Siberian larch forests and the ion content of thaw lakes form a geochemically functional entity

Ulrike Herzschuh[1,2], Luidmila A. Pestryakova[3], Larissa A. Savelieva[4], Liv Heinecke[1], Thomas Böhmer[1], Boris K. Biskaborn[1], Andrei Andreev[5], Arne Ramisch[1], Avery L.C. Shinneman[6] & H. John B. Birks[7,8,9]

Siberian larch forests growing on shallow permafrost soils have not, until now, been considered to be controlling the abiotic and biotic characteristics of the vast number of thaw-lake ecosystems. Here we show, using four independent data sets (a modern data set from 201 lakes from the tundra to taiga, and three lake-core records), that lake-water geochemistry in Yakutia is highly correlated with vegetation. Alkalinity increases with catchment forest density. We postulate that in this arid area, higher evapotranspiration in larch forests compared with that in the tundra vegetation leads to local salt accumulation in soils. Solutes are transported to nearby thaw lakes during rain events and snow melt, but are not fully transported into rivers, because there is no continuous groundwater flow within permafrost soils. This implies that potentially large shifts in the chemical characteristics of aquatic ecosystems to known warming are absent because of the slow response of catchment forests to climate change.

[1] Periglacial Research Section, Alfred Wegener Institute for Polar and Marine Research, Telegrafenberg A43, 14473 Potsdam, Germany. [2] Institute for Earth and Environmental Science, University of Potsdam, Karl- Liebknecht-Strasse 24-25, 14476 Potsdam-Golm, Germany. [3] Department for Geography and Biology, Northeastern Federal University of Yakutsk, Belinskogo 58, 67700 Yakutsk, Russia. [4] Department of Geomorphology, Saint-Petersburg State University, Universitetskaya nab., 7/9, 199034 Saint-Petersburg, Russia. [5] Institute of Geology and Mineralogy, University of Cologne, Zülpicher Strasse 49a, 50674 Köln, Germany. [6] Department of Biology, University of Washington, Seattle, Washington 98195, USA. [7] Department of Biology and Bjerknes Centre for Climate Research, University of Bergen, PO Box 7803, N-5020 Bergen, Norway. [8] Environmental Change Research Centre, University College London, Gower Street, London WC1E 6BT, UK. [9] School of Geography and the Environment, Oxford University, South Parks Road, Oxford OX1 3QY, UK. Correspondence and requests for materials should be addressed to U.H. (email: Ulrike.Herzschuh@awi.de).

In studies on the ecological impacts of global warming, research has concentrated on the role of thaw lakes in the regional organic carbon cycle as sedimentary sinks[1] and as methane sources[2]. Their role in salt and inorganic carbon cycling within the permafrost ecosystems has not, however, been investigated.

