused by template age~~. A more detailed technical evaluation can be found in S2. ~~Taphonomic biases in sedaDNA are still not well understood and part of ongoing research, especially for lake sediments (Alsos et al., 2015).~~

Compared to the number of vascular plant taxa (58) and bryophytes (4) recorded by pollen analysis, the sedaDNA approach recorded a higher number of both vascular plants (134) and bryophytes (20). ~~Compared to the number of vascular plant taxa (58) and bryophytes (20) recorded by pollen analysis, the sedaDNA approach recorded a higher number of vascular plants (134) and a similar number of bryophytes (20). After~~ Next to technical biases and taphonomy, the lower number of taxa recorded by pollen can be explained by ~~two major reasons:~~the sampling effect and the taxonomic resolution. ~~First,~~ T~~t~~he counts of sequences and pollen grains differed by several orders of magnitude, which is reflected in rarefaction curves of sedaDNA reaching saturation while those of pollen do not. This indicates that our sequencing depth was adequate for sedaDNA. The number of pollen counts was guided by pollen records for the Laptev region usually ranging between 100 and 600 counts (Andreev et al., 2011). However, for future studies a higher sampling effort should be considered. ~~to mitigate the effect of the low taxonomic richness~~. ~~Second, t~~The recorded richness also depends on the taxonomic resolution, which ~~. For sedaDNA the resolution~~ depends on the marker employed for sedaDNA. ~~In our study, t~~The resolution of the *trn*L P6 loop marker (Taberlet et al., 2007) allowed ~~made it possible~~ to assign 78 % of the retrieved sequences to species or genus level, while 71 % of pollen, ~~and~~ spores and algae were identified to a similar taxonomic level. This is in the range of~~similar to~~ other sedaDNA studies focusing on Arctic vegetation from permafrost sediments (Taberlet et al., 2007: 90 % up to genus level, Sønstebø et al., 2010: 83 % for the oldest and 68 % for the youngest sample, Jørgensen et al., 2012: 81 %, Willerslev et al., 2014: 80 %). SedaDNA analysis was able to resolve 21 sequence types from Poaceae and Cyperaceae on the species or genus level. Although some of the sequence types within these families cannot be resolved~~distinguish between 18 sequence types from Poaceae and 12 from Cyperaceae, although several closely related taxa within Poaceae and Cyperaceae and especially Salicaeae, could not be distinguished because there is not enough~~ due to insufficient ~~interspecific~~

variation in the marker ~~target~~ region~~. Most of these closely related taxa, especially in the Poaceae and Cyperaceae,~~ they cannot be distinguished by pollen analysis either ~~and the resolution is often restricted to family level~~ (Birks, 2001). A higher resolution would ~~be able to~~ provide a better estimate of the taxonomic richness and thus greater insight into local environmental conditions~~ since different species of the same genus can have different habitat preferences~~. Additionally, we found that bryophytes are highly under-represented in our datasets, despite being reported to be highly abundant, diverse and functionally very important members of modern polygonal landscapes (Zibulski et al., 2013, 2016). Epp et al. (2012) developed a ~~more specific~~ marker for bryophyte metabarcoding with an approximately 10 % higher resolution than the *trn*L P6 loop~~marker used in this study~~. Their marker, however, ~~suffered from~~had a low amplification success rate in late Pleistocene samples. They highlighted two probable causes, which ~~may also~~might hold true for the *trn*L P6 marker. First, the main bryophyte biomass is typically found above-ground, whereas roots are suspected to contribute the majority of vascular plant DNA in soil (Levy-Booth et al., 2007; Willerslev et al., 2003; Yoccoz et al., 2012), and second, the presence of secondary metabolites may increase DNA degradation rates after cell lysis (e.g. Xie and Lou, 2009).

~~In addition to sequence counts, richness in the~~Lastly, sedaDNA richness ~~sequence data~~ depends ~~highly~~ on the completeness of the reference database ~~used to assign the taxonomic name~~ (Jørgensen et al., 2012; Parducci et al., 2013; Pedersen et al., 2013). By using the Arctic-Boreal reference database (Soininen et al., 2015; Sønstebø et al., 2010; Willerslev et al., 2014) we were able to increase the resolution for many taxa, plus the EMBL database allowed us to examine our sequences for possible contamination by food or cultivated plants.

5.2 Environmental conditions during the pre-LGM (54–51 kyr BP, 18.9–8.35 m) and composition of deposited organic matter

The major part of the core below the ice-wedge encompasses sediments deposited before the LGM. According to sedaDNA and pollen this part can be divided into two zones; however the boundary between these zones slightly differs between the terrestrial and swamp and aquatic datasets and lies between 15 m and 17 m. The zonation matches the major structure presented in the PCA biplots of sedaDNA and pollen. The pollen record from the whole core portrays an open landscape at the regional

scale. Below the ice wedge the core shows fluctuating ages associated with high standard deviations, which can be explained either by the radiocarbon dating method being at its limit or by reworking of the sediments.

The deepest part of the core from 18.9 m to ~16 m ~~matches well with the deepest cryolithological unit presented in Schirrmeister et al. (2016). S~~sedaDNA reflects a local terrestrial flora with low taxonomic richness, comprising taxa such as *Plantago*, *Puccinellia* and *Potentilla,* with *Carex aquatilis* as the only wetland plant retrieved at 18.15 m depth. The pollen record of this zone is characterized by high proportions of shrub, Poaceae and tree pollen, but with low proportions of *Larix*. As *Larix* pollen has a very limited dispersal capacity owing to its size, weight and low quantity (Niemeyer et al., 2015; Sjögren et al., 2008), *Larix* stands are inferred to have been in the regional vicinity of the coring site, ~~and~~ possibly even closer than they are today. In contrast to sedaDNA, high proportions of *Pediastrum*, ~~*Botrycoccus*~~*Botryococcus* and *Zygnema*-type algae and highest proportions of *Potamogeton* pollen overall imply the presence of a shallow pond (Andreev et al., 2002; Kienast et al., 2005). Only in this zone do the sedaDNA and pollen records show such distinct differences. In the deepest 2.5 m the amount of redeposited pre-Quaternary pollen and spores is highest (up to ~5%) among all samples. According to the sedimentary and hydrogeochemical results presented in Schirrmeister et al. (2016), this part of the core implies an ancient active layer. An active layer is prone to disturbances such as erosion, cryoturbation (through seasonal thawing and refreezing) and potentially also grazing, all of which allow for redeposition of older material.

At depths from approximately 16 m until the ice-wedge at 8.35 m, sedaDNA and pollen reveal high taxonomic richness. SedaDNA portrays high diversity among grasses and forbs~~, a rich herbaceous flora~~ including swamp and aquatic taxa and high proportions of *Salix*. In several samples the sedaDNA record is dominated by Cyperaceae and other swamp and aquatic taxa, with mostly negative values on the ratio plot, especially for the sedaDNA dataset. This indicates~~ing~~ wet conditions on a local scale and probably enhancing organic matter accumulation, as cold and anoxic conditions reduce decomposition rates (Davidson et al., 2000). These findings are supported by the sedimentary, hydrogeochemical (Schirrmeister et al., 2016) and biomarker analyses (Stapel et al., 2016)~~(Stapel et al., in review)~~ performed on the same core. Schirrmeister et al. (2016) and Stapel et al. (2016)~~Stapel et al. (in review)~~

identified less decomposed organic matter at depths of 10 m and between 11.2 m and 15 m from higher total organic carbon (TOC) content, higher hydrogen index, lower $\delta^{13}C$ values and high concentrations of branched Glycerol dialkyl glycerol tetraether (br-GDGTs, microbial membrane compounds). The taxonomic composition of sedaDNA comprised typical taxa of low-centred polygonal depressions such as *Stuckenia*, *Hippuris* and *Caltha palustris* indicating the presence of a shallow pond (Kienast et al., 2008). Intermittently, the absence of hydrophytes and increasing proportions of Poinae (*Arctofila fulva* / *Dupontia fisheri*) and Cyperaceae indicate times without a pond and hence point towards temporal fluctuations in the hydrology of the depression. The temporal scale on which these fluctuations occurred however cannot be assessed due to cryoturbation in the sediments and the large uncertainties of the dating results.

As shown by recent studies of low-centred polygons, microtopographical differences resulting from the moisture gradient between the dry uplifted ridge and the wet depression shape the local plant community (de Klerk et al., 2009, 2011; Zibulski et al., 2016). According to the sedaDNA results, *Carex* probably occupied the major part of the polygon, whereas the ridge was likely covered by *Salix* ~~shrubs~~ along with Poaceae and ~~herbs~~forbs (de Klerk et al., 2009, 2011, 2014, Minke et al., 2007, 2009; Teltewskoi et al., 2016). The transitional zone from the ridge to the depression is characterized by an increase in moisture and was probably occupied by taxa such as *Carex*, *Eriophorum*, ~~*Saxifraga hirculus*,~~ *Comarum* and *Pedicularis* (de Klerk et al., 2009, 2014; Savelieva et al., 2013; Zibulski et al., 2016). However, vegetation surveys along transects through modern low-centred polygons and temporal reconstructions from short cores (de Klerk et al., 2009, 2011, 2014, Minke et al., 2007, 2009; Teltewskoi et al., 2016; Zibulski et al., 2016) show high proportions of ~~Ericales (e.g.~~ *Vaccinium*~~ spp.~~, *Ledum palustre*, *Empetrum nigrum*~~)~~ and *Betula nana* ssp. *exilis* and therefore display a different ~~plant assemblage to~~composition of Holocene polygons in comparison to our findings. This indicates that Holocene polygon mires ~~may~~ might differ from those of the pre-LGM.

The palynological ~~pollen and NPP~~ record in this zone (16–8.35 m) comprises *Potamogeton* pollen, ~~and~~ *Pediastrum*, ~~*Botrycoccus*~~*Botryococcus* and *Zygnema*-type algae, which supports ~~the sedaDNA results and~~ the presence of a shallow pond (Andreev et al., 2002; Kienast et al., 2005). ~~Taxa such as~~ *Artemisia*, *Dryas* and~~,~~ Papaveraceae ~~and *Saxifraga*~~ indicate overall dry environmental conditions with probably

more severe winters than today, while *Potamogeton* indicates warmer summers (Kienast et al., 2001, 2005). Annual precipitation of less than 250 mm and rapidly falling temperatures in winter must have occurred to allow thermal cracking of the soil to keep the active layer sufficiently shallow for the formation of ice-wedges and ridges that enclose low-centred polygons (Minke et al., 2007). Low relative pollen proportions of trees ~~(also derived by long distance dispersal~~ from extra-regional stands ~~(Birks, 2001; van der Knaap, 1987))~~ and shrubs with high proportions of grasses, sedges and ~~herbs~~forbs are consistent with other published pollen and macrofossil analyses in this time interval and region (Andreev et al., 2011; Kienast et al., 2001; Sher et al., 2005). The recorded pollen spectra~~um~~ from our core furthermore tally~~ies~~ with studies from~~–~~ the central Laptev region~~Bol'shoy Lyakhovsky island, Bykovsky Peninsula and Kurungnakh Island~~, in which ~~present~~ pollen records are dominated by Cyperaceae and Poaceae with a constant presence of *Salix* and high abundances of *Artemisia* and Caryophyllaceae for 55 to 40 kyr BP (Andreev et al., 2011 and references therein). The decreasing proportions of *Larix* pollen with decreasing depth presented here may point towards a ~~southward~~ retreat of *Larix* stands or a reduction in pollen productivity through unfavourable environmental conditions.