Larch forests, the major vegetation cover of Siberia (Fig. 1), are assumed to protect permafrost soils from thawing. Permafrost in turn creates a high water table that allows larch forests to grow, even under strongly arid conditions[3]. It has been ignored that such a coupled system can cause strong epigenetic salt accumulation because of the following mechanisms. First, precipitation, representing the major ion source in central Yakutia, continuously contributes dissolved ion input of about $8–19\ \mathrm{mg\ l^{-1}}$ (mainly SO_4^{2-}, HCO_3^-, Na^+, Ca^{2+} and Mg^{2+})[4]. In addition, some ions are delivered from the soil weathering,
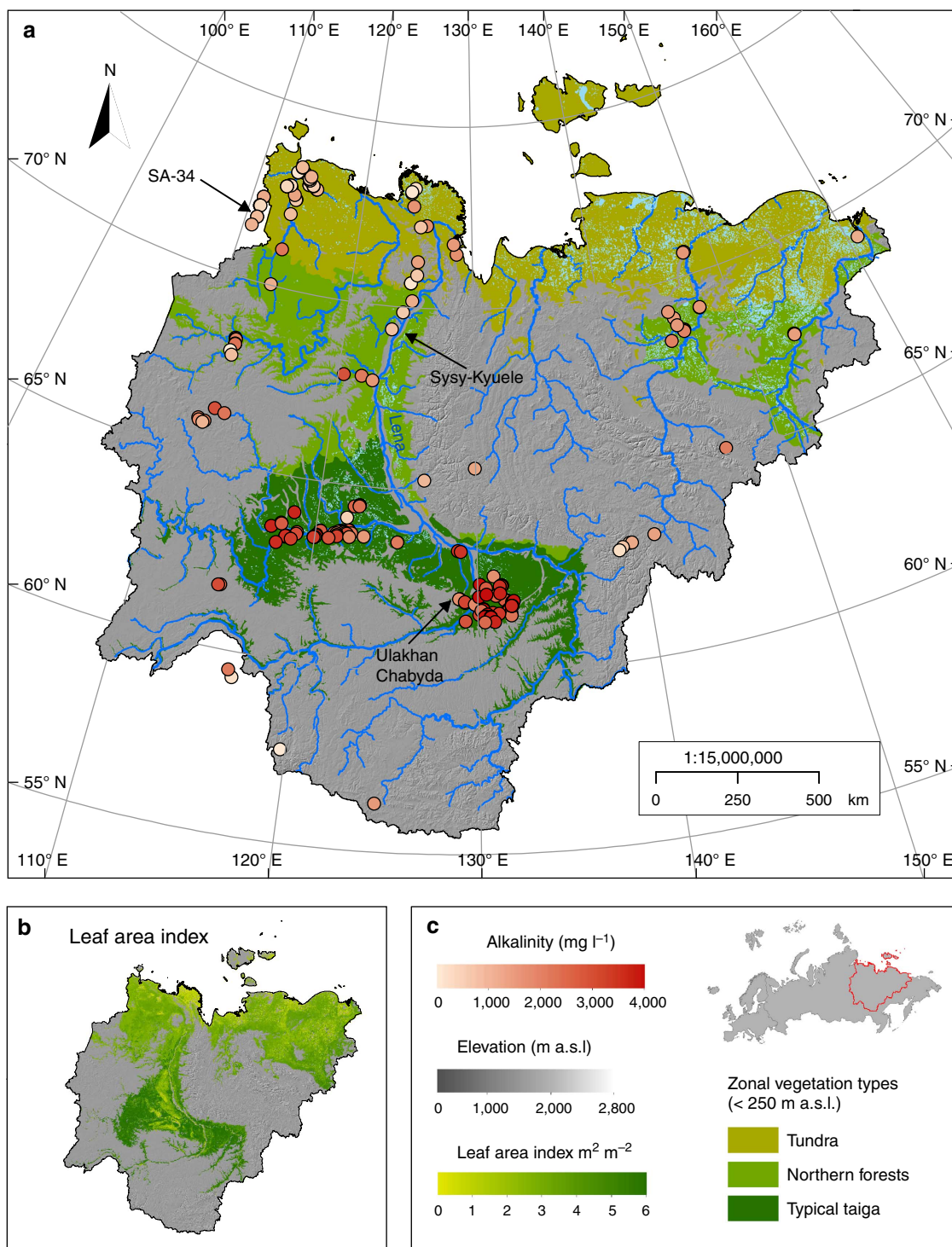
Figure 1 | Study area in Yakutia. (**a**) The distribution of zonal vegetation types in the lowlands (<250 m a.s.l.) of Yakutia; also indicated are the 201 sampling sites and measured alkalinity (HCO_3^- $\mathrm{mg\ l^{-1}}$). Sites of sediment core studies are indicated by arrows. (**b**) Satellite-derived leaf-area index of the vegetation period (mean 10 July 2000–10 August 2009) of the Yakutian lowlands[34]. (**c**) The outline of Yakutia and its location in Eurasia.
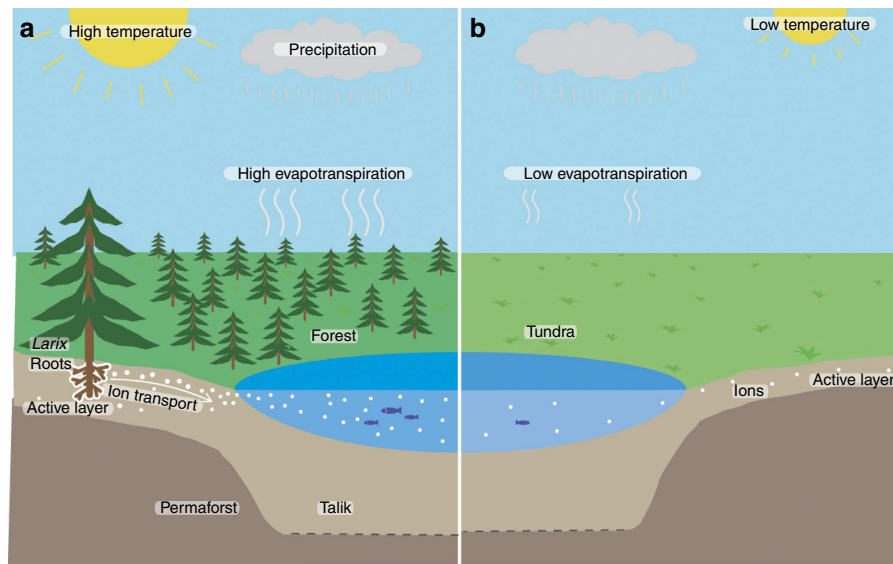
Figure 2 | Conceptual model visualization. We assume that in the forest sites (panel **a**) due to higher evapotranspiration, relative to tundra sites (panel **b**), and a lag in a continuous groundwater flow due to permafrost, ions supplied by precipitation become enriched in nearby thermokarst lakes.

although the rate is low[5]. Second, potential evapotranspiration exceeds precipitation. This particularly applies to the central Yakutian lowlands where there is a two-fold higher potential evaporation relative to the actual precipitation[6]. Third, no continuous groundwater flow exists that connects the remote forest sites with the river systems and, thus, the minerals are not permanently removed from the forest sites. As a result, frozen pale soils with a diagnostic illuvial carbonate and salt horizon have developed, particularly in the lowlands[6]. However, despite aridity, the soluble salt content is rather low (also indicated by the growth of *Larix*, which tolerates only moderate salt levels[7]). These characteristics inhibit the feedback-driven functional chain from salt accumulation, through forest retreat and enhanced soil warming, to permafrost decay. Hence, the question arises what are the mechanisms that result in low salt and carbonate contents in *Larix* forests.

Using four independent data sets, we show that *Larix* forests, permafrost and the degree of ion content of thaw lakes form an interdependent entity (Fig. 2), which can be described by the following conceptual model. *Larix* growth in summer is largely limited by the soil moisture[8]. However, interannual variations in evapotranspiration in *Larix* forests are low, as the trees use meltwater from lower soil depths (supplied as rain during the previous summer)[3,9]. To avoid overconcentration in tissues during water uptake, plants actively leave ions in the rooting zone[10]. Salts are temporally stored in the soil and solutes are washed into thermokarst depressions (alases) with spring meltwater[11] or during wet summers[3,9]. Accordingly, *Larix* forest soils show a much lower salt content than grassland soils in the alases, despite a higher evapotranspiration in forests[9,12]. This implies that epigenetic accumulated salt in permafrost depressions does not only originate locally but is also continuously supplied from the surrounding forests. The strength of long-term salt accumulation in thaw lakes is, together with the frequency of lake drainage or (where occurring) outflow volume, a function of the strength of evapotranspiration in addition to evaporation from the lake surface. The strength of evapotranspiration (which is related to the relative contribution of latent heat in the energy budget) is primarily a function of the photosynthetically active leaf area (Fig. 1) and, therefore, increases markedly across the tundra–taiga

ecotone[13], as confirmed by the land-surface model results for Yakutia[14].

Results

Modern lake chemistry and forest–tundra transition in space. The Yakutian lowlands, where ice-rich permafrost produces strong thermokarst processes[15], have one of the highest thaw-lake densities. According to our 201 lake data set (Fig. 1), most lakes are shallow but the total ion content and alkalinity are highly variable (Table 1). According to our conceptual model, the tundra–forest transition in space today should be paralleled by an increase in ion content in thaw lakes. We find that the present-day lake water chemistry is significantly related to the remote sensing-derived leaf-area index in lake catchments, that is, the leaf-area index uniquely explains more variation in water ion composition than July temperature or lake morphological parameters do in redundancy analyses (RDA, Table 2, Supplementary Fig. S1), the multivariate analogue to multiple linear regression (see Methods). RDA reveals that of the 13.0% of the total explained variation in the modern diatom spectra from these lakes (Table 2); 10.5% is uniquely explained by a representative set of chemical variables (HCO_3^-, Cl^- and Si), whereas the morphological variables explain a much lower percentage (2.4%). Among the chemical variables, alkalinity explains the most variation, indicating its potential for palaeoecological reconstructions (Supplementary Figs S2 and S3, and Supplementary Table S1).

Lake Sysy Kyuele core and Holocene forest–tundra transition. We have investigated the Holocene vegetation and alkalinity changes from a sediment core from the Lake Sysy Kyuele (Figs 1 and 3a) located in the northern Lena River lowlands presently surrounded by sparse northern forests. Pollen data (Fig. 3a) indicate that following late-glacial tundra vegetation, *Betula* expanded massively during the early Holocene. *Larix* appeared about 11,000 calibrated years before present (cal yr BP = 1,950) and its maximum extent was between 9,500 and 5,000 cal yr BP. The slight decrease of the *Larix* cover was paralleled by an expansion of wet understory vegetation as indicated by an increasing Ericaceae and *Sphagnum*. Variation partitioning using RDA reveals that of the 68.9% explained variation in the diatom

Table 1 | Characteristics of sampled lakes and vegetation type for the Yakutian lowlands.

	Tundra*	Northern forest	Typical Taiga	Lowlands†
Number of lakes sampled	61	23	103	187‡
Conductivity range in $\mu S\,cm^{-1}$ of sampled lakes (median)	8–306 (39)	47–632 (199)	31–7,743 (666)	8–7,743 (255)
Alkalinity range in $mg\,l^{-1}$ of sampled lakes (median)	2–134 (16)	18–272 (98)	9–3,721 (281)	2–3,721 (98)
Median depth range (m) of sampled lakes	2.3	3.0	3.1	2.8
Leaf-area index range of catchments in $m^2\,m^{-2}$ (median)[33]	0.33–1.5 (1.1)	1.1–2.1 (1.4)	0.7–4.2 (3.3)	0.6–4.2 (2.2)
Total area below 250 m a.s.l. in Yakutia (km^2)[32]	484,592	301,543	265,043	1,051,178
Number of lakes $>0.1\,km^2$ §	59,673	29,296	22,681	111,650
Sum lake area of lakes $>0.1\,km^2$ §	64,316	29,497	14,076	107,889
Cover% lakes $>0.1\,km^2$ §	13.5	9.8	5.3	10.3
Inorganic carbon in lake water (Pg C) estimated for all lakes $>0.1\,km^2$ §	0.00063	0.00165	0.0018	0.0045

*This also includes forest-tundra.
†Representing 81% of total area below 250 m a.s.l. of 1,291,966 km^2, because azonal vegetation types, for example, along rivers were not considered.
‡In addition, 14 lakes have been sampled from the upland and mountainous areas.
§For lakes $>1\,km^2$ inferred from the database[20] and for lakes $<1\,km^2$ estimated from a power-law function (see Supplementary Methods).

Table 2 | Results of redundancy analyses.

Modern lake ion composition	Explained variation (%)	Modern lake diatom assemblage	Explained variation (%)	Sysy Kyuele fossil diatom assemblages	Explained variation (%)
All env var.	45.2*	All env var.	13.0*	All env var.	68.9*
LAI	8.0*	Chemical var. (HCO_3^-, Cl^-, Si)	10.5*	Pollen (PC 1 and 2)	9.4*
T_{July}	5.3*	Depth and area	2.4*	Insolation (Jan. and July)	5.3*
Depth and area	1.4*			Fluvial (sand %)	1.5*
				Depth (Fe/Mn)	1.4*

chemical var., chemical variables; env var., environmental variables; Jan., January; LAI, leaf-area index; PC, principal component.
Explained variation of compositional data in a series of redundancy analysis by all environmental variables or variable groups together, and each variable or variable group uniquely.
*Significant at the $P<0.05$ level.

data (Table 2), vegetation (scores of principal component analysis (PCA) axes 1 and 2 based on fossil pollen data) uniquely explains a higher amount of variation than insolation (January and July for 69°N), fluvial activity (indicated by sand content), or lake depth (assessed by Fe/Mn ratio). However, most variation is explained by a combination of several environmental variables, indicating a strong coupling of the various components in the system. Reconstructed values of HCO_3^- (based on a newly developed regional diatom–HCO_3^- transfer function, see Methods) are about $100\,mg\,l^{-1}$ during the early Holocene but decrease markedly later with a particularly strong decrease $\sim5,500$ cal yr BP to the present values of about $15\,mg\,l^{-1}$. The phase of *Larix* and *Betula* maximum cover is characterized by the highest lake alkalinity, which, according to our conceptual model (Fig. 2), originates from higher evapotranspiration rates and enhanced ion import to the lake from surrounding soils and reduced run-off.