5.3 Environmental conditions during the post-LGM (11.4–9.7 kyr BP (13.4–11.1 cal kyr BP)) and composition of deposited organic matter

The permafrost core was drilled at the top of a Yedoma hill (Schirrmeister et al. 2016). Wind and rain probably eroded most of the Holocene deposits resulting in a hiatus between the sample of the modern core top and the second sample at 0.25 m depth (11.1 cal kyr BP). The upper part of the core consists of sediments dated to the transition from the late glacial to the early Holocene (13.4–11.1 cal kyr BP). As emphasized in Andreev et al. (2011) records of the late glacial transition are rare because of active thermoerosion. Hence, our results provide valuable information about the vegetation history in this region and organic matter composition. ~~The upper part of the core from the ice wedge up to approximately 0.25 m depth includes the late glacial transition to the early Holocene (approximately 11.4 to 9.7 kyr BP, (13.4–11.1 cal kyr BP)). As emphasized in Andreev et al. (2011) records of the late glacial transition are rare because of active thermoerosion. Hence, our results provide valuable information about the vegetation history in this region and organic matter composition.~~

The sedaDNA results imply profound changes after the LGM, which is displayed in the major structure of the terrestrial PCA biplot. First, the local taxonomic richness decreased strongly. Second, the taxonomic composition of the local flora changed towards shrub tundra and was mainly characterized by high proportions of *Betula, Salix* and *Equisetum* with a low diversity among Poaceae and herbaceousforbs plants with only a subset of the formerly present Cyperaceae. Highly fluctuating proportions, especially between Poaceae and Cyperaceae, indicate fluctuating moisture conditions but not the presence of a pond. In relatively drier periods, indicated by positive values in the ratio plot, the organic matter comprises mostly Poaceae, which were represented by a different composition in comparison to the pre-LGM, with Agrostidinae and *Poa* MOTU2. During moister periods, indicated by negative values in the ratio plot, increased proportions of wetland plants such as *Eriophorum*, *Equisetum* and *Ranunculus* were recorded. The high proportions of *Equisetum* in this time interval are supported by our spore pollen record as well as in six post-LGM samples with a median proportion of 38 % in the palaeogenetic study of Willerslev et al. (2014) and the palynological review of Andreev et al. (2011). with NPP percentages below 1 %.

The pollen analysis shows that the same dominant taxa detected by the sedaDNA approach characterize the area on a regional scale and implies shrub tundra with *Salix*, *Betula* and *Alnus* dwarf willow, dwarf birch and alder (Andreev et al., 2011). The equal relative proportions between pollen from trees and shrubs and pollen from herbsforbs, grasses and sedges indicate climate amelioration during the early Holocene (Andreev et al., 2011) with increased humidity after the marine transgression (Kienast et al., 2001). Published pollen records for 11.7–10.6 kyr BP are dominated by Cyperaceae and Poaceae. Shrub pollen increased in the Laptev Sea region approximately at 9 kyr BP (Andreev et al., 2011), while in the Khorogor Valley near Tiksi an increase especially of *Betula* pollen up to 60% was recorded already at 11.54 ± 60 kyr BP (Grosse et al., 2007; Khg-11). In this study the increase is recorded at 11.4 ± 50 kyr BP and therefore matches well with the pollen data of the Khorogor Valley. Shrub pollen increased at approximately 9 kyr BP (Andreev et al., 2011), with up to 60 % of *Betula* in the Khorogor Valley near Tiksi (Andreev et al., 2011; Grosse et al., 2007). These results match well with the pollen data presented here.

The uppermost samples are dominated by Poaceae followed by Cyperaceae pollen and show high proportions of *Alnus*, *Betula* and *Salix* ~~alder, dwarf birch and dwarf willow shrubs,~~ along with Ericales and Ranunculaceae ~~and~~ but low proportions of *Artemisia*, reflecting the modern pollen spectrum (Andreev et al., 2011; CAVM Team, 2003). In contrast, the sedaDNA surface sample is characterized by *Agrostidinae* ~~MOTUs 1 and 2~~, *Eriophorum* MOTU1 and *Carex aquatilis*. This most likely reflects their ~~high below ground~~root biomass in this sample. Taken together, both proxies reflect the tussock-sedge, dwarf-shrub~~, (moss)~~ tundra according to the division of the Circumarctic Vegetation Map (CAVM Team, 2003).

6. Conclusions

We demonstrate that combining sedaDNA with palynological analyses ~~pollen and NPPs~~ for palaeobotanical reconstructions offers valuable insights into the taxonomic composition of plant organic matter and its accumulation conditions. The proxies complement each other and differences between the records result mainly from differences in the spatial resolution, the taxonomic resolution capacities and the sampling effort.

The sedaDNA record presents a high number of taxa at a high resolution throughout the core allowing a detailed characterization of plant community changes in the deposited sediments. Furthermore, the high taxonomic resolution allowed for inferences of hydrological changes at the coring site. Our findings support the general view that sedaDNA represents the local flora and predominantly local below-ground biomass.

Pollen, spores and algae ~~and NPPs~~ prove to be important for a more complete representation of the taxonomic composition on a local to regional scale and to assess the environmental conditions. They are suitable in tracing hydrological changes at the coring site, even though the signal can be buffered by the regional and extra-regional pollen input and hence is not as pronounced as with sedaDNA.

For sediments deposited from 54 to 51 kyr BP, the pollen record suggests an open landscape shaped by relatively dry environmental conditions. Plant organic matter composition, derived from sedaDNA, reveals high taxonomic richness among grasses, sedges and ~~herbs~~forbs, including hydropyhtes with a constant presence of ~~*Salix* dwarf willows~~. The fluctuating dominance of swamp and aquatic taxa in both

proxies suggests the presence of a low-centred polygon, which frequently carried a pond in the depression, implying that severe winters prevailed, which allowed for frost cracking and polygon development. Comparisons with studies of Holocene polygons indicate that pre-LGM plant communities of low-centred polygons differed from Holocene ones.

During the post-LGM (11.4–9.7 kyr BP (13.4–11.1 cal kyr BP)), sedaDNA shows an almost complete taxonomic turnover to shrub tundra with less taxonomic richness in comparison to the pre-LGM deposits. Similar to the sedaDNA record, pollen also implies shrub tundra during the post-LGM, but in contrast to sedaDNA, the taxonomic richness decreased only to a minor extent while proportional changes among dominant taxa are more decisive.

Future developments will probably comprise technical refinements to reduce biases in sedaDNA data as the relevant sequencing technologies are progressing quickly and sequence databases are continuously enhanced with new genomic reference sequences. Biases currently introduced through the PCR step will potentially be reduced as DNA extracts will increasingly be sequenced directly - all together providing a more comprehensive understanding of past ecosystems.

Data availability

https://www.pangaea.de/ (follows after acceptance and includes the DNA and pollen, spore and algae ~~and NPP~~ datasets)

Supplement

1. S1-S3_msBK8~~upplement_Zimmermann~~.pdf
   - S1 Sequencing and sequence filtering
   - S2 Technical evaluation and authenticity of the data
   - References
   - S3 Figures
2. ~~S4_DNA_dataset.xlsx~~S4_negative_controls.xlsx
3. ~~S5_cultivated_nonArctic_plants.xlsx~~S5_DNA_dataset.xlsx

4. ~~S6_negative_controls.xlsx~~S6_exotic_DNA_sequences.xlsx
5. S7_Pollen_counts.xlsx

Author contributions

H. H. Zimmermann performed the core sub-sampling, the genetic laboratory work and all bioinformatics and statistical analyses, guided by L. S. Epp, U. Herzschuh and K. R. Stoof-Leichsenring. E. Raschke counted pollen and NPPs. P. P. Overduin coordinated the field work in 2012 and collected the core. G. Schwamborn and L. Schirrmeister opened the core and advised on sub-sampling. H. Zimmermann wrote the manuscript that all co-authors commented on.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We are grateful to the German Federal Ministry of Education and Research (BMBF) for funding this study as part of the joint German-Russian research project CARBOPERM (03G0836B, 03G0836F). We thank our colleagues who helped during fieldwork in 2012 as part of the Russian-German Cooperation SYSTEM LAPTEV SEA. L. S. Epp is supported by the German Research Council (DFG grant EP98/2-1 to L.S.E.). We gratefully acknowledge Jonas Grünwald, for assistance with the sub-sampling, Daronja Trense, for pollen sample preparation, Liv Heinecke for the age-depth model and Bennet Juhls for compiling the maps. Finally, the paper benefited by English language correction from Cathy Jenks.

References

Aiken, S. G., Dallwitz, M. J., Consaul, L. L., McJannet, C. L., Boles, R. L., Argus, G. W., Gillett, J. M., Scott, P. J., Elven, R., LeBlanc, M. C., Gillespie, L. J., Brysting, A. K., Solstad, H. and Harris, J. G.: Flora of the Canadian Arctic Archipelago: Descriptions, Illustrations, Identification, and Information

Feldfunktion geändert

Retrieval. NRC Research Press, National Research Council of Canada, Ottawa, [online] Available from: http://nature.ca/aaflora/data/index.htm (Accessed 20 July 2016), 2007.

Alsos, I. G., Coissac, E., Edwards, M., Foreid, M. K., Gielly, L., Sjoegren, P., Taberlet, P. and Yoccoz, N.: Plant DNA in sediments: to which degree do they represent the flora?, GENOME, 58(5), 186, 2015.

Andreev, A., Schirrmeister, L., Siegert, C., Bobrov, A. A., Demske, D., Seiffert, M. and Hubberten, H. W.: Paleoenvironmental changes in Northeastern Siberia during the Late Quaternary - evidence from pollen records of the Bykovsky Peninsula, Polarforschung, 70, 13–25, 2002.

Andreev, A. A., Schirrmeister, L., Tarasov, P. E., Ganopolski, A., Brovkin, V., Siegert, C., Wetterich, S. and Hubberten, H.-W.: Vegetation and climate history in the Laptev Sea region (Arctic Siberia) during Late Quaternary inferred from pollen records, Quat. Sci. Rev., 30(17–18), 2182–2199, doi:10.1016/j.quascirev.2010.12.026, 2011.

Andrews, S.: FastQC: a quality control tool for high throughput sequence data, [online] Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc, 2010.

Arnold, L. J., Roberts, R. G., MacPhee, R. D. E., Haile, J. S., Brock, F., Möller, P., Froese, D. G., Tikhonov, A. N., Chivas, A. R., Gilbert, M. T. P. and Willerslev, E.: Paper II – Dirt, dates and DNA: OSL and radiocarbon chronologies of perennially frozen sediments in Siberia, and their implications for sedimentary ancient DNA studies, Boreas, 40(3), 417–445, doi:10.1111/j.1502-3885.2010.00181.x, 2011.