Lake Ulakhan Chabyda core and transitions among forest types. Evapotranspiration differs markedly among different forest types. Comparisons among *Larix*- and *Pinus*-dominated stands show that in accordance with the higher leaf-area index of *Larix* compared with that of the *Pinus*, the sapflow rate of single *Larix* trees is higher[16]. We assume that differences in mean ion content of lakes from different forest types (from low to high—*Pinus*, *Larix* and *Betula,* Supplementary Fig. S1), at least partly, originate from differences in evapotranspiration rates, as confirmed by differences in soil carbonate and soluble salt content among sites[11]. Our sediment analyses at Ulakhan Chabyda in the central Yakutian lowlands support this hypothesis. In its surroundings,

Larix was probably dominant and *Betula* obtained high cover from around 9,500–5,500 cal yr BP (Fig. 3b). Invasion of *Pinus* started from about 7,000 cal yr BP, and from about 6,000 cal yr the lake's surroundings were covered by *Pinus–Larix* forests similar to the present day. Modern analogue matching for fossil Ulakhan Chabyda samples shows that good analogues are available from the central Yakutian *Larix* forest sites for the period since $\sim6,000$ cal yr BP when modern conditions became established around Ulakhan Chabyda (Supplementary Fig. S4). However, the closest modern analogues for the earlier period occur in the present-day central Mongolian *Larix–Betula* forest zone[17]. Diatom-based reconstructions show that HCO_3^- concentrations temporarily reached $900\,mg\,l^{-1}$ during the early Holocene, with a marked decrease from 7,100 to 4,500 cal yr BP to $\sim150\,mg\,l^{-1}$ today. Accordingly, we assume that the much higher abundance of *Betula* during the early-to-mid Holocene (also seen at Sysy Kyuele), and related higher evapotranspiration, contributed to the higher lake ion content. At Ulakhan Chabyda, this trend was enhanced by the mid-Holocene *Pinus* invasion on sandy soils characterized by a low permafrost table that was probably formerly covered by *Betula*.

Time scales of change. To test whether the millennial-scale processes mirrored by the records described above apply to centennial to decadal time scales, which would enhance the implications of our conceptual model for future ecosystem changes, we investigated a short core from a small lake in the Popigai catchment, in the area between northwestern Yakutia and Krasnoyarski Krai, presently surrounded by the northern *Larix* forests (Fig. 1). A comparison of the pollen and diatom results
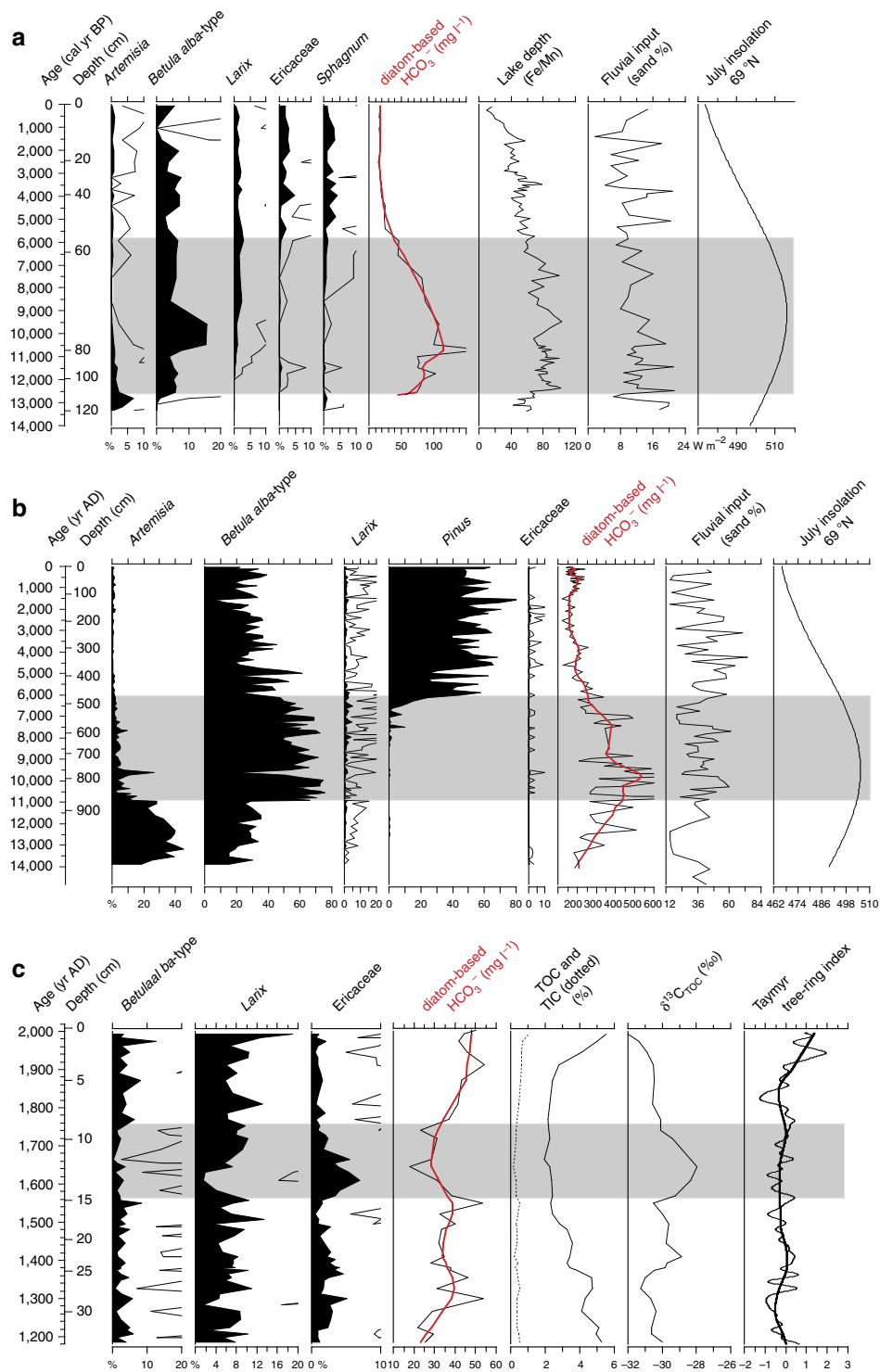
Figure 3 | Results of sediment investigations from three sites in northern Siberia. (a) Sysy Kyuele, lower Lena river lowlands covered by open northern *Larix* forests. **(b)** Ulakhan Chabyda, central Yakutian lowlands in the typical *Larix* taiga zone. **(c)** Lake SA-34, upper Popigai River on the border between Yakutia and Krasnoyarski Krai, surrounded by northern *Larix* forests. In the interpretation of our pollen data, we assume that *Artemisia* indicates openness of vegetation, *Betula alba*-type and *Larix* indicate forest extent and composition, and Ericaceae and *Sphagnum* indicate the extent of wetland vegetation either in the form of mire patches or in the forest understory. Changes in lake depth and fluvial activity are inferred from the variation in Fe/Mn and sand%. Alkalinity reconstructions are presented in mg HCO_3^- l^{-1}. It is based on the application of a modern transfer function (model validation by jackknifing yields a root-mean square error of prediction (RMSEP) of $0.3\log + 1$ mg l^{-1}) to fossil diatom assemblages. Major trends are summarized by a LOESS (locally estimated scatterplot smoothing smoother) (span: 0.25). For Lake SA-34, the reconstructed alkalinity changes were validated by changes in the inorganic carbon supply as indicated by $\delta^{13}C$ of organic carbon. July insolation (Fig. 2a,b)[44] or Taymyr tree-ring width chronology (Fig. 2c)[19] are plotted as independent records of summer temperature changes. Chronologies developed on the basis of radiocarbon dates justify the interpretation of millennial-scale changes at Sysy Kyuele and Ulakhan Chabyda. For Lake SA-34, an age-depth model based on $^{210}Pb/^{137}Cs$ and radiocarbon dates provides a reliable basis to interpret centennial-scale changes. Highlighted in gray are phases of specific interest discussed in the text.

reveals a significant correlation between changes in vegetation and phytoplankton (Fig. 3c). Expansion of Ericaceae in the forest understory of the northern *Larix* forests during the seventeenth century, indicating moister soil conditions, is paralleled by decreasing lake alkalinity as indicated by very low diatom-based HCO_3^- concentrations, low total inorganic carbon (TIC) values and high $\delta^{13}C$ of total organic carbon (TOC), the latter originating from a limitation in carbon supply in the water column[18]. When compared with a regional tree-ring chronology[19], this minimum coincides with below-average ring widths, potentially originating from cold summers on the onset of the Little Ice Age. The twentieth century, in contrast, was characterized by high summer temperatures and high *Larix* percentages, and was compared with the seventeenth century average, a doubling of lake HCO_3^- concentrations.

Discussion

Our results have several implications that are not yet considered with respect to the status of larch taiga ecosystems (and of northern forests in general) in response to future global warming.

According to our modern analyses, thaw lakes represent a major reservoir of inorganic carbon. GIS analyses (Supplementary Methods) reveal that about 40% of the Yakutia ($\sim$size of continental Europe) are lowlands (Fig. 1) typically covered by ice-rich permafrost sediments[15]. According to the global lake database[20] for lakes $>1\,km^2$ and that predicted from a power-law fitted to lakes between 1 and $0.1\,km^2$, the tundra, northern forests and typical taiga lowlands cover $\sim$13%, 10% and 5%, respectively, of the catchments of lakes with an area $>0.1\,km^2$. If we assume these lakes are, respectively, 3.1, 3.0 and $2.3\,m$ deep (median depths in our database), the present-day dissolved inorganic carbon content approximately sums to $0.004\,Pg\,C$ when we apply the observed median HCO_3^- concentrations of 16, 98 and $278\,mg\,l^{-1}$ for the tundra, northern forests and taiga lakes. If we assume that the lake area and depth and the lake-to-latitude distribution during the early or mid-Holocene were similar to today (which represents a rather conservative estimate, as there is evidence that most lakes were deeper and also that thermokarst activity was higher[21]), and if we assume that the inorganic carbon load in these lakes has increased fivefold (as indicated by changes in our two investigated lakes and in another lake from the Lena Delta[22]), an extra $\sim$$0.02\,Pg\,C$ is stored in the inorganic carbon reservoir of Yakutian permafrost lakes. In contrast to the rather fixed organic carbon reservoir in permafrost, this fraction is rapidly exchangeable in a climate-driven vegetation change.

If terrestrial vegetation is a more direct driver of aquatic ecosystems than thermal regime, this has important implications for aquatic communities under rapid warming. In contrast to an expected immediate response to warming[23], the time-lagged response of vegetation to climate will result in a climate–aquatic ecosystem disequilibrium. In particular, high-latitude forest establishment is assumed to lag significantly behind a temperature rise[24], which might result in the formation of new 'interim' or novel terrestrial ecosystems dominated by shrubby vegetation[25]. The implication for aquatic environments in the futures is, however, unknown.

Even now, freshwater availability is a matter of concern in central Yakutia[26], a problem that will be of particular relevance to areas farther north under future warming and related vegetation transition[9].

In general, our conceptual model (Fig. 2) contradicts earlier assumptions that with forest expansion during the Holocene, soil acidity and, thus, humic-acid export to lakes increases[27]. Our results support inferences from other studies[28] that processes other than tree immigration into catchments, such as mire expansion or fire, were the drivers of observed lake acidification.

The question arises whether the functional coupling between lake ion content and forest density, which we propose in Yakutia is a circum-arctic phenomenon. Treeline transect studies of permafrost-influenced lakes in Alaska and the Northwestern Territories of Canada[29–31] likewise show that ion content and/or alkalinity increase from tundra to forest sites. However, the observed gradients are much smaller than those in Yakutia. These patterns have not been interpreted in terms of a causal relationship with vegetation evapotranspiration rates. They have been tentatively interpreted as a result of higher catchment weathering and organic matter decomposition rates, and increased precipitation[29–31]. Because of the limited comparability of the physical and ecological conditions in eastern Siberia and northwestern North America (for example, *Picea* and *Abies* instead of *Larix* as dominant trees and no continuous permafrost in forested areas in the North American Arctic), it remains speculative whether our proposed functional entity between forests and thaw-lake ion contents is applicable to other arctic sites.

In conclusion, we propose that interactions between permafrost, larch forest and thaw lakes represent a missing link in understanding the mineral and inorganic carbon cycling in the lowlands of Siberian taiga forests and that thaw lakes represent a major and, on millennial to decadal scales, highly variable carbon reservoir.