Bellemain, E., Davey, M. L., Kauserud, H., Epp, L. S., Boessenkool, S., Coissac, E., Geml, J., Edwards, M., Willerslev, E., Gussarova, G., Taberlet, P. and Brochmann, C.: Fungal palaeodiversity revealed using high-throughput metabarcoding of ancient DNA from arctic permafrost, Environ. Microbiol., 15(4), 1176–1189, doi:10.1111/1462-2920.12020, 2013.

Bennett, B.: Determination of the number of zones in a biostratigraphical sequence, New Phytol., 132, 155–170, 1996.

Beug, H.-J.: Leitfaden der Pollenbestimmung, Verlag Dr. Friedrich Pfeil, München., 2004.

Binladen, J., Gilbert, M. T. P., Bollback, J. P., Panitz, F., Bendixen, C., Nielsen, R. and Willerslev, E.: The Use of Coded PCR Primers Enables High-Throughput Sequencing of Multiple Homolog Amplification Products by 454 Parallel Sequencing, PLoS ONE, 2(2), e197, doi:10.1371/journal.pone.0000197, 2007.

Birks, H. H.: Plant macrofossils, in Tracking Environmental Change Using Lake Sediments, vol. 3 Terrestrial, Algal, and Siliceous Indicators, edited by J. P. Smol, H. J. B. Birks, W. M. Last, R. S. Bradley, and K. Alverson, Springer Netherlands, Dordrecht. [online] Available from: http://link.springer.com/10.1007/0-306-47668-1 (Accessed 29 July 2016), 2001.

Blaauw, M., Christen J.A.: Flexible paleoclimate age-depth models using an autoregressive gamma process, Bayesian Anal. 6, 457–474, doi: 10.1214/ba/1339616472, 2011

Boessenkool, S., McGlynn, G., Epp, L. S., Taylor, D., Pimentel, M., Gizaw, A., Nemomissa, S., Brochmann, C. and Popp, M.: Use of ancient sedimentary DNA as a novel conservation tool for high-altitude tropical biodiversity, Conserv. Biol., 28(2), 446–455, doi:10.1111/cobi.12195, 2014.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P. and Coissac, E.: OBITools: a Unix-inspired software package for DNA metabarcoding, Mol. Ecol. Resour., 16, 176–182, doi:10.1111/1755-0998.12428, 2016.

CAVM Team: Circumpolar Arctic Vegetation Map (1:7,500,000 scale), Conservation of Arctic Flora and Fauna (CAFF) Map No. 1. U.S. Fish and Wildlife Service, Anchorage, Alaska. [online] Available from: http://www.geobotany.uaf.edu/cavm/, 2003.

Champlot, S., Berthelot, C., Pruvost, M., Bennett, E. A., Grange, T. and Geigl, E.-M.: An efficient multistrategy DNA decontamination procedure of PCR reagents for hypersensitive PCR applications, PLoS ONE, 5(9), e13042, doi:10.1371/journal.pone.0013042, 2010.

Charkin, A. N., Dudarev, O. V., Semiletov, I. P., Kruhmalev, A. V., Vonk, J. E., Sánchez-García, L., Karlsson, E. and Gustafsson, Ö.: Seasonal and interannual variability of sedimentation and organic matter distribution in the Buor-Khaya Gulf: the primary recipient of input from Lena River and coastal erosion in the southeast Laptev Sea, Biogeosciences, 8(9), 2581–2594, doi:10.5194/bg-8-2581-2011, 2011.

Cornwell, W. K., Cornelissen, J. H. C., Amatangelo, K., Dorrepaal, E., Eviner, V. T., Godoy, O., Hobbie, S. E., Hoorens, B., Kurokawa, H., Pérez-Harguindeguy, N., Quested, H. M., Santiago, L. S., Wardle, D. A., Wright, I. J., Aerts, R., Allison, S. D., Van Bodegom, P., Brovkin, V., Chatain, A., Callaghan, T. V., Díaz, S., Garnier, E., Gurvich, D. E., Kazakou, E., Klein, J. A., Read, J., Reich, P. B., Soudzilovskaia, N. A., Vaieretti, M. V. and Westoby, M.: Plant species traits are the predominant control on litter decomposition rates within biomes worldwide, Ecol. Lett., 11(10), 1065–1071, doi:10.1111/j.1461-0248.2008.01219.x, 2008.

Cox, P. A.: Hydrophilous pollination, Annu. Rev. Ecol. Syst., 19(1), 261–279, doi:10.1146/annurev.es.19.110188.001401, 1988.

Czudek, T. and Demek, J.: Thermokarst in Siberia and its influence on the development of lowland relief, Quat. Res., 1(1), 103–120, doi:10.1016/0033-5894(70)90013-X, 1970.

Davidson, E. A. and Janssens, I. A.: Temperature sensitivity of soil carbon decomposition and feedbacks to climate change, Nature, 440(7081), 165–173, doi:10.1038/nature04514, 2006.

Davidson, E. A., Trumbore, S. E. and Amundson, R.: Biogeochemistry: soil warming and organic carbon content, Nature, 408(6814), 789–790, 2000.

De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E. and Taberlet, P.: DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet, Mol. Ecol. Resour., 14(2), 306–323, doi:10.1111/1755-0998.12188, 2014.

Duchkov, A. D., Zheleznjak, M. and Sokolova, L. S.: Electronic geothermal atlases of Asian Russia, Sci. Cold Arid Reg., 6(5), 0479–0483, doi:10.3724/SP.J.1226.2014.00479, 2014.

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., Erséus, C., Gusarov, V. I., Edwards, M. E., Johnsen, A., Stenøien, H. K., Hassel, K., Kauserud, H., Yoccoz, N. G., Bråthen, K. A., Willerslev, E., Taberlet, P., Coissac, E. and Brochmann, C.: New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems, Mol. Ecol., 21(8), 1821–1833, doi:10.1111/j.1365-294X.2012.05537.x, 2012.

Epp, L. S., Gussarova, G., Boessenkool, S., Olsen, J., Haile, J., Schrøder-Nielsen, A., Ludikova, A., Hassel, K., Stenøien, H. K., Funder, S., Willerslev, E., Kjær, K. and Brochmann, C.: Lake sediment multi-taxon DNA from North Greenland records early post-glacial appearance of vascular plants and accurately tracks environmental changes, Quat. Sci. Rev., 117(0318), 152–163, doi:10.1016/j.quascirev.2015.03.027, 2015.

Grimm, E. C.: CONISS : A FORTRAN 77 Program for stratigraphically constrained cluster analysis by the method of incremental sum of squares, Comput. Geosci., 13(I), 13–35, 1987.

Grosse, G., Schirrmeister, L., Siegert, C., Kunitsky, V. V., Slagoda, E. A., Andreev, A. A. and Dereviagyn, A. Y.: Geological and geomorphological evolution of a sedimentary periglacial landscape in Northeast Siberia during the Late Quaternary, Geomorphology, 86(1–2), 25–51, doi:10.1016/j.geomorph.2006.08.005, 2007.

Grosse, G., Romanovsky, V., Jorgenson, T., Walter Anthony, K., Brown, J. and Overduin, P. P.: Vulnerability and feedbacks of permafrost to climate change, EOS Trans. Am. Geophys. Union, 92(9), 73–80, doi:10.1002/ppp.689.Rowland, 2011.

Günther, F., Overduin, P. P., Makarov, A. S. and Grigoriev, M. N.: Russian-German cooperation SYSTEM LAPTEV SEA : the expeditions Laptev Sea - Mamontov Klyk 2011 & Buor Khaya 2012, Alfred Wegener Institute for Polar and Marine Research, Bremerhaven. [online] Available from: hdl:10013/epic.41834, 2013a.

Günther, F., Overduin, P. P., Sandakov, A. V., Grosse, G. and Grigoriev, M. N.: Short- and long-term thermo-erosion of ice-rich permafrost coasts in the Laptev Sea region, Biogeosciences, 10, 4297–4318, doi:10.5194/bg-10-4297-2013, 2013b.

Gustafsson, Ö., van Dongen, B. E., Vonk, J. E., Dudarev, O. V. and Semiletov, I. P.: Widespread release of old carbon across the Siberian Arctic echoed by its large rivers, Biogeosciences, 8(6), 1737–1743, doi:10.5194/bg-8-1737-2011, 2011.

Haile, J., Holdaway, R., Oliver, K., Bunce, M., Gilbert, M. T. P., Nielsen, R., Munch, K., Ho, S. Y. W., Shapiro, B. and Willerslev, E.: Ancient DNA chronology within sediment deposits: Are Paleobiological reconstructions possible and is DNA leaching a factor?, Mol. Biol. Evol., 24(4), 982–989, doi:10.1093/molbev/msm016, 2007.

Haile, J., Froese, D. G., MacPhee, R. D. E., Roberts, R. G., Arnold, L. J., Reyes, A. V., Rasmussen, M., Nielsen, R., Brook, B. W., Robinson, S., Demuro, M., Gilbert, M. T. P., Munch, K., Austin, J. J., Cooper, A., Barnes, I., Möller, P. and Willerslev, E.: Ancient DNA reveals late survival of mammoth and horse in interior Alaska, Proc. Natl. Acad. Sci., pnas.0912510106, doi:10.1073/pnas.0912510106, 2009.

Hansen, A. J., Mitchell, D. L., Wiuf, C., Paniker, L., Brand, T. B., Binladen, J., Gilichinsky, D. A., Rønn, R. and Willerslev, E.: Crosslinks rather than strand breaks determine access to ancient DNA sequences from frozen sediments, Genetics, 173(2), 1175–1179, doi:10.1534/genetics.106.057349, 2006.

Heck, K. L., van Belle, G. and Simberloff, D.: Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size, Ecology, 56(6), 1459–1461, doi:10.2307/1934716, 1975.

Hobbie, S. E.: Effects of plant species on nutrient cycling, Trends Ecol. Evol., 7(10), 336–339, doi:10.1016/0169-5347(92)90126-V, 1992.

Hubberten, H. W., Andreev, A., Astakhov, V. I., Demidov, I., Dowdeswell, J. A., Henriksen, M., Hjort, C., Houmark-Nielsen, M., Jakobsson, M., Kuzmina, S., Larsen, E., Lunkka, J. P., Lysså, A., Mangerud, J., Möller, P., Saarnisto, M., Schirrmeister, L., Sher, A. V., Siegert, C., Siegert, M. J. and Svendsen, J. I.: The periglacial climate and environment in northern Eurasia during the last glaciation, Quat. Sci. Rev., 23(11–13), 1333–1357, doi:10.1016/j.quascirev.2003.12.012, 2004.

Hugelius, G., Strauss, J., Zubrzycki, S., Harden, J. W., Schuur, E. A. G., Ping, C.-L., Schirrmeister, L., Grosse, G., Michaelson, G. J., Koven, C. D., O'Donnell, J. A., Elberling, B., Mishra, U., Camill, P., Yu, Z., Palmtag, J. and Kuhry, P.: Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps, Biogeosciences, 11(23), 6573–6593, doi:10.5194/bg-11-6573-2014, 2014.