Methods

Modern sediment material. We collected water and surface sediments from 201 lakes scattered throughout Yakutia (Fig. 1) along the whole latitudinal vegetation gradient[32]: 61 lakes are located in the arctic tundra (21 in the forest–tundra zone), 23 lakes in the northern *Larix* forest and 103 lakes in a typical taiga zone (15 in *Pinus–Larix* forests, 13 in *Betula–Larix* forests and 75 in *Larix* forests). In addition, 14 lakes from mountain forests were sampled. Samples from lakes surrounded by the azonal river valley vegetation were not included. Lakes were sampled in summer 2000–2009. Surface sediments were obtained from the deepest part of each lake using a sediment grab or a gravity corer. The deepest part of each lake was located using a portable echo sounder.

Environmental parameters of modern sampling sites. Limnological parameters, such as pH, electrical conductivity, lake area, maximum lake depth and Secchi depth, were measured in the field. Water samples were collected from 0.5 m below the surface of each lake, immediately before the retrieval of the sediment samples. The water samples were kept cool during field work and were analysed immediately on return. For ion analysis, water was passed through a cellulose-acetate filter. Samples for cation analysis were then acidified with HNO_3, whereas samples for anion analysis and residue samples were stored cool. The cation content of the water was analysed by inductively coupled plasma optical emission spectrometry (Perkin-Elmer Optima 3000 XL), whereas the anion content was determined by ion chromatography (Dionex DX-320). Alkalinity was measured using a Metrohm Basic Titrino 794 and the results were evaluated against those obtained in the field using the Alkalinity AL7 titration test kit of Macherey-Nagel. Ion balance was calculated for each sample to validate the analytical methods, resulting in deviations of less than $\pm\,5\%$ for most samples. We are aware that PO_4^{3-} concentrations might have altered significantly during the period between sampling and analysis, but we include these data in the numerical analysis, as they provide general information on important nutrient concentrations.

We used the $1\,km^2$-MODIS data sets from the ESA DUE Permafrost project[33] to calculate the leaf-area index for our studied sites using the 10-year (2000–2009) mean (excluding all zero values) of the main vegetation period (weekly, 07/10–8/10) for a circular buffer zone of 30 km diameter. Finally, the square-root of the median of all cells was used for further analyses. All this work was done with ArcGIS 10. Mean July temperatures were extracted from a gridded global data set[34].

Diatom analyses. Diatom analyses were conducted on about 0.5 g of sediment. The calcareous and organic components were removed by heating with HCl (10%) and H_2O_2 (30%). Cleaned diatom samples were mounted on microscope slides with Naphrax. About 500 diatom valves per slide were counted along randomly selected transects. The samples were analysed using a Zeiss microscope equipped with a differential interference contrast at a magnification of $\times\,1,000$. Water and diatom data of the 103 lakes from Central Yakutia are published by Pestryakova *et al.*[35]

Sediment core material. A core of 121 cm length was collected from Sysy Kyuele (69.404°N; 123.828°E; 81 m a.s.l.) in summer 2009. The sediment consists of a sequence from greyish silt to brown organic-rich gyttja. The age-depth model, sedimentological data and the diatom data are published by Biskaborn et al.[36] The 1,060-cm long sediment core from Ulakhan Chabyda (61.98°N; 129.39°E, 218 m a.s.l.) consists of silty lake mud. Pollen and diatom data are published by Andreev et al.[37] and Pestryakova et al.[35], respectively.

A 34-cm length sediment core was collected from lake SA-34 (71.501°N and 110.900°E, 85 m a.s.l.), using a gravity corer. Granulometric and organic carbon analyses (not shown) indicate that sediments consist of silty lake mud without major changes. ^{210}Pb measurements indicate that the upper 5.7 cm accumulated after 1860 AD and within error ranges the 1963-AD-^{137}Cs-maximum verifies this chronology (Supplementary Figs S5, S6, S7, Supplementary Tables S2, S3, S4 and Supplementary Note 1). The chronology of the lower part of the core is based on the age-depth model using radiocarbon dates. We used the BACON software[38] on twelve ^{14}C dates to construct a robust age-depth model (for details, see Supplementary Note 1). The reliability of the inferred ages from the lower part is supported by an almost similar age-depth model generated from the extrapolation of the mean sedimentation rate inferred from ^{210}Pb dating of the upper part of the core.