Hurlbert, S. H.: The nonconcept of species diversity: A critique and alternative parameters, Ecology, 52(4), 577–586, doi:10.2307/1934145, 1971.

Iversen, C. M., Sloan, V. L., Sullivan, P. F., Euskirchen, E. S., McGuire, A. D., Norby, R. J., Walker, A. P., Warren, J. M. and Wullschleger, S. D.: The unseen iceberg: plant roots in arctic tundra, New Phytol., 205(1), 34–58, doi:10.1111/nph.13003, 2015.

John, D. M., Whitton, B. A. and Brook, A. J., Eds.: The freshwater algal flora of the British Isles: An identification guide to freshwater and terrestrial algae, Cambridge University Press, Cambridge., 2002.

Jørgensen, T., Haile, J., Möller, P., Andreev, A., Boessenkool, S., Rasmussen, M., Kienast, F., Coissac, E., Taberlet, P., Brochmann, C., Bigelow, N. H., Andersen, K., Orlando, L., Gilbert, M. T. P. and Willerslev, E.: A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability, Mol. Ecol., 21(8), 1989–2003, doi:10.1111/j.1365-294X.2011.05287.x, 2012.

Juggins, S.: rioja: Analysis of Quaternary science data, R package version 0.7-3., 2012.

Kanz, C., Aldebert, P., Althorpe, N., Baker, W., Baldwin, A., Bates, K., Browne, P., van den Broek, A., Castro, M., Cochrane, G., Duggan, K., Eberhardt, R., Faruque, N., Gamble, J., Garcia Diez, F., Harte, N., Kulikova, T., Lin, Q., Lombard, V., Lopez, R., Mancuso, R., McHale, M., Nardone, F., Silventoinen, V., Sobhany, S., Stoehr, P., Tuli, M. A., Tzouvara, K., Vaughan, R., Wu, D., Zhu, W. and Apweiler, R.: The EMBL nucleotide sequence database, Nucleic Acids Res., 33(DATABASE ISS.), 29–33, doi:10.1093/nar/gki098, 2005.

Kienast, F., Siegert, C., Dereviagin, A. and Mai, D. H.: Climatic implications of Late Quaternary plant macrofossil assemblages from the Taymyr Peninsula , Siberia, Glob. Planet. Change, 31, 265–281, 2001.

Kienast, F., Schirrmeister, L., Siegert, C. and Tarasov, P.: Palaeobotanical evidence for warm summers in the East Siberian Arctic during the last cold stage, Quat. Res., 63(3), 283–300, doi:10.1016/j.yqres.2005.01.003, 2005.

Kienast, F., Tarasov, P., Schirrmeister, L., Grosse, G. and Andreev, A. A.: Continental climate in the East Siberian Arctic during the last interglacial: Implications from palaeobotanical records, Glob. Planet. Change, 60(3–4), 535–562, doi:10.1016/j.gloplacha.2007.07.004, 2008.

de Klerk, P., Donner, N., Joosten, H., Karpov, N. S., Minke, M., Seifert, N. and Theuerkauf, M.: Vegetation patterns, recent pollen deposition and distribution of non-pollen palynomorphs in a polygon mire near Chokurdakh (NE Yakutia, NE Siberia), Boreas, 38(1), 39–58, doi:10.1111/j.1502-3885.2008.00036.x, 2009.

de Klerk, P., Donner, N., Karpov, N. S., Minke, M. and Joosten, H.: Short-term dynamics of a low-centred ice-wedge polygon near Chokurdakh (NE Yakutia, NE Siberia) and climate change during the last ca 1250 years, Quat. Sci. Rev., 30(21–22), 3013–3031, doi:10.1016/j.quascirev.2011.06.016, 2011.

de Klerk, P., Teltewskoi, A., Theuerkauf, M. and Joosten, H.: Vegetation patterns, pollen deposition and distribution of non-pollen palynomorphs in an ice-wedge polygon near Kytalyk (NE Siberia), with some remarks on Arctic pollen morphology, Polar Biol., 37(10), 1393–1412, doi:10.1007/s00300-014-1529-3, 2014.

van der Knaap, W. O.: Transported pollen and spores on Spitsbergen and Jan Mayen, Pollen Spores, 24, 449–453, 1987.

Knoblauch, C., Beer, C., Sosnin, A., Wagner, D. and Pfeiffer, E.-M.: Predicting long-term carbon mineralization and trace gas production from thawing permafrost of Northeast Siberia, Glob. Change Biol., 19(4), 1160–1172, doi:10.1111/gcb.12116, 2013.

Lachenbruch, A. H.: Mechanics of thermal contraction cracks and ice-wedge polygons in permafrost, Geol. Soc. Am. Spec. Pap., 70, 1–66, doi:10.1130/SPE70-p1, 1962.

Leffingwell, E. de K.: Ground-ice wedges: The dominant form of ground-ice on the north coast of Alaska, J. Geol., 23(7), 635–654, 1915.

Levy-Booth, D. J., Campbell, R. G., Gulden, R. H., Hart, M. M., Powell, J. R., Klironomos, J. N., Pauls, K. P., Swanton, C. J., Trevors, J. T. and Dunfield, K. E.: Cycling of extracellular DNA in the soil environment, Soil Biol. Biochem., 39(12), 2977–2991, doi:10.1016/j.soilbio.2007.06.020, 2007.

Lydolph, M. C., Jacobsen, J., Arctander, P., Gilbert, M. T. P., Gilichinsky, D. A., Hansen, A. J., Willerslev, E. and Lange, L.: Beringian paleoecology inferred from permafrost-preserved fungal DNA, Appl. Environ. Microbiol., 71(2), 1012–1017, doi:10.1128/AEM.71.2.1012-1017.2005, 2005.

Lyell, C.: Principles of geology: Being an inquiry how far the former changes of the earth's surface are referable to causes now in operation, Murray, London., 1830.

MacArthur, H.: On the relative abundance of bird species, Proc. Natl. Acad. Sci. USA, 43, 293–295, 1957.

Mensing, S., Smith, J., Burkle Norman, K. and Allan, M.: Extended drought in the Great Basin of western North America in the last two millennia reconstructed from pollen records, Quat. Int., 188(1), 79–89, doi:10.1016/j.quaint.2007.06.009, 2008.

Minke, M., Donner, N., Karpov, N. S., de Klerk, P. and Joosten, H.: Distribution, diversity, development and dynamics of polygon mires: Examples from NE Yakutia (NE Siberia), Peatl. Int., 1, 36–40, 2007.

Minke, M., Donner, N., Karpov, N., de Klerk, P. and Joosten, H.: Patterns in vegetation composition, surface height and thaw depth in polygon mires in the Yakutian Arctic (NE Siberia): a

microtopographical characterisation of the active layer, Permafr. Periglac. Process., 20(4), 357–368, doi:10.1002/ppp.663, 2009.

Mogensen, H. L.: The Hows and Whys of Cytoplasmic Inheritance in Seed Plants, Am. J. Bot., 83(3), 383, doi:10.2307/2446172, 1996.

Moore, P. D., Webb, J. A. and Collison, M. E.: Pollen Analysis, 2nd ed., Blackwell Scientific Publications, Oxford, UK., 1991.

Niemeyer, B., Klemm, J., Pestryakova, L. A. and Herzschuh, U.: Relative pollen productivity estimates for common taxa of the northern Siberian Arctic, Rev. Palaeobot. Palynol., 221, 71–82, doi:10.1016/j.revpalbo.2015.06.008, 2015.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O´Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H. and Wagner, H.: vegan: Community Ecology Package. R package version 2.0-2., 2011.

Parducci, L., Matetovici, I., Fontana, S. L., Bennett, K. D., Suyama, Y., Haile, J., Kjaer, K. H., Larsen, N. K., Drouzas, A. D. and Willerslev, E.: Molecular- and pollen-based vegetation analysis in lake sediments from central Scandinavia, Mol. Ecol., 22(13), 3511–3524, doi:10.1111/mec.12298, 2013.

Pedersen, M. W., Ginolhac, A., Orlando, L., Olsen, J., Andersen, K., Holm, J., Funder, S., Willerslev, E. and Kjær, K. H.: A comparative study of ancient environmental DNA to pollen and macrofossils from lake sediments reveals taxonomic overlap and additional plant taxa, Quat. Sci. Rev., 75, 161–168, doi:10.1016/j.quascirev.2013.06.006, 2013.

Pedersen, M. W., Ruter, A., Schweger, C., Friebe, H., Staff, R. A., Kjeldsen, K. K., Mendoza, M. L. Z., Beaudoin, A. B., Zutter, C., Larsen, N. K., Potter, B. A., Nielsen, R., Rainville, R. A., Orlando, L., Meltzer, D. J., Kjær, K. H. and Willerslev, E.: Postglacial viability and colonization in North America's ice-free corridor, Nature, 537(7618), 45–49, 2016.

Preston, C. D. and Croft, J. M.: Aquatic plants in Britain and Ireland: a joint project of the Environment Agency Institute of Terrestrial Ecology and the Joint Nature Conservation Committee, Harley, Colchester., 1997.

R Core Team: R: A language and environment for statistical computing, Vienna, Austria. [online] Available from: http://www.r-project.org/, 2014.

Rawlence, N. J., Lowe, D. J., Wood, J. R., Young, J. M., Churchman, G. J., Huang, Y.-T. and Cooper, A.: Using palaeoenvironmental DNA to reconstruct past environments: progress and prospects, J. Quat. Sci., 29(7), 610–626, doi:10.1002/jqs.2740, 2014.

Romanovsky, V. E., Drozdov, D. S., Oberman, N. G., Malkova, G. V., Kholodov, A. L., Marchenko, S. S., Moskalenko, N. G., Sergeev, D. O., Ukraintseva, N. G., Abramov, A. A., Gilichinsky, D. A. and Vasiliev, A. A.: Thermal state of permafrost in Russia, Permafr. Periglac. Process., 21(March), 136–155, doi:10.1002/ppp.683, 2010.

Savelieva, L. A., Raschke, E. A. and Titova, D. V.: Photographic Atlas of Plants and Pollen of the Lena River Delta, St. Petersburg State University, Saint-Petersburg., 2013.

Sayers, E. W., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., DiCuccio, M., Edgar, R., Federhen, S., Feolo, M., Geer, L. Y., Helmberg, W., Kapustin, Y., Landsman, D., Lipman, D. J., Madden, T. L., Maglott, D. R., Miller, V., Mizrachi, I., Ostell, J., Pruitt, K. D., Schuler, G. D., Sequeira, E., Sherry, S. T., Shumway, M., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusova, T. A., Wagner, L., Yaschenko, E. and Ye, J.: Database resources of the National Center for Biotechnology Information, Nucleic Acids Res., 37(Database issue), D5-15, doi:10.1093/nar/gkn741, 2009.

Schirrmeister, L., Grosse, G., Wetterich, S., Overduin, P. P., Strauss, J., Schuur, E. A. G. and Hubberten, H.-W.: Fossil organic matter characteristics in permafrost deposits of the northeast Siberian Arctic, J. Geophys. Res., 116, doi:10.1029/2011JG001647, 2011a.