Statistical analyses. All lake-water chemical and sedimentary variables were transformed using $\log(x+1)$ (with the exception of pH) and were standardized. Diatom data were expressed as relative abundances. Species were retained in numerical analysis if they occur in at least six modern lake samples with an abundance of $>0.5\%$, or if they occur in at least one lake with an abundance of $>20\%$. The following analyses were performed: principal component analysis (PCA), which was run on the ion composition data to investigate relationships among the different ions. On the basis of the arrow length and position in the ordination biplot, a set of three variables (HCO_3^-, Cl^- and Si) was selected for further analyses to avoid effects from collinearity of the various chemical parameters (a full list of all parameters and the biplot is shown in the Supplementary Fig. S1). RDA, which is a powerful constrained ordination method to detect the explanatory power of a single or a group of predictor variables for a multivariate data set of response variables that have predominantly linear relationships with the predictor variables, was executed. Three sets of RDA were run to investigate the relationships as follows: first, between the modern lake ion composition and environmental variables (leaf-area index, mean July temperature, and lake depth and area); second, between modern diatom data and site-specific environmental characteristics (selected chemical variables and lake depth and area); and third, between fossil diatoms from Sysy Kyuele and independent proxies for vegetation (samples scores on the first two PCA axes of pollen data that were found to be significant when testing against a broken-stick model), lake depth (Fe/Mn ratio) and fluvial activity (sand %). In a first step, all environmental variables were included to infer the total amount of variation that can be explained by the environmental 'predictor' variable set. In a second step, we performed variation partitioning to explore the amount of variation in the compositional data that is uniquely explained by each single environmental variable (by including the other variables as covariables). For the modern diatom data set, additional RDA runs, including each chemical parameter separately, were performed to identify the most relevant environmental variable that could be used for quantitative environmental inferences. Scaling focuses on intervariable correlations in all ordinations. Monte Carlo permutation tests to indicate significance at the $P<0.05$ level were performed under the full model, with 999 permutations. PCA and RDAs were performed using CANOCO version 4.5 (ref. 39), and the broken-stick model[40] was carried out using the biodiversity R package. Quantitative transfer functions for inferring alkalinity (that is, the concentration of HCO_3^-) from fossil diatom assemblages were developed using weighted-averaging partial least-squares regression. Surface samples from sites that were later used for reconstructions were excluded from the calibration set. In addition to our data from Yakutia, a diatom data set of 26 Mongolian lakes[17] was included to improve potentially the reliability of the reconstructions of late glacial–early Holocene conditions at Ulakhan Chabyda where a high abundance of *Cyclotella radiosa* was recorded, which has no modern analogues in the modern Yakutian data set. Weighted-averaging partial least-squares regressions were implemented with C2 software version 1.3 (ref. 41). A one-component model was selected, as it has a high r^2 (0.77) and a low root-mean square error of prediction ($0.30 \log(x+1)$ mg l^{-1}) (Supplementary Fig. S8), and as the performance of a two-component model was not a statistically significant improvement over the one-component model (randomization t-test $P>0.13$). The final transfer functions were applied to fossil diatom spectra from Sysy Kyuele and Ulakhan Chabyda. The reconstructed values of the core tops of $1.2 \log(x+1)$ mg l^{-1} for Sysy Kyuele and $2.3 \log(x+1)$ mg l^{-1} for Ulakhan Chabyda are in reasonably good agreement with the present-day measured HCO_3^- concentrations of 1.1 and $2.4 \log(x+1)$ mg l^{-1}, respectively. As, according to the pollen data, vegetation variability during the last millennium at SA-34 did not cover the whole environmental range of present-day Yakutia, a submodel was used for the reconstruction of lake SA-34's alkalinity, where only sites from northern forests, tundra and mountainous forests (99 lakes) were included, which yielded a slightly lower root-mean square error of prediction of $0.26 \log(x+1)$ mg l^{-1} but markedly improved model predictions at the low mineralized sites. To assess the reliability of the reconstructed values, analogue matching was performed with squared-chord distance (SCD) as the dissimilarity measure[42,43]. Plots of minimum distance versus age indicate (Supplementary Fig. S4) that almost all samples have very good or good analogues in the modern data set (that is, meet the 1% percentile located at 0.56 SCD or the 2.5% percentile at 0.69 SCD, respectively). A few samples from Ulakhan Chabyda around 7,000 cal yr BP are characterized by a high *Navicula diluviana* abundance, which is absent from the modern Yakutian diatom flora and thus have only poor analogues in the modern diatom data set. Furthermore, high *Ellerbeckia arenaria* abundances observed in some of the levels in the lake SA-34 record are rarely found in modern assemblages. However, these samples do not impact the general inference of the alkalinity reconstructions presented here.

References

1. Cole, J. J. et al. Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems* **10**, 171–184 (2007).
2. Walter, K. M., Smith, L. C. & Chapin, III F. S. Methane bubbling from northern lakes: present and future contributions to the global methane budget. *Phil. Trans. R. Soc. A* **365**, 1657–1676 (2007).
3. Ohta, T. In: *Permafrost Ecosystems: Siberian Larch Forests. Ecological studies* Vol. 209 (eds Osawa, A. et al.) 245–269 (Springer, 2010).
4. State Report. Environment and environmental protection activity in Saha Republic (Yakutia) in 2001. State Report of Saha Republic Russian Federation (in Russian) (2003).
5. Desyatkin, R. V., Lesovaya, S. N., Okoneshnikova, M. V. & Zaitseva, T. S. Palevye (pale) soils of Central Yakutia: genetic specificity, properties, and classification. *Eur. Soil Sci.* **44**, 1304–1314 (2011).
6. Chevychelov, A. P. & Bosikov, N. P. in *The Far North: Plant Biodiversity and Ecology of Yakutia, 1, Plant and Vegetation 3* (eds Troeva, E. I. et al.) (Springer, 2010).
7. Carter, M. R. Effects of sulphate and chloride soil salinity on growth and needle composition of Siberian larch. *Can. J. Plant Sci.* **60**, 903–910 (1980).
8. Ochta, T. et al. Interannual variation of water balance and summer evapotranspiration in an eastern Siberian larch forest over a 7-year period (1998–2006). *Agri. Forest Meteorol.* **148**, 1941–1959 (2008).
9. Lopez, C. M. L. et al. Effect of increased rainfall on water dynamics of larch (Larix cajanderi) forest in permafrost regions, Russia: an irrigation experiment. *J. Forest Res.* **15**, 365–373 (2010).
10. Grimaldi, C., Thomas, Z., Fossey, M., Fauvel, Y. & Merot, P. High chloride concentrations in the soil and groundwater under an oak hedge in the west of France: an indicator of evapotranspiration and water movement. *Hydrol. Processes* **23**, 1865–1873 (2009).
11. Lopez, C. M. L. et al. Epigenetic salt accumulation and water movement in the active layer of central Yakutia in eastern Siberia. *Hydrol. Processes* **21**, 103–109 (2007).
12. Lopez, C. M. L. et al. Comparison of carbon and water vapor exchange of forest and grassland in permafrost regions, Central Yakutia, Russia. *Agri. Forest Meteorol.* **148**, 1968–1977 (2008).
13. Behringer, J., Chapin, III F. S., Thompson, C. C. & McGuire, A. D. Surface energy exchanges along a tundra-forest transition and feedbacks to climate. *Agri. Forest Meteorol.* **131**, 143–161 (2005).
14. Park, H., Yamazaki, T., Yamamoto, K. & Ohta, T. Tempo-spatial characteristics of energy budget and evapotranspiration in the eastern Siberia. *Agri. Forest Meteorol.* **148**, 1990–2005 (2008).
15. Czudek, T. & Demek, J. Thermokarst in Siberia and its influence on the development of lowland relief. *Quat. Res.* **1**, 103–120 (1970).
16. Hamada, S. et al. Hydrometrological behaviour of pine and larch forests in eastern Siberia. *Hydrol. Processes* **18**, 23–39 (2004).
17. Shinneman, A. L., Edlund, M. B., Almendinger, J. E. & Soninkhishig, N. Diatoms as indicators of water quality in Western Mongolian lakes: a 54-site calibration set. *J. Paleolimnol.* **42**, 373–389 (2009).
18. Brenner, M., Whitmore, T. J., Curtis, J. H., Hodell, D. A. & Schelske, C. L. Stable isotope (δ^{13}C and δ^{15}N) signatures of sedimented organic matter as indicators of historic lake trophic state. *J. Paleolimnol.* **22**, 205–221 (1999).
19. Briffa, K. R. et al. Trends in recent temperature and radial tree growth spanning 2000 years across northwest Eurasia. *Phil. Trans. R. Soc. B Biol. Sci.* **363**, 2269–2282 (2008).
20. Lehner, B. & Döll, P. Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol.* **296**, 1–22 (2004).
21. Katamura, F., Fukuda, M., Bosikov, N. P. & Desyatki, R. V. Charcoal records from thermokarst deposits in central Yakutia, eastern Siberia: implications for forest fire history and thermokarst development. *Quat. Res.* **71**, 36–40 (2009).
22. Laing, T. E., Rühland, K. M. & Smol, J. P. Past environmental and climatic changes related to tree-line shifts inferred from fossil diatoms from a lake near the Lena River Delta, Siberia. *Holocene* **9**, 547–557 (1999).
23. Smol, J. P. et al. Climate-driven regime shifts in the biological communities of arctic lakes. *Proc. Natl Acad. Sci. USA* **102**, 4397–4402 (2004).
24. Lloyd, A. H. Ecological histories from Alaskan tree lines provide insight into future change. *Ecology* **86**, 1687–1695 (2005).