Schirrmeister, L., Kunitsky, V., Grosse, G., Wetterich, S., Meyer, H., Schwamborn, G., Babiy, O., Derevyagin, A. and Siegert, C.: Sedimentary characteristics and origin of the Late Pleistocene Ice Complex on north-east Siberian Arctic coastal lowlands and islands – A review, Quat. Int., 241(1–2), 3–25, doi:10.1016/j.quaint.2010.04.004, 2011b.

Schirrmeister, L., Schwamborn, G., Overduin, P. P., Strauss, J., Fuchs, M. C., Grigoriev, M., Yakshina, I., Rethemeyer, J., Dietze, E. and Wetterich, S.: Yedoma Ice Complex of the Buor Khaya Peninsula (southern Laptev Sea), Biogeosciences Discuss., 1–36, doi:10.5194/bg-2016-283, 2016.

Schnell, I. B., Bohmann, K. and Gilbert, M. T. P.: Tag jumps illuminated – reducing sequence-to-sample misidentifications in metabarcoding studies, Mol. Ecol. Resour., 15(6), 1289–1303, doi:10.1111/1755-0998.12402, 2015.

Sher, A. V., Kuzmina, S. A., Kuznetsova, T. V. and Sulerzhitsky, L. D.: New insights into the Weichselian environment and climate of the East Siberian Arctic, derived from fossil insects, plants, and mammals, Quat. Sci. Rev., 24(5–6), 533–569, doi:10.1016/j.quascirev.2004.09.007, 2005.

Simpson, G. L.: Analogue methods in palaeoecology: Using the analogue package, J. Stat. Softw., 22(2), 1–29, doi:10.18637/jss.v022.i02, 2007.

Simpson, G. L. and Oksanen, J.: analogue: Analogue and weighted averaging methods for palaeoecology. [online] Available from: http://cran.r-project.org/package=analogue, 2016.

Sjögren, P., van der Knaap, W. O., Huusko, A. and van Leeuwen, J. F. N.: Pollen productivity, dispersal, and correction factors for major tree taxa in the Swiss Alps based on pollen-trap results, Rev. Palaeobot. Palynol., 152(3–4), 200–210, doi:10.1016/j.revpalbo.2008.05.003, 2008.

Sjögren, P., Edwards, M. E., Gielly, L., Langdon, C. T., Croudace, I. W., Merkel, M. K. F., Fonville, T. and Alsos, I. G.: Lake sedimentary DNA accurately records 20th Century introductions of exotic conifers in Scotland, New Phytol., n/a-n/a, doi:10.1111/nph.14199, 2016.

Soininen, E. M., Gauthier, G., Bilodeau, F., Berteaux, D., Gielly, L., Taberlet, P., GUssarova, G., Bellemain, E., Hassel, K., Stenøien, H. K., Epp, L., Schrøder-Nielsen, A., Brochmann, C. and Yoccoz, N. G.: Highly overlapping winter diet in two sympatric lemming species revealed by DNA metabarcoding, Plos One, 10(1), e0115335, doi:10.1371/journal.pone.0115335, 2015.

Sønstebø, J. H., Gielly, L., Brysting, a K., Elven, R., Edwards, M., Haile, J., Willerslev, E., Coissac, E., Rioux, D., Sannier, J., Taberlet, P. and Brochmann, C.: Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate., Mol. Ecol. Resour., 10(6), 1009–1018, doi:10.1111/j.1755-0998.2010.02855.x, 2010.

Stapel, J. G., Schirrmeister, L., Overduin, P. P., Wetterich, S., Strauss, J., Horsfield, B. and Mangelsdorf, K.: Microbial lipid signatures and substrate potential of organic matter in permafrost deposits: Implications for future greenhouse gas production, J. Geophys. Res. Biogeosciences, 121(10), 2016JG003483, doi:10.1002/2016JG003483, 2016.

Stockmarr, J.: Tablets with spores used in absolute pollen analysis, Pollen et Spores, 13:615-621, 1971

Strauss, J., Schirrmeister, L., Wetterich, S., Borchers, A. and Davydov, S. P.: Grain-size properties and organic-carbon stock of Yedoma Ice Complex permafrost from the Kolyma lowland, northeastern Siberia, Glob. Biogeochem. Cycles, 26, 1–12, doi:10.1029/2011GB004104, 2012.

Strauss, J., Schirrmeister, L., Grosse, G., Wetterich, S., Ulrich, M., Herzschuh, U. and Hubberten, H.-W.: The deep permafrost carbon pool of the Yedoma region in Siberia and Alaska, Geophys. Res. Lett., 40, 6165–6170, doi:10.1002/2013GL058088, 2013.

Strauss, J., Schirrmeister, L., Mangelsdorf, K., Eichhorn, L., Wetterich, S. and Herzschuh, U.: Organic-matter quality of deep permafrost carbon – a study from Arctic Siberia, Biogeosciences, 12(7), 2227–2245, doi:10.5194/bg-12-2227-2015, 2015.

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C. and Willerslev, E.: Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding, Nucleic Acids Res., 35(3), doi:10.1093/nar/gkl938, 2007.

Teltewskoi, A., Beermann, F., Beil, I., Bobrov, A., de Klerk, P., Lorenz, S., Lüder, A., Michaelis, D. and Joosten, H.: 4000 years of changing wetness in a permafrost polygon peatland (Kytalyk, NE

Siberia): A comparative high-resolution multi-proxy study, Permafr. Periglac. Process., 27(1), 76–95, doi:10.1002/ppp.1869, 2016.

Thomsen, P. F. and Willerslev, E.: Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity, Biol. Conserv., 183, 4–18, doi:10.1016/j.biocon.2014.11.019, 2015.

Wagner, D., Gattinger, A., Embacher, A., Pfeiffer, E.-M., Schloter, M. and Lipski, A.: Methanogenic activity and biomass in Holocene permafrost deposits of the Lena Delta, Siberian Arctic and its implication for the global methane budget, Glob. Change Biol., 13(5), 1089–1099, doi:10.1111/j.1365-2486.2007.01331.x, 2007.

Willerslev, E., Hansen, A. J., Binladen, J., Brand, T. B., Gilbert, M. T. P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D. A. and Cooper, A.: Diverse Plant and Animal Genetic Records from Holocene and Pleistocene Sediments, Science, 300(5620), 791–795, doi:10.1126/science.1084114, 2003.

Willerslev, E., Hansen, A. J. and Poinar, H. N.: Isolation of nucleic acids and cultures from fossil ice and permafrost, Trends Ecol. Evol., 19(3), 141–147, doi:10.1016/j.tree.2003.11.010, 2004a.

Willerslev, E., Hansen, A. J., Rønn, R., Brand, T. B., Barnes, I., Wiuf, C., Gilichinsky, D., Mitchell, D. and Cooper, A.: Long-term persistence of bacterial DNA, Curr. Biol., 14(1), R9–R10, doi:10.1016/j.cub.2003.12.012, 2004b.

Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M. E., Lorenzen, E. D., Vestergard, M., Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L. S., Pearman, P. B., Cheddadi, R., Murray, D., Brathen, K. A., Yoccoz, N., Binney, H., Cruaud, C., Wincker, P., Goslar, T., Alsos, I. G., Bellemain, E., Brysting, A. K., Elven, R., Sonstebo, J. H., Murton, J., Sher, A., Rasmussen, M., Ronn, R., Mourier, T., Cooper, A., Austin, J., Moller, P., Froese, D., Zazula, G., Pompanon, F., Rioux, D., Niderkorn, V., Tikhonov, A., Savvinov, G., Roberts, R. G., MacPhee, R. D. E., Gilbert, M. T. P., Kjaer, K. H., Orlando, L., Brochmann, C. and Taberlet, P.: Fifty thousand years of Arctic vegetation and megafaunal diet, Nature, 506(7486), 47–51, doi:10.1038/nature12921, 2014.

Xie, C.-F. and Lou, H.-X.: Secondary metabolites in bryophytes: An ecological aspect, Chem. Biodivers., 6(3), 303–312, doi:10.1002/cbdv.200700450, 2009.

Yoccoz, N. G., Bråthen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., Von Stedingk, H., Brysting, A. K., Coissac, E., Pompanon, F., Sønstebø, J. H., Miquel, C., Valentini, A., De Bello, F., Chave, J., Thuiller, W., Wincker, P., Cruaud, C., Gavory, F., Rasmussen, M., Gilbert, M. T. P., Orlando, L., Brochmann, C., Willerslev, E. and Taberlet, P.: DNA from soil mirrors plant taxonomic and growth form diversity, Mol. Ecol., 21(15), 3647–3655, doi:10.1111/j.1365-294X.2012.05545.x, 2012.

Zibulski, R., Herzschuh, U., Pestryakova, L. A., Wolter, J., Müller, S., Schilling, N., Wetterich, S., Schirrmeister, L. and Tian, F.: River flooding as a driver of polygon dynamics: modern vegetation data and a millennial peat record from the Anabar River lowlands (Arctic Siberia), Biogeosciences, 10(8), 5703–5728, doi:10.5194/bg-10-5703-2013, 2013.

Zibulski, R., Herzschuh, U. and Pestryakova, L. A.: Vegetation patterns along micro-relief and vegetation type transects in polygonal landscapes of the Siberian Arctic, J. Veg. Sci., 27(2), 377–386, doi:10.1111/jvs.12356, 2016.

Figures

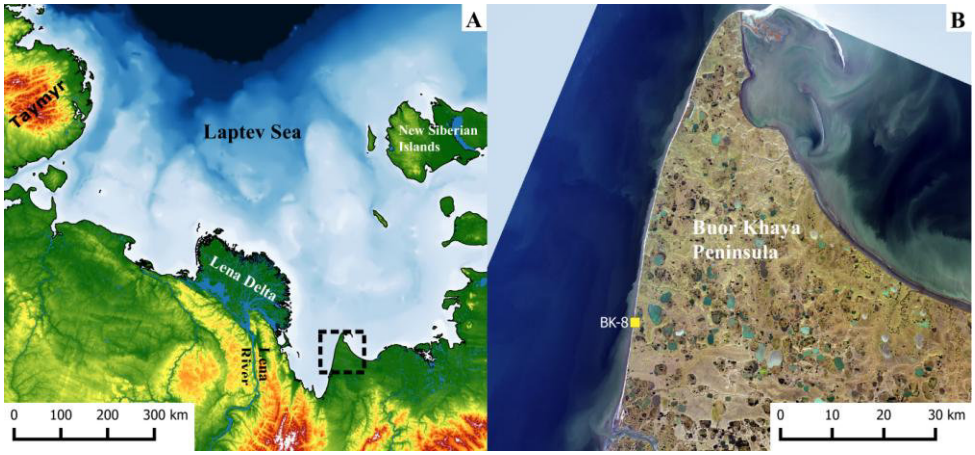

Figure 1: (A) Geographical setting of the study site in the southern Laptev Sea with the Buor Khaya Peninsula framed by dashed lines and (B) the position of the BK-8 drilling site on the western coast of the peninsula (compiled by Bennet Juhls).

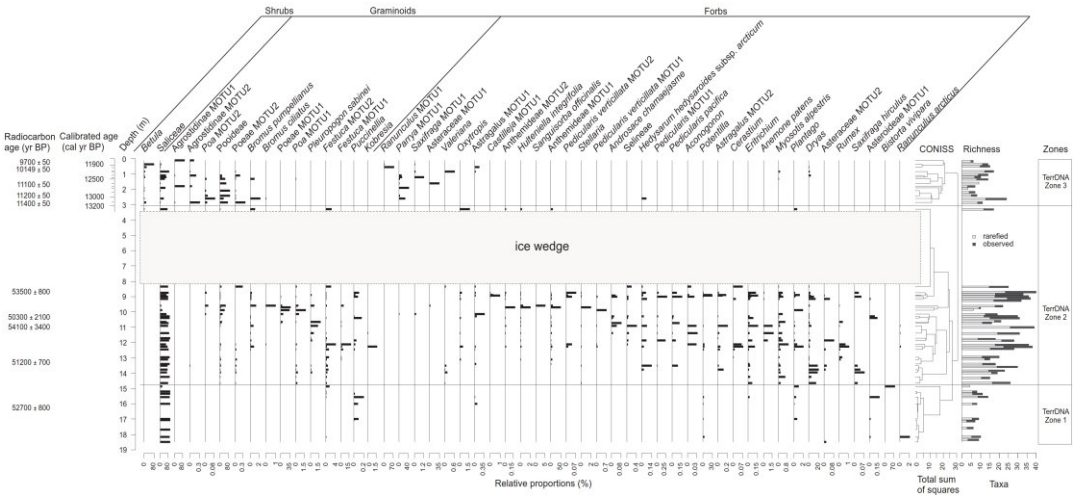

Figure 2: Stratigram of terrestrial sedaDNA (subset for statistical analyses) with relative proportions in percent of the taxa in each sample as horizontal bars, CONISS dendrogram, observed (dark grey) and rarefied (light grey) taxonomic richness. Uncalibrated radiocarbon ages are given with standard error in years before present (yr BP) and calibrated ages from the age-depth model in calendar years before present (cal BP). The grey area indicates the ice-wedge where no samples were taken. Horizontal lines indicate borders between zones. Scaling is taxon specific for better visibility of low percentages. ~~Uncalibrated radiocarbon ages are given in years before present (yr BP).~~

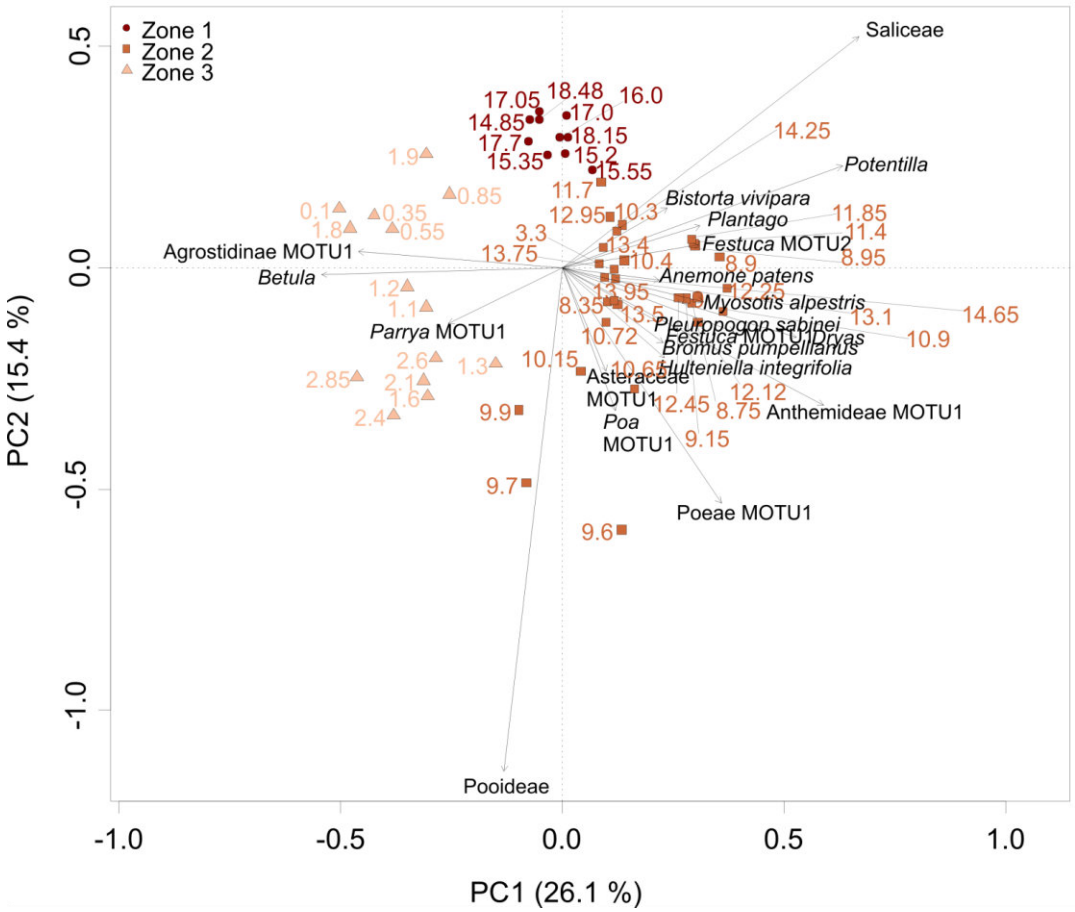

Figure 3: PCA biplot of terrestrial taxa from plant sedaDNA, showing a subset of 20 taxa which explained most of the variance in the dataset. Sample depths (m) are coloured according to their corresponding zone. Explained variances of the two principal components (PC) are shown in brackets.

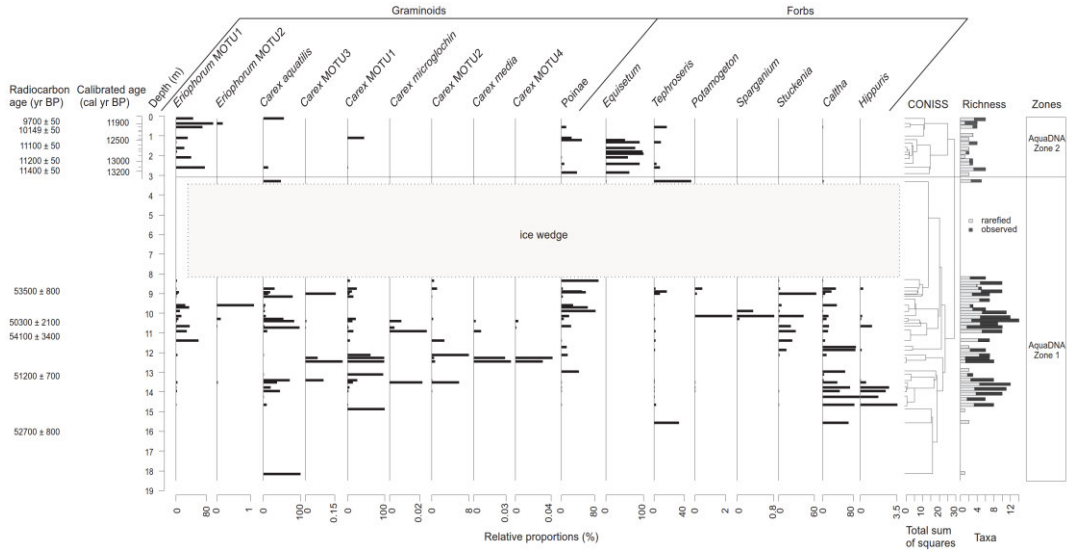

Figure 4: Stratigram of swamp and aquatic sedaDNA (subset for statistical analyses) with relative proportions in percent of the taxa in each sample as horizontal bars, CONISS dendrogram, observed (dark grey) and rarefied (light grey) taxonomic richness. Uncalibrated radiocarbon ages are given with standard error in years before present (yr BP) and calibrated ages from the age-depth model in calendar years before present (cal BP). The grey area indicates the ice-wedge where no samples were taken. The horizontal line indicates the border between the zones. Scaling is taxon specific for better visibility of low percentages. ~~Uncalibrated radiocarbon ages are given in years before present (yr BP).~~

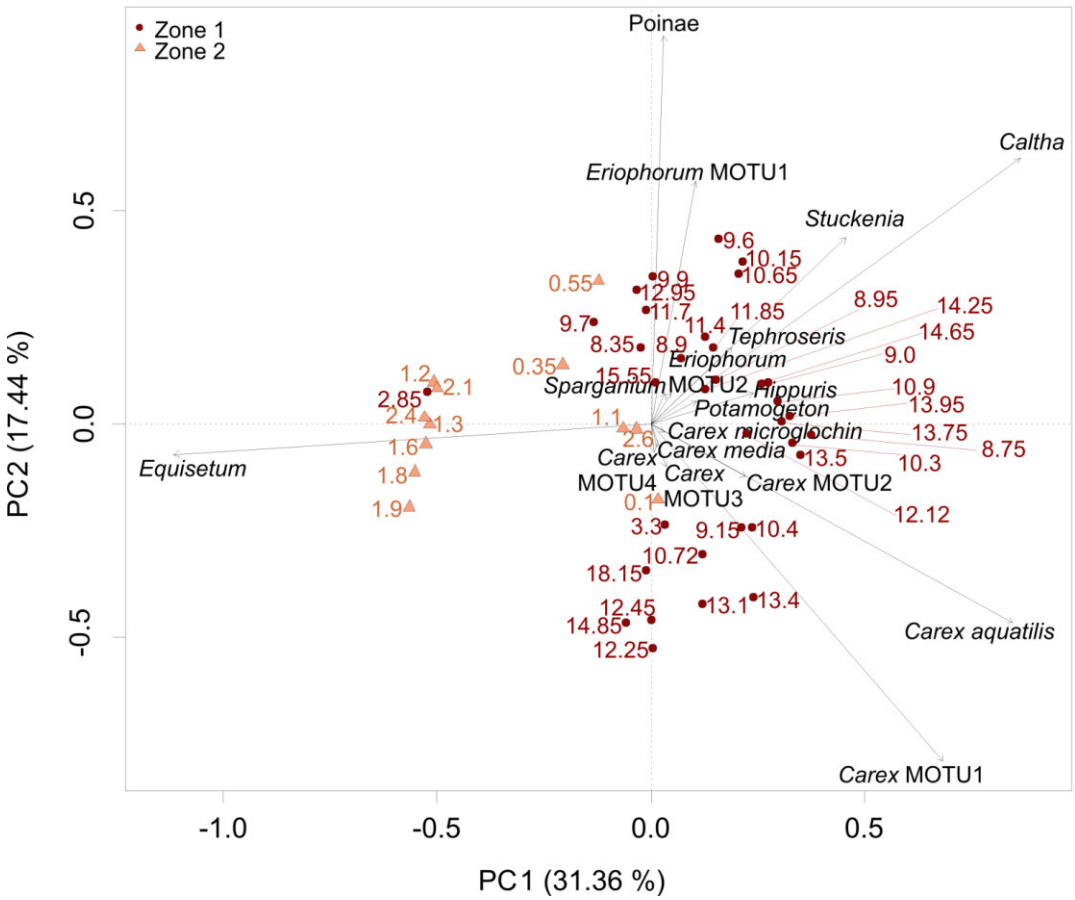

Figure 5: PCA biplot of swamp and aquatic taxa from plant sedaDNA. Sample depths (m) are coloured according to their corresponding zone. Explained variances of the two principal components (PC) are shown in brackets.