7

25. Myers-Smith, I. *et al.* Shrub expansion in tundra ecosystems: dynamics, impacts and research priorities. *Environ. Res. Lett.* **6,** 045509 (2011).

26. Savinov, D. D. *et al. Alas Ecosystems: Structure, Function and Dynamics* 264 (Nauka, Novosibirsk, 2005).

27. Seppä, H. & Weckström, J. Holocene vegetation and limnological changes in the Fennoscandian tree-line area as documented by pollen and diatom records from Lake Tsuolbmajavri, Finland. *Ecoscience* **6,** 621–635 (1999).

28. Rosén, P. & Hammarlund, D. Effects of climate, fire, and vegetation development on Holocene changes in total organic carbon concentration in three boreal forest lakes in northern Sweden. *Biogeosciences* **4,** 975–984 (2007).

29. Pienitz, R., Smol, J. P. & Lean, R. S. Physical and chemical limnology of 24 lakes located between Yellowknife and Contwoyto Lake, Northwest Territories (Canada). *Can. J. Fisheries Aqua. Sci.* **54,** 347–358 (1997).

30. Rühland, K. & Smol, J. P. Limnological characteristics of 70 lakes spanning arctic treeline from Coronation Gulf to Great Slave Lake in the Central Northwest Territories, Canada. *Int. Rev. Hydrobiol.* **83,** 183–203 (1998).

31. Gregory-Eaves, I., Smol, J. P., Finney, B. P., Lean, D. R. S. & Edwards, M. E. Characteristics and variation in lakes along a north-south transect in Alaska. *Arch. Hydrobiol.* **147,** 193–223 (2000).

32. Matveev, I. A. (eds) *Agricultural Atlas of the Republic Sakha (Yakutia)* 115 (Nauka, Moscow, 1989).

33. Urban, M., Herold, M., Hese, S., Pöcking, S. & Schmullius, C. *ESA DUE Permafrost - Pan-Arctic Land Cover Product (incl. GlobCarbon LAI and Fire Affected Area Information)* Vienna University of Technology (Institute of Photogrammetry and Remote Sensing, Vienna, 2011).

34. New, M., Lister, D., Hulme, M. & Makin, I. A high-resolution data set of surface climate over global land areas. *Clim. Res.* **21,** 1–25 (2002).

35. Pestryakova, L. A., Herzschuh, U., Wetterich, S. & Ulrich, M. Present-day variability and Holocene dynamics of permafrost-affected lakes in central Yakutia (Eastern Siberia) inferred from diatom records. *Quat. Sci. Rev.* **51,** 56–70 (2012).

36. Biskaborn, B. K., Herzschuh, U., Bolshiyanov, D., Savelieva, L. & Diekmann, B. Environmental variability in northeastern Siberia during the last ~13300 years inferred from lake diatoms and sediment-geochemical parameters. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **329–330,** 22–36 (2012).

37. Andreev, A. A., Klimanov, V. A. & Sulerzhitsky, L. D. Vegetation and climate history of Central Yakutia during the Holocene and Late Pleistocene. *Botan. Z.* **87,** 86–98 (in Russian) (2002).

38. Blaauw, M. & Christen, J. A. Flexible palaeoclimate age-depth models using an Autoregressive Gamma Process. *Bayesian Anal.* **6,** 457–474 (2011).

39. Ter Braak, C. J. F. & Šmilauer, P. *CANOCO Reference Manual and CanoDraw for Windows User's Guide: Software for Canonical Community Ordination (Version 4.5)* (Microcomputer Power, Ithaca, NY, 2002).

40. Legendre, P. & Legendre, L. *Numerical Ecology* 2nd English Edition (Elsevier, 1998).

41. Juggins, S. *User Guide C2—A Program for Analysing and Visualizing Palaeoenvironmental Data* (Department of Geography, University of Newcastle, UK, 2003).

42. Birks, H. J. B. In: *Statistical Modelling of Quaternary Science Data, Technical Guide 5* (eds Maddy, D. & Brew, J. S.) 161–254 (Quaternary Research Association, Cambridge, 1995).

43. Birks, H. J. B. Numerical tools in palaeolimnology—progress, potentialities, and problems. *J. Paleolimnol.* **20,** 307–332 (1998).

44. Laskar, J. *et al.* A long term numerical solution for the insolation quantities of the Earth. *Astron. Astrophys.* **428,** 261–285 (2004).

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Herzschuh, U. *et al.* Siberian larch forests and the ion content of thaw lakes form a geochemically functional entity. *Nat. Commun.* 4:2408 doi: 10.1038/ncomms3408 (2013).