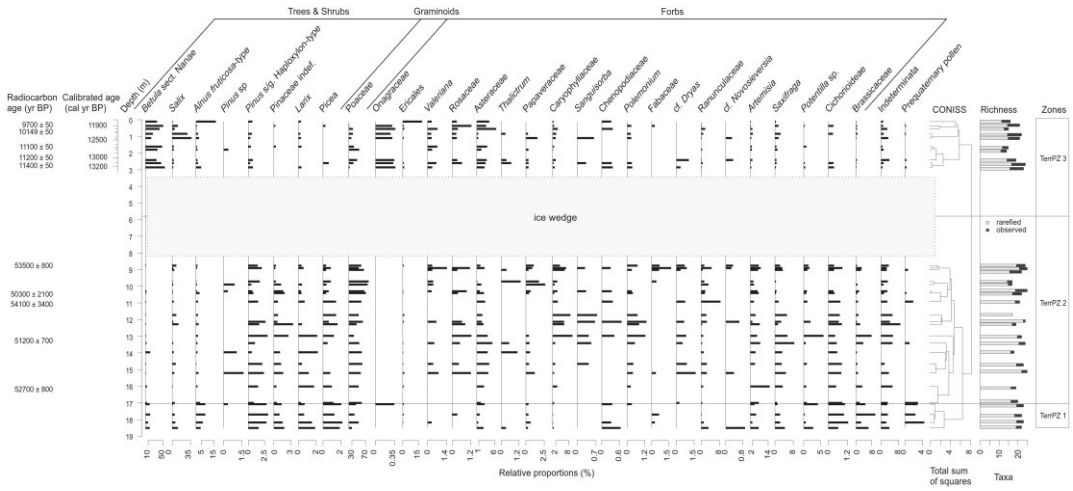

Figure 6: Stratigram of terrestrial pollen (subset for statistical analyses) with relative proportions in percentage of the taxa in each sample as horizontal bars, CONISS dendrogram, observed (dark grey) and rarefied (light grey) taxonomic richness. Uncalibrated radiocarbon ages are given with standard error in years before present (yr BP) and calibrated ages from the age-depth model in calendar years before present (cal BP). The grey area indicates the ice-wedge where no samples were taken. Horizontal lines indicate borders between palynological zones. Scaling is taxon specific for better visibility of low percentages. ~~Uncalibrated radiocarbon ages are given in years before present (yr BP).~~

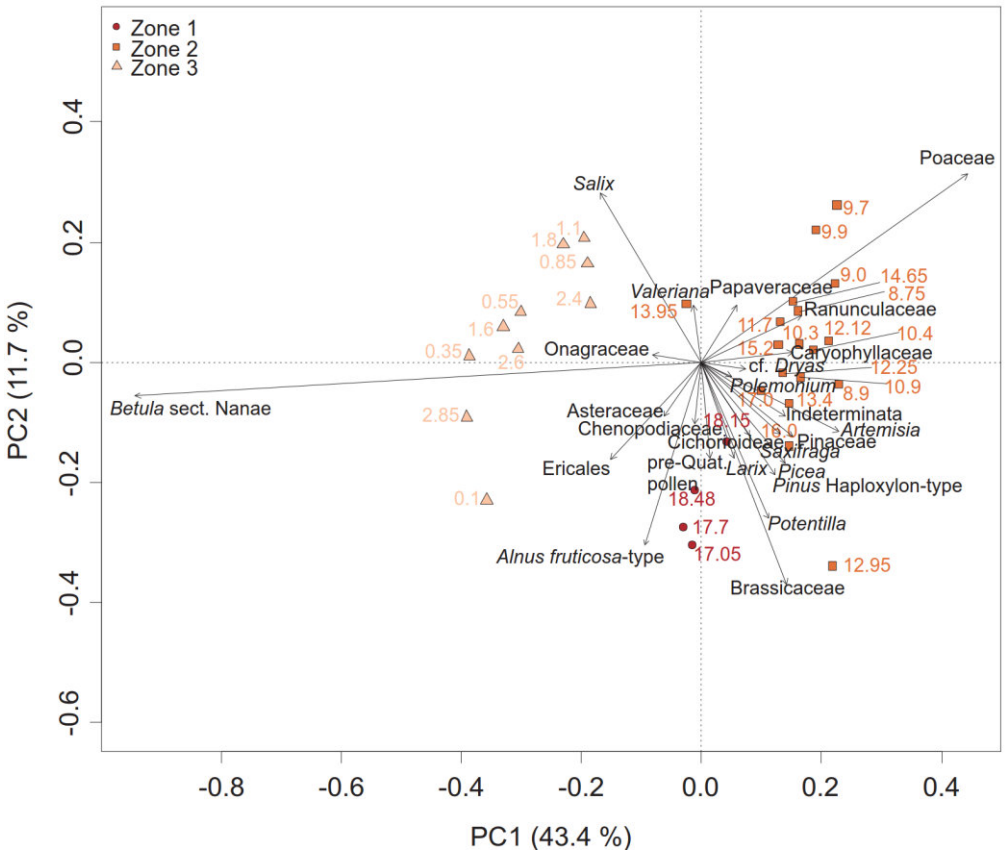

Figure 7: PCA biplot of terrestrial taxa from pollen analysis, showing a subset of 2~~5~~3 taxa, indeterminate and pre-Quaternary pollen, which explained most of the variance in the dataset. Sample depths (m) are coloured according to their corresponding zone. Explained variances of the two principal components (PC) are shown in brackets.

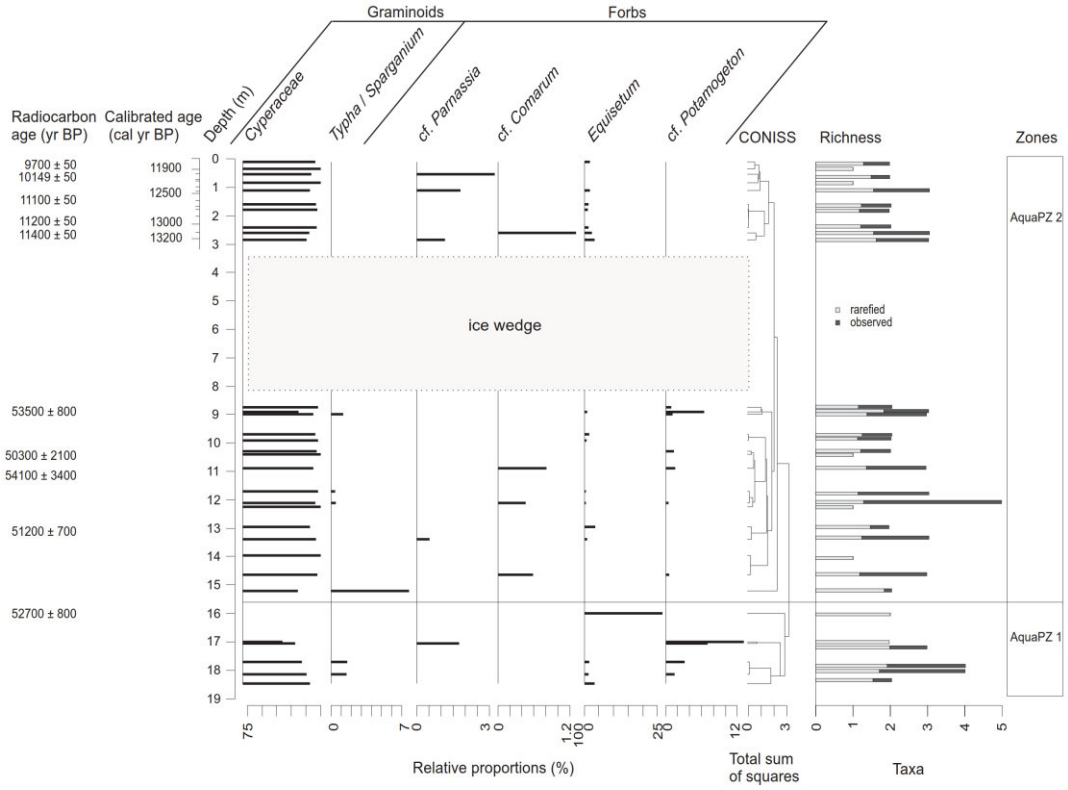

Figure 8: Stratigram of swamp and aquatic pollen/spores with relative proportions of the taxa in each sample as horizontal bars and CONISS dendrogram. The grey area indicates the ice-wedge where no samples were taken. Uncalibrated radiocarbon ages are given with standard error in years before present (yr BP) and calibrated ages from the age-depth model in calendar years before present (cal BP). The horizontal line indicates the border between the palynological zones. Scaling is taxon specific for better visibility of low percentages. ~~Uncalibrated radiocarbon ages are given in years before present (yr BP).~~

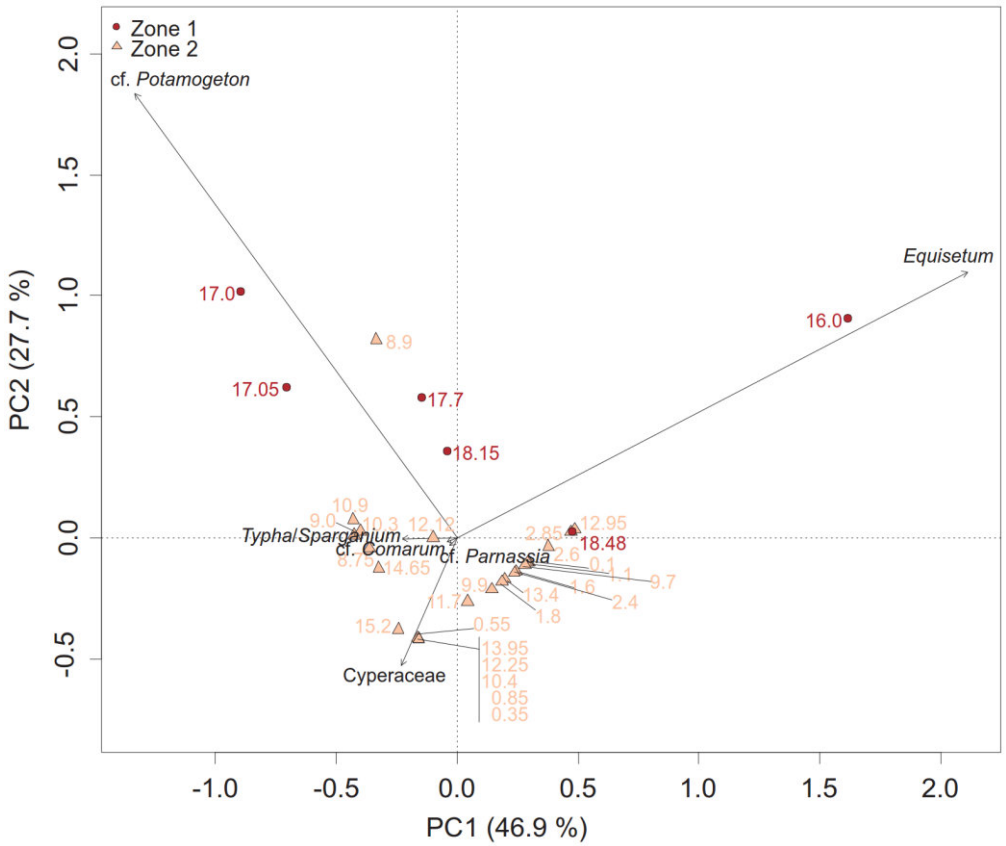

Figure 9: PCA biplot of swamp and aquatic taxa from pollen analysis. Sample depths (m) are coloured according to their corresponding zone. Explained variances of the two principal components (PC) are shown in brackets.

5

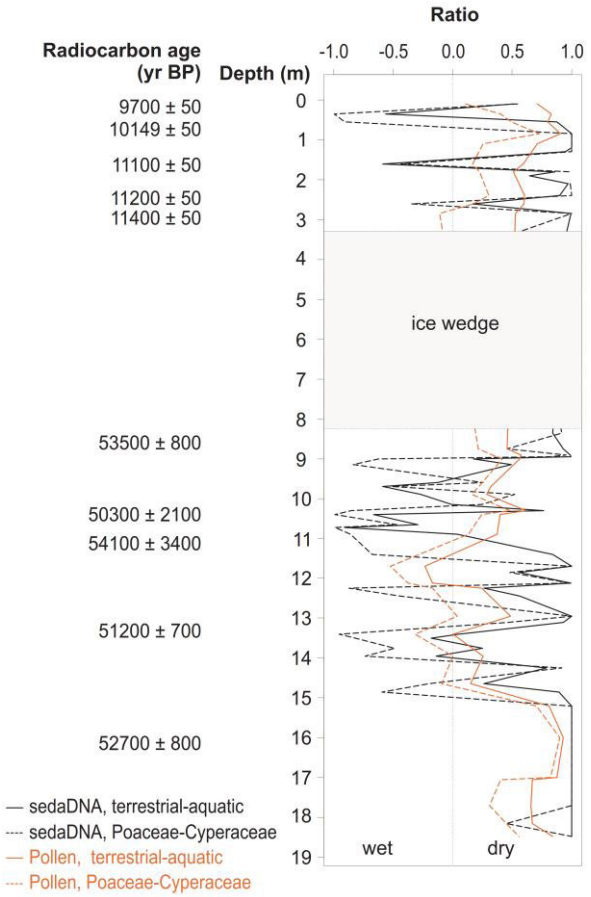

Figure 10: Terrestrial-aquatic (Eq. 1) (line) and Poaceae-Cyperaceae ratios (Eq. 2) (dashed line) of plant sedaDNA (black) and pollen (red). Positive ratios indicate a higher contribution of terrestrial or Poaceae sequence counts or pollen grains and hence drier conditions. Negative ratios indicate a higher contribution of swamp and aquatic or Cyperaceae sequence counts or pollen grains and hence wetter conditions. Uncalibrated radiocarbon ages are given in years before present (yr BP).

Tables

Table 1: Cryolithological units from the bottom to the top with sample depths below surface (m), radiocarbon ages given as years before present (yr BP), calibrated radiocarbon ages (cal kyr BP) and characteristics according to Schirrmeister et al. (2016).

Unit	Depth (m)	Radiocarbon ages (yr BP)	± (yr)	Calibrated ages (2σ) (cal yr BP)	Cryolithological characteristics
1 (18.9–16.0 m)	16.13	52,700	800	not calculable	horizontally bedded grey-brown, silty fine-grained sand; electric conductivity up to 14 mS cm–1 resulting from downward freezing of the ground from the surface
2 (16.0–9.95 m)	13.03	51,200	700	not calculable	horizontally bedded grey-brown, silty fine-grained sand
	11.15	54,100	3,400	not calculable	
	10.53	50,300	2,100	not calculable	
3 (9.95–8.35 m)	8.83	53,500	800	not calculable	inclined contact zone between the ice-wedge and encasing sediment, grey-brown, silty fine-grained sand
4 (8.35–3.4 m)					ice-wedge
5 (3.4–0.35 m)	2.93	11,400	50	13,096-13,304	grey-brown, silty, fine-grained sand
	2.5	11,200	50	12,996-13,199	
	1.7	11,100	50	12,827-13,090	
	0.7	10,149	50	11,603-12,048	
6 (0.35–0.0 m)	0.3	9,700	50	11,075-11,238	active layer, brown, silty fine-grained sand, and modern top
	0.1	modern			

Table 2: Number of sequence counts for bryophyte taxa detected by sedaDNA in the corresponding sample depths.

Depth (m)	Brachytheciaceae	Bryaceae MOTU1	Bryaceae MOTU2	*Bryum* MOTU1	*Cosmarium botrytis*	Dicranales	*Drepanocladus*	*Funaria*	*Hamatocaulis vernicosus*	Hypnales MOTU1	Hypnales MOTU2	Hypnales MOTU3	*Leptobryum pyriforme*	*Platydictya confervoides*	Polytrichaceae	Pottiaceae	*Prychostomum pendulum*	*Sanionia uncinata*	*Sphagnum*	*Sphagnum russowii*
0.35	0	0	0	14	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.60	0	0	0	27	60	0	0	0	0	0	0	12	42	0	0	0	0	0	0	0
2.85	0	0	0	0	0	0	0	0	82	0	0	0	0	0	0	0	0	0	0	0
3.30	0	0	0	0	0	0	0	0	0	0	148	0	0	0	39	0	157	89	3171	3810
8.35	0	12	0	41	0	0	0	0	0	0	0	0	0	0	0	13	0	0	0	0
8.75	0	14	19	203	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8.90	10	0	0	27	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0
8.95	0	0	0	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9.00	0	0	0	17	0	0	15	0	0	236	0	0	0	56	0	0	0	0	0	0
9.60	0	0	0	0	0	0	0	0	0	75	0	0	0	0	0	0	0	0	0	0
10.40	0	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0

Table 3: Number of spores, algae, pre-Quaternary spores and total pollen concentration ~~non-pollen palynomorphs~~ for each sample.

Depth (m)	*Botrychium*	*Osmunda*	Polypodiophyta~~aceae~~	Bryales	*Encalypta*	*Riccia*	*Sphagnum*	*Huperzia selago*	*Lycopodium* sp.	*L. annotinum*-type	*L. clavatum-type*	*L. complanatun*	*L. pungens*	*Selaginella involvens*	*S. rupestris*	*Cercophora*	*Gelasinospora*	*Sordaria*	*Sporormiella*	*Valsaria*	*Glomus*	*Botryococcus*	*Pediastrum*	*Spirogyra*	*Zygnema*-type	pre-Quaternary spores	Total pollen concentration
0.10	0	0	0	0	0	0	6	0	0	0	0	1	1	0	0	0	0	0	0	0	0	2	1	0	1	0	13899
0.35	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	13	0	0	0	0	0	1	0	4	2	28585
0.55	0	0	0	4	0	0	0	0	0	0	0	0	0	0	1	0	38	0	0	0	1	0	0	0	4	0	11992
0.85	0	2	2	0	2	1	0	0	0	0	0	0	0	0	1	0	37	0	0	0	0	1	0	0	1	0	19649
1.10	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	9	0	0	0	0	1	0	0	13	0	20472
1.60	0	0	0	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3	0	0	0	0	38090
1.80	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	48965
2.40	0	0	0	8	0	1	2	1	0	0	0	0	0	0	0	1	5	0	0	0	0	2	0	0	2	0	24976
2.60	0	0	0	0	1	1	1	1	0	0	0	0	0	0	1	0	3	0	0	0	0	0	1	0	1	0	28102
2.85	0	0	0	1	1	1	1	1	0	0	0	0	0	0	1	1	10	0	0	0	0	1	1	0	0	0	23954
8.75	0	0	2	1	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	7	3	0	1	0	8166
8.90	1	0	0	2	1	0	3	0	0	1	0	0	0	0	5	1	0	0	1	0	5	7	5	0	1	1	11294
9.00	0	0	0	1	0	0	0	1	0	1	0	0	0	0	2	0	0	0	0	0	0	3	2	0	6	0	38324
9.70	0	0	0	8	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	2	0	3	0	20328
9.90	0	0	0	25	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	6	0	22580
10.30	1	0	1	0	1	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	6	5	0	4	0	27350
10.40	0	0	1	4	1	0	1	0	0	0	0	1	1	0	1	0	0	0	0	0	0	6	5	0	3	0	14078
10.90	0	0	3	0	1	1	2	1	0	0	0	0	0	0	1	0	0	0	0	0	0	9	8	0	2	0	9203
11.70	0	0	0	4	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2	0	8	0	13912
12.12	0	0	1	3	1	1	1	1	0	0	0	0	0	0	2	0	0	0	1	0	0	2	12	0	2	1	29315
12.25	0	0	2	7	1	0	0	0	1	0	0	0	0	0	4	0	0	0	0	0	0	6	5	0	4	0	11327
12.95	0	1	2	16	0	2	1	0	0	0	0	0	0	0	1	0	0	0	27	0	0	3	49	0	3	0	10792
13.40	0	0	1	10	1	1	3	1	1	0	0	0	1	0	1	0	0	0	2	0	0	5	4	0	4	1	11221
13.95	0	0	1	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	6	0	0	2	0	4170
14.65	1	0	4	0	2	1	2	0	0	0	0	0	1	0	1	0	0	0	0	0	3	8	0	0	2	0	8575
15.20	0	0	1	10	4	1	2	0	0	0	0	0	1	0	1	0	0	4	0	0	2	4	3	0	7	0	10691
16.00	0	1	2	4	1	0	7	0	0	0	0	0	1	3	0	0	0	0	0	0	3	4	1	0	1	0	12148
17.00	0	0	2	9	0	0	6	0	0	0	2	0	1	0	2	0	0	0	9	1	1	7	10	0	0	0	6710
17.05	0	0	5	1	4	0	7	0	0	0	0	0	0	0	1	0	0	0	2	0	7	9	12	0	2	1	4974
17.70	0	0	2	2	2	0	7	0	0	0	0	1	0	0	1	0	0	0	2	0	1	13	16	0	1	3	13223
18.15	1	1	4	0	1	1	6	0	1	0	2	0	1	0	3	0	0	0	7	0	1	17	63	0	3	1	9691
18.48	1	1	4	20	3	0	8	0	1	0	1	0	0	0	1	0	0	0	7	0	12	14	21	0	3	0	3